Multiresolution Analysis

by Adrian Mariano
Introduction

A wavelet basis is a basis for $L^2(\mathbb{R})$ generated by a single function $\psi$. All of the elements of the basis are translates or scalings of the original $\psi$ function. Wavelet bases have useful applications because they have both time and frequency localization unlike the ubiquitous Fourier basis. Some of their uses include characterization of the local properties of functions and providing compact approximations of functions.

Preliminaries

Define the Fourier transform to be

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{i\xi x} \, dx.$$ 

Throughout this paper I will omit limits on integrals which are over the reals and I will omit limits on sums which are over all the integers, so

$$\int f(x) \, dx = \int_{-\infty}^{\infty} f(x) \, dx$$

and

$$\sum_{n} f(n) = \sum_{n=-\infty}^{\infty} f(n).$$

We will be working over the vector space of Lebesgue square-integrable functions $L^2(\mathbb{R})$, which I will abbreviate $L^2$. Similarly, $\ell^2$ will be the set of square summable sequences of complex numbers:

$$\{c_n\} \in \ell^2 \iff \sum_{n} |c_n|^2 < \infty.$$ 

Multiresolution Analysis

A multiresolution analysis consists of a sequence of vector spaces $V_i \subset L^2$ and a scaling function $\phi(x)$ which satisfy the following properties.

1. The vector spaces are nested so that

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots$$

and they are scaled copies of each other so

$$f(x) \in V_i \iff f(2x) \in V_{i+1}$$
2. The union of all of the vector spaces gives all the square-integrable functions:
\[ \bigcup_{i} V_i = L^2. \]

3. The intersection of all of the spaces is the zero function:
\[ \bigcap_{i} V_i = \{0\} \]

4. Integer translates of the scaling function form an orthonormal basis for the space \( V_0 \).
\[ \{\sqrt{2^i} \phi(2^ix - k)\} \text{ is an orthonormal basis for } V_i. \]

The last property implies that the \( V_i \) are invariant under translations:
\[ f(x) \in V_i \implies f(x - 2^{-i}k) \in V_i \]
where \( k \) is an integer. Let \( \phi_k = \phi(x - k) \) and let \( \phi_{i,k} = \sqrt{2^i} \phi(2^ix - k) \). In this notation, \( \{\phi_k\} \) is a basis for \( V_0 \) and \( \{\phi_{i,k}\} \) is a basis for \( V_i \).

**Example**

Let \( V_i \) be the set of functions which are constant on each interval \( [2^{-i}n, 2^{-i}(n+1)] \) where \( n \) is an integer. Take for the scaling function (see figure 1)
\[
\phi(x) = \begin{cases} 
1 & x \in [0,1] \\
0 & \text{otherwise}
\end{cases}
\]

To check that this is a multiresolution analysis, we must verify each of the properties:
1. The vector spaces are obviously nested as required. Lower resolution spaces are constant across larger intervals, so they are also constant across smaller ones. The scaling property is obviously satisfied.

2. The set \( \mathcal{U}_i \) is dense in \( L^2 \) so any \( L^2 \) function can be expressed as the limit of functions in \( \mathcal{U}_i \).

3. Suppose \( f \) is in \( \mathcal{V}_i \) for all \( i \). Then \( f \) must be constant on arbitrarily large intervals. The only such function which is also in \( L^2 \) is the zero function, so this property is satisfied.

4. The \( \phi_{i,k} \) are obviously orthonormal and it is obvious that they span \( \mathcal{V}_i \).

All of the properties are satisfied, so this choice of the \( \mathcal{V}_i \) and \( \phi \) is a multiresolution analysis.

**Starting with the scaling function**

It is easier to start with a function and build the vector spaces around it than it is to construct the vector spaces and then find a scaling function. We would like to begin with \( \phi(x) \) and then define a multiresolution analysis as the span of dilations of \( \phi \).

We cannot use an arbitrary \( \phi \). One obvious condition is that the \( \phi_n \) be orthonormal. Fortunately, even if our \( \phi \) of choice does not generate an orthonormal set, there is a trick which fixes this problem.

**Theorem 1 (Characterization of Orthonormal Bases)** The set \( \{ \phi(x-k) : k \in \mathbb{Z} \} \) is orthonormal if and only if

\[
\sum_n |\hat{\phi}(\xi + 2\pi n)|^2 = \frac{1}{2\pi}.
\]

**Proof:** A set \( \{ \phi_k \} \) is orthonormal if and only if for all sequences \( \{c_k\} \in \ell^2 \), we have

\[
\| \sum_k c_k \phi_k \|^2 = \sum_k |c_k|^2.
\]

By using Plancherel's Theorem, we get

\[
\sum_k |c_k|^2 = \| \sum_k c_k \phi_k \|^2 = \| \sum_k c_k \hat{\phi}_k \|^2 = \int \left| \sum_k c_k e^{ik\zeta} \hat{\phi}(\zeta) \right|^2 d\zeta
\]

\[
= \sum_n \int_{2\pi n}^{2\pi(n+1)} \left| \sum_k c_k e^{-ik\zeta} \hat{\phi}(\zeta) \right|^2 d\zeta
\]

\[
= \sum_n \int_0^{2\pi} \left| \sum_k c_k e^{-ik\zeta} \hat{\phi}(\xi + 2\pi n) \right|^2 d\xi.
\]
By Parseval's Theorem,
\[ \sum_k |c_k|^2 = \frac{1}{2\pi} \int_0^{2\pi} \left| \sum_k c_k e^{-ik\xi} \right|^2 d\xi. \]
Therefore,
\[ \frac{1}{2\pi} \int_0^{2\pi} \left| \sum_k c_k e^{-ik\xi} \right|^2 d\xi = \int_0^{2\pi} \left| \sum_k c_k e^{-ik\xi} \right|^2 \sum_n \left| \hat{\phi}(\xi + 2\pi n) \right|^2 d\xi. \]
Since this is true for all choices of \( \{c_k\} \), it follows that the \( \phi_k \) form an orthonormal basis if and only if \( \sum_n \left| \hat{\phi}(\xi + 2\pi n) \right|^2 = \frac{1}{2\pi} \) almost everywhere. \( \blacksquare \)

Using the characterization of orthonormal bases, we can think of taking an arbitrary \( \phi \) and trying to use it to generate an orthonormal basis. As long as
\[ 0 < A < \sum_n \left| \hat{\phi}(\xi + 2\pi n) \right|^2 < B < \infty \]
we can construct a new family of functions by setting
\[ \hat{\phi}^* = \frac{\hat{\phi}}{\sqrt{2\pi \sum_n \left| \hat{\phi}(\xi + 2\pi n) \right|^2}}. \]
This new set of functions is orthonormal:
\[ \sum_n \left| \hat{\phi}^*(\xi + 2\pi n) \right|^2 = \sum_n \left| \frac{\hat{\phi}(\xi + 2\pi n)}{\sqrt{2\pi \sum_k \left| \hat{\phi}(\xi + 2\pi k + 2\pi n) \right|^2}} \right|^2 = \sum_n \left| \frac{\hat{\phi}(\xi + 2\pi n)}{\sqrt{2\pi \sum_k \left| \hat{\phi}(\xi + 2\pi k) \right|^2}} \right|^2 = \frac{1}{2\pi}. \]
Because \( k \) and \( n \) range over all the integers, the inside sum is independent of \( n \), so
\[ \sum_n \left| \frac{\hat{\phi}(\xi + 2\pi n)}{\sqrt{2\pi \sum_k \left| \hat{\phi}(\xi + 2\pi k + 2\pi n) \right|^2}} \right|^2 = \frac{\sum_n \left| \hat{\phi}(\xi + 2\pi n) \right|^2}{2\pi \sum_k \left| \hat{\phi}(\xi + 2\pi k) \right|^2} = \frac{1}{2\pi}. \]
The space spanned by the \( \phi^*_k \) is given by
\[ V_0^* = \left\{ f : f = \sum_n a_n^* \phi_n^* \right\} \]
\[ = \left\{ f : \hat{f} = \alpha_1 \hat{\phi} \text{ where } \alpha_1 \text{ is 2\pi periodic} \right\} \]
\[ = \left\{ f : \hat{f} = \alpha_2 \hat{\phi} \text{ where } \alpha_2 \text{ is 2\pi periodic} \right\} \]
where the last step holds because $\sum_n |\hat{\phi}(\xi + 2\pi n)|^2$ is $2\pi$ periodic. Therefore,

$$V_0^* = \left\{ f : f = \sum_n a_n \phi_n, \text{ with } \{a_n\} \in \ell^2 \right\} = V_0.$$

Now we know that we can begin with an arbitrary $\phi$, produce an new $\phi^*$ which yields an orthonormal basis. Orthonormality alone is not sufficient, however. We must make several other assumptions as well.

**Theorem 2 (Construction of Multiresolution Analysis)** Let $\phi$ be a function such that the following hold:

$$\{\phi_k\} \text{ forms an orthonormal set}, \quad (1)$$

there exists a sequence $\{c_n\} \in \ell^2$ such that

$$\phi(x) = \sum_n c_n \phi(2x - n), \quad (2)$$

$$|\phi(x)| \leq \frac{c}{(1 + |x|)^{1+\alpha}} \quad (3)$$

where $\alpha > 0$ and

$$\hat{\phi}(0) = \frac{1}{\sqrt{2\pi}}. \quad (4)$$

Then the vector spaces

$$V_i = \text{span}\{\phi_{i,k} : k \in \mathbb{Z}\}$$

together with the scaling function $\phi$ form a multiresolution analysis.

**Proof:** We need to check the four properties for a multiresolution analysis. By the orthonormality condition on $\phi$, property 4 of a multiresolution analysis is obviously satisfied. Property 1 is easy to verify. First, we will show that $V_i \subset V_{i+1}$. Let $f \in V_i$. We would like to show that

$$f(x) = \sum_n b_n \phi_{i+1,n}(x).$$

We have

$$f(x) = \sum_n a_n \phi_{i,n}(x)$$

$$= \sqrt{2} \sum_n a_n \phi(2x - n)$$

$$= \sqrt{2} \sum_n a_n \sum_k c_k \phi(2^{i+1}x - 2n - k)$$
(by equation 2)
\[
= \sum_n \sum_k \frac{1}{\sqrt{2}} a_n c_k \phi_{i+1}(x - 2n - k)
\]
\[
= \sum_j b_j \phi_{i+1}(x - j)
\]

where
\[
b_j = \frac{1}{\sqrt{2}} \sum_m f_m c_{j-2m}.
\]

This makes sense because the sequences are square summable and the sums are over all the integers.

We must also check that the vector spaces scale correctly. If \( f(x) \in V_i \) then

\[
f(2x) = \sum_n f_n \phi_{i,n}(2x)
\]
\[
= \sum_n \sqrt{2^i} f_n \phi_{i+1}(2^{i+1}x - n)
\]
\[
= \sum_n \frac{f_n}{\sqrt{2}} \phi_{i+1,n}(x).
\]

Therefore \( f(x) \in V_i \implies f(2x) \in V_{i+1} \). Reading the equations backwards shows implication the other direction, so property 1 for a multiresolution analysis is satisfied.

The two limiting properties are somewhat more difficult to check. Let \( P_j \) be projection onto \( V_j \). We want to show that for any \( f \in L^2 \)

\[
\lim_{j \to +\infty} P_j f = f \quad \text{and} \quad \lim_{j \to -\infty} P_j f = 0.
\]

**Claim:** It suffices to assume that \( f \) is bounded, \( C^1 \) and vanishes outside the interval \([-K, K]\). **Proof of claim:** For any \( f \in L^2 \) and any \( \epsilon > 0 \) there exists a bounded, \( C^1 \) function \( g \) with compact support such that \( \|f - g\| < \epsilon/2 \). Then

\[
\|P_j f\| \leq \|P_j(f - g)\| + \|P_j g\| < \epsilon/2 + \|P_j g\|
\]

so if \( P_j g \to 0 \) as \( j \to -\infty \) then \( \lim_{j \to -\infty} \|P_j f\| < \epsilon/2 \) and \( \epsilon \) was arbitrary, so \( P_j f \to 0 \).

As \( j \to +\infty \), we have

\[
\|P_j f - f\| \leq \|P_j(f - g)\| + \|P_j g - g\| + \|g - f\| < \epsilon + \|P_j g - g\|
\]

and if \( \|P_j g\| \to g \) then we have \( \|P_j f - f\| \to \epsilon \) as \( j \to +\infty \). \( \Box \)
Next, we will examine the projection operation.

\[ P_j f(x) = \sum_n \langle f, \phi_{j,n}(x) \phi(2^j x - n) \rangle \]
\[ = \sum_n \int f(y) 2^j \phi(2^j y - n) \phi(2^j x - n) \, dy \]
\[ = \int 2^j M(2^j x, 2^j y) f(y) \, dy \]

where

\[ M(x, y) = \sum_n \phi(x - n) \phi(y - n). \]

The last step involves switching the summation and the integral. **Claim**: This is permissible because

\[ |M(x, y)| \leq \frac{C_1}{1 + |(x - y)|^{1+\alpha}} \]

which implies absolute convergence. **Proof of claim**: Since \( M(x + m, y + m) = M(x, y) \) when \( m \) is an integer, we can assume that \( x \in [0, 1] \). Then \( |x - n| \approx |n| \) and \( |x - y| \approx |y| \) for large \( y \). Then by condition 3 we have

\[ \sum_n |\phi(x - n) \phi(y - n)| \leq C \sum_{|y - n| < |y|/2} \frac{1}{(1 + |n|)^{1+\alpha} (1 + |y - n|)^{1+\alpha}} \]
\[ + C \sum_{|y - n| \geq |y|/2} \frac{1}{(1 + |n|)^{1+\alpha} (1 + |y - n|)^{1+\alpha}}. \]

Now note that \( |y - n| < |y|/2 \implies |y| - |n| < |y|/2 \implies |y|/2 < |n| \). Therefore, the expression becomes:

\[ |M(x, y)| \leq C \sum_{|y - n| < |y|/2} \frac{1}{(1 + |y|/2)^{1+\alpha} (1 + |y - n|)^{1+\alpha}} \]
\[ + C \sum_{|y - n| \geq |y|/2} \frac{1}{(1 + |n|)^{1+\alpha} (1 + |y|/2)^{1+\alpha}} \]
\[ = \frac{C}{(1 + |y|/2)^{1+\alpha}} \sum_n \left( \frac{1}{(1 + |y - n|)^{1+\alpha}} + \frac{1}{(1 + |n|)^{1+\alpha}} \right). \]

In the last step, we have expanded the limits on both summations to range over all \( n \). The inequality holds because all the terms are positive. The summation is bounded by a constant which is independent of \( y \) since it is a convergent sum.
We can also change the $|y|/2$ into $|y|$, letting the constant take up the difference. This yields
\[ M(x, y) \leq \frac{C'}{1 + |y|^{1+\alpha}} \approx \frac{C'}{1 + |x - y|^{1+\alpha}} \]
which is the desired result. 

Consider what happens as $j \to -\infty$. We have
\[ |P_j f(x)| \leq C_2 \int_{|y| \leq K} \frac{2^j}{(1 + 2^j |x - y|)^{1+\alpha}} \, dy \]
using the fact that $f$ is bounded and compactly supported with $C_2 = C_1 \max(f)$. If $|x| > 2K$ then since $y \in [-K, K]$ we have
\[ \frac{1}{(1 + 2^i |x - y|)^{1+\alpha}} \leq \frac{1}{(1 + 2^i (|x| - K))^{1+\alpha}}. \]
We would like to have the bound
\[ \frac{1}{(1 + 2^i (|x| - K))^{1+\alpha}} \leq \frac{C_3}{(1 + 2^i |x|)^{1+\alpha}} \]
for some $C_3$. This is equivalent to
\[ \left( \frac{1 + 2^i |x|}{1 + 2^i (|x| - K)} \right)^{1+\alpha} \leq C_3. \]
Calculating the maximum of the left hand side for $|x| > 2K$ yields a finite value for $C_3$. Now we have
\[ |P_j f(x)| \leq \int_{|y| < K} \frac{C_2 2^j}{(1 + 2^j |x|)^{1+\alpha}} \, dy \]
\[ = \frac{2K C_2 2^j}{(1 + 2^j |x|)^{1+\alpha}} \]
In the other case, $|x| \leq 2K$ and we also assume $j < 0$. Obviously $1 + 2^j |x - y| \geq 1$. Using the restriction on $x$, we also have
\[ 1 + 2^j |x| \leq 1 + 2^j 2K. \]
Dividing through is okay since the terms are positive, so
\[ \frac{1 + 2^j |x|}{1 + 2^j 2K} \leq 1. \]
Now we know that
\[ 1 + 2^j |x - y| \geq \frac{1 + 2^j |x|}{1 + 2^j 2K} \geq \frac{1 + 2^j |x|}{1 + 2K} \]
8
where the last step makes use of our assumption that \( j < 0 \). From this we can conclude that

\[
\frac{2^j}{(1 + 2^j |x - y|^{1+\alpha})} \leq \frac{2^j (1 + 2K)^{1+\alpha}}{(1 + 2^j |x|^{1+\alpha})}.
\]

Again, we have an expression independent of \( y \), so integration is simply multiplication by \( 2K \) and we have the bound

\[
|P_j f(x)| \leq \int_{|y| < K} \frac{C^{2^j}}{(1 + 2^j |x|)^{1+\alpha}} \, dy = \frac{2KC^{2^j}}{(1 + 2^j |x|)^{1+\alpha}}.
\]

With the bound now valid for all \( x \), we can write

\[
\int |P_j f(x)|^2 \, dx \leq C 2^{2j} \int (1 + 2^j |x|)^{-2-2\alpha} \, dx = C 2^j \int (1 + |y|)^{-2-2\alpha} \, dy.
\]

The integral in the last expression does not depend on \( j \), so we have

\[
\lim_{j \to -\infty} \int |P_j f(x)|^2 \, dx = 0
\]

as desired. This proves property 3.

We have only one property left to show: \( P_j f \to f \) as \( j \to \infty \). Claim: Condition 4 implies that

\[
\int M(x, y) \, dy = 1
\]  \hspace{1cm} (5)

for all values of \( x \). Proof of claim: By the orthonormality of the \( \phi_k \) we have

\[
\sum_k |\hat{\phi}(2\pi k)|^2 = \frac{1}{2\pi}.
\]

Therefore, condition 4 implies that

\[
\hat{\phi}(2\pi k) = 0 \text{ for all integers } k \neq 0.
\]

By the Poisson summation formula,

\[
\sum_n \phi(x - n) = \sqrt{2\pi} \sum_n \hat{\phi}(2\pi n) = \sqrt{2\pi} \hat{\phi}(0) = \int \phi(x) \, dx.
\]
So,

\[
\int M(x, y) \, dy = \sum_n \phi(x - n) \int \phi(y - n) \, dy \\
= \left( \sum_n \phi(x - n) \right) \int \phi(y) \, dy \\
= \int \phi(x) \, dx \int \phi(y) \, dy \\
= |\int \phi(x) \, dx|^2 = 2\pi |\hat{\phi}(0)|^2 = 1.
\]

This proves the claim. \(\square\)

From above, we have for \(j > 0\) and \(|x| > 2K\)

\[
|P_j f(x)| \leq \frac{C2^j}{(1 + 2^j |x|)^{1+\alpha}} \\
\leq \frac{C2^j 2^{\alpha j}}{(1 + 2^j |x|)^{1+\alpha}} \\
\leq \frac{C2^j 2^{\alpha j}}{(2^j |x|)^{1+\alpha}} \\
= \frac{C}{|x|^{1+\alpha}}
\]

and for \(|x| \leq 2K\) we get

\[
|P_j f(x)| \leq C \int_{|y| < K} \frac{2^j}{(1 + 2^j |x - y|)^{1+\alpha}} \, dy \\
\leq \int_{-\infty}^{\infty} \frac{1}{(1 + |z|)^{1+\alpha}} \, dz \\
= C'.
\]

Combining these two results, we have

\[
|P_j f(x)| \leq \frac{C}{(1 + |x|)^{1+\alpha}}
\]

independent of \(j\). Therefore, by the Dominated Convergence Theorem, it suffices to show that \(P_j f(x) \to f(x)\) pointwise. We have

\[
|P_j f(x) - f(x)| = \left| \int 2^j M(2^j x, 2^j y) f(y) \, dy - f(x) \right| \\
= \left| \int 2^j M(2^j x, 2^j y) (f(y) - f(x)) \, dy \right| \\
\text{(by equation 5)}
\]
\[
\leq C \int 2^j (1 + 2^j |x - y|)^{-1-\alpha} |f(x) - f(y)| \, dy \\
\leq C_1 \int_{|x - y| \leq 1} 2^j (1 + 2^j |x - y|)^{-1-\alpha} |x - y| \, dy \\
+ C_2 \int_{|x - y| > 1} 2^j (1 + 2^j |x - y|)^{-1-\alpha} \, dy.
\]

The first part of the last step is justified by the Mean Value Theorem, which states that for some interval \([a, b]\) there exists \(c \in [a, b]\) such that

\[f(b) - f(a) = f'(c)(b - a).\]

We wish to apply the theorem to any interval \([x, y]\) or \([y, x]\) (as appropriate) of length one to get a bound. This works because the derivative of \(f\) is bounded, so we can take the maximum value to get an inequality that holds everywhere.

Now we make the substitution \(w = 2^j (x - y)\) and get

\[
|P_j f(x) - f(x)| = C_1 \int_{|w| \leq 2^j} |w|(1 + |w|)^{-1-\alpha} 2^{-j} \, dw \\
+ C_2 \int_{|w| > 2^j} (1 + |w|)^{-1-\alpha} \, dw \\
\leq C_1 2^{-j} \int_{|w| \leq 2^j} (1 + |w|)^{-\alpha} \, dw + C_2 \frac{2}{\alpha (2^j + 1)^\alpha} \\
\leq C_1 2^{-j} 2^{j(1-\alpha)} + C_2 2^{-j\alpha} \\
\leq C 2^{-j\alpha}.
\]

In the penultimate step, we assumed that \(\alpha < 1\) which is okay because lower values of \(\alpha\) are a weaker restriction on the decay of \(f\). Taking limits, we conclude that

\[\lim_{j \to \infty} |P_j f(x) - f(x)| = 0\]

which proves property 2 for a multiresolution analysis.

**Wavelet Construction**

The multiresolution analysis allows us to approximate a signal at a desired resolution. Projecting an arbitrary function \(f\) from \(L^2\) onto \(V_i\) gives the best approximation (in the least squares sense) of \(f\) in the space \(V_i\). By comparing the projections at two adjacent detail levels, we can get a description of the detail at a particular resolution. It is a description of the information which is lost when we pass from a higher resolution to a lower resolution.
Let \( O_i \) be the orthogonal complement of \( V_i \) in \( V_{i+1} \). Then
\[
V_{i+1} = V_i \oplus O_i
\]
where \( \oplus \) is the direct sum. It follows that
\[
\bigoplus_i O_i = L^2.
\]
If \( f(x) \in O_i \) then \( f(x) \perp V_i \). This implies that \( f(2x) \perp V_{i+1} \). However, \( f(x) \)
is an element of \( V_{i+1} \) so \( f(2x) \in V_{i+2} \). Therefore, \( f(2x) \in O_{i+1} \). A similar argument works in the other direction so we can conclude that the \( O_i \) inherit the scaling properties of the \( V_i \):
\[
f(x) \in O_i \iff f(2x) \in O_{i+1}.
\]
A wavelet basis is a basis for \( L^2 \) generated by taking scalings and dilations of a single function \( \psi \). Given \( \phi \) we would like \( \{ \phi(2^ix - k) : k \text{ and } i \text{ integers} \} \) to be a basis. The main theorem of multiresolution analysis states that there is an orthonormal basis for \( O_i \) and gives a construction for that basis. We can then use that basis to project a function onto \( O_i \). This produces the detail signal. Because of the scaling property of the \( O_i \), this basis can be scaled to all of the \( V_i \) to obtain an orthonormal wavelet basis. Also because of the scaling property, it is sufficient to find the basis for \( O_0 \).

Since \( V_0 \subset V_1 \) it is possible to write \( \phi \) in terms of the basis for \( V_1 \):
\[
\phi = \sum_n h_n \phi_{1,n} \quad \text{where } h_n = \langle \phi, \phi_{1,n} \rangle.
\]
Now set
\[
H(\xi) = \frac{1}{\sqrt{2}} \sum_n h_n e^{-in\xi}
\]
so that
\[
\hat{\phi}(\xi) = H(\xi/2)\hat{\phi}(\xi/2). \tag{6}
\]
Let
\[
\hat{\psi}(\xi) = e^{i\xi/2}H(\xi/2 + \pi)\hat{\phi}(\xi/2). \tag{7}
\]
As in the case of \( \phi \), I will write \( \psi_n(x) = \psi(x - n) \) and \( \psi_{i,n} = \sqrt{2} \psi(2^ix - n) \). With these definitions, we can state the big theorem.

**Theorem 3 (Wavelet Construction)** The set \( \{ \psi_n \} \) is an orthonormal basis for \( O_0 \).

**Proof:** In order to show that \( \psi \) generates a wavelet basis, we must prove three things:
• The $\psi_n$ are orthonormal
• $\psi$ is a member of $O_0$
• The set $\{\psi_n\}$ spans $O_0$

**Orthonormality** It is sufficient to show that $\sum_k |\hat{\psi}(x + 2\pi k)|^2 = \frac{1}{2\pi}$.

\[
\sum_k |\hat{\psi}(x - 2\pi k)|^2 = \sum_k |H(\xi/2 + \pi k + \pi)|^2 |\hat{\phi}(\xi/2 + \pi k)|^2
\]
\[
= \sum_{\text{k odd}} |H(\xi/2 + \pi k + \pi)|^2 |\hat{\phi}(\xi/2 + \pi k)|^2
\]
\[
+ \sum_{\text{k even}} |H(\xi/2 + \pi k + \pi)|^2 |\hat{\phi}(\xi/2 + \pi k)|^2.
\]

Noting that $H$ is periodic with period $2\pi$ we can see that

\[
\sum_k |\hat{\psi}(x - 2\pi k)|^2 = |H(\xi/2)|^2 \sum_n |\hat{\phi}(\xi/2 + \pi + 2\pi n)|^2
\]
\[
+ |H(\xi/2 + \pi)|^2 \sum_n |\hat{\phi}(\xi/2 + 2\pi n)|^2.
\]

Because the $\phi_k$ are orthonormal, both of the summations sum to $\frac{1}{2\pi}$ so we get

\[
\sum_k |\hat{\psi}(x + 2\pi k)|^2 = \frac{1}{2\pi} \left(|H(\xi/2)|^2 + |H(\xi/2 + \pi)|^2\right).
\]

To complete the proof, we simply need to show that

\[
|H(\xi/2)|^2 + |H(\xi/2 + \pi)|^2 = 1.
\]

We begin with

\[
\sum_n |\hat{\phi}(\xi + 2\pi n)|^2 = \frac{1}{2\pi}.
\]

By substituting equation 6 with $\zeta = \xi/2$ we get

\[
\sum_n |H(\zeta + \pi n)|^2 |\hat{\phi}(\zeta + \pi n)|^2 = \frac{1}{2\pi}.
\]

Splitting into even and odd terms again and using the orthonormality of the $\phi_n$, we conclude that

\[
|H(\zeta)|^2 + |H(\zeta + \pi)|^2 = 1.
\]

Therefore, the $\psi_n$ form an orthonormal set.
is in the right space} First show that $\psi \in V_1$. For this to be true, we must have
\[ \psi = \sum_n g_n \phi_{1,n}. \]

This implies
\[ \hat{\psi}(\xi) = \frac{1}{\sqrt{2}} \sum_n g_n e^{-in\xi/2} \hat{\phi}(\xi/2). \]

The term $\sum_n g_n e^{-in\xi/2}$ can represent any $4\pi$ periodic function. Therefore, since $H(\xi/2 + \pi)e^{i\xi/2}$ is $4\pi$ periodic, the equation is satisfied and we can conclude that $\psi \in V_1$.

Next show that $\psi \perp \phi_{0,k}$ for all $k$. This will imply that $\psi \perp V_0$ which combined with $\psi \in V_1$ implies $\psi \in O_0$.

We must show that $\langle \psi, \phi_{0,k} \rangle = 0$ for all $k$. By Plancherel's theorem
\[ \langle \psi, \phi_{0,k} \rangle = \int \hat{\psi}(\xi) e^{-i\xi k} \hat{\phi}(\xi) \, d\xi = \sum_n \int_0^{2\pi} e^{i\xi k} \hat{\psi}(\xi + 2\pi n) \hat{\phi}(\xi + 2\pi n) \, d\xi = \int_0^{2\pi} e^{i\xi k} \sum_n \hat{\psi}(\xi + 2\pi n) \hat{\phi}(\xi + 2\pi n) \, d\xi. \]

Since $e^{i\xi k}$ is a basis for $L^2[0, 2\pi]$ this expression is zero for all $k$ if and only if
\[ \sum_n \hat{\psi}(\xi + 2\pi n) \hat{\phi}(\xi + 2\pi n) = 0. \]

Dividing the sum into even and odd terms, the left hand side becomes
\[ \sum_j \hat{\psi}(\xi + 4\pi j) \hat{\phi}(\xi + 4\pi j) + \sum_j \hat{\psi}(\xi + 4\pi j + 2\pi) \hat{\phi}(\xi + 4\pi j + 2\pi). \]

Now using the definition of $\psi$ and $\hat{\phi}(\xi) = H(\xi/2)\hat{\phi}(\xi/2)$, we get
\[ \sum_j e^{i(\xi + 4\pi j)/2} H(\xi/2 + 2\pi j + \pi) \hat{\phi}(\xi/2 + 2\pi j) H(\xi/2 + 2\pi j) \hat{\phi}(\xi/2 + 2\pi j) \]
\[ + \sum_j e^{i(\xi + 4\pi j + 2\pi)/2} H(\xi/2 + 2\pi j + 2\pi) \hat{\phi}(\xi/2 + 2\pi j + \pi) H(\xi/2 + 2\pi j + \pi) \hat{\phi}(\xi/2 + 2\pi + \pi). \]

Using the periodicity of $H$ and the fact that $\sum_m |\hat{\phi}(\xi + 2\pi m)|^2 = \frac{1}{2\pi}$ we can remove the summations and get
\[ \frac{1}{2\pi} \left( e^{i\xi/2} H(\xi/2 + \pi) H(\xi/2 + \pi) e^{i\pi} e^{i\xi/2} H(\xi/2) H(\xi/2) \right). \]
But \( e^{i\pi} = -1 \) so this equals zero which is what we wanted to show.

We have proven that \( \psi \perp V_0 \) and \( \psi \in V_1 \). Therefore, \( \psi \in O_0 \) as desired.

**The \( \psi_n \) span \( O_0 \)** We want to show that the \( \psi_n \) span \( O_0 \), the orthogonal complement of \( V_0 \) in \( V_1 \). Since \( \phi_n \) span \( V_0 \), this is equivalent to showing that \( \{ \phi_n \} \cup \{ \psi_n \} \) spans \( V_1 \). Suppose \( f \) is a function in \( V_1 \) so
\[
\hat{f}(\xi) = A(\xi/2)\widehat{\phi}(\xi/2) \quad \text{with} \quad A(\xi) = \frac{1}{\sqrt{2}} \sum_n a_n e^{-in\xi}.
\]

Let
\[
B(\xi) = A(\xi/2)\overline{H(\xi/2)} + A(\xi/2 + \pi)\overline{H(\xi/2 + \pi)}
\]
\[
C(\xi) = e^{-i\xi/2} (A(\xi/2)H(\xi/2 + \pi) - A(\xi/2 + \pi)H(\xi/2)).
\]

Then \( B \) and \( C \) are periodic with period \( 2\pi \) since \( A \) and \( H \) are periodic. Therefore, we can write
\[
B(\xi) = \sum_n b_n e^{-in\xi} \quad \text{and} \quad C(\xi) = \sum_n c_n e^{-in\xi}.
\]

By using equation 8 we can get
\[
B(\xi)H(\xi/2) + C(\xi)e^{i\xi/2}H(\xi/2 + \pi) = A(\xi/2)|H(\xi/2)|^2 + A(\xi/2)|H(\xi/2 + \pi)|^2 = A(\xi/2).
\]

Multiplying through by \( \widehat{\phi}(\xi/2) \) and using equations 6 and 7 gives
\[
B(\xi)\widehat{\phi}(\xi) + C(\xi)\hat{\psi}(\xi) = A(\xi/2)\widehat{\phi}(\xi/2) = \hat{f}(\xi)
\]
and inverting the Fourier transform produces
\[
\sum_n b_n \phi_n(x) + \sum_n c_n \psi_n(x) = f(x).
\]

Therefore \( \psi_n(x) \in O_0 \) as desired.

This completes the proof that the \( \psi_{i,n} \) form an orthonormal wavelet basis. ■

Next, we need to invert the Fourier transform to find a direct expression for \( \psi \).
\[
\hat{\psi}(\xi) = \frac{1}{\sqrt{2}} e^{i\xi/2} \sum_k h_k e^{i(k/2 + \pi)x} \hat{\phi}(\xi/2)
\]
\[
= \frac{1}{\sqrt{2}} \sum_k (-1)^k \overline{h_k} e^{i(k+1)/2}\phi(\xi/2).
\]
Inverting the transform,
\[ \psi(x) = \sqrt{2} \sum_k (-1)^k \tilde{h}_k \phi(2x - (-k - 1)). \]

Making the substitution \( n = -k - 1 \), and using the fact that \( n - 1 \) and \( -n - 1 \) have the same parity,
\[ \psi(x) = \sum_n (-1)^{n-1} \tilde{h}_{-n-1} \phi_{1,n}(x). \] (9)

This construction of the wavelet is not unique. It is easy to verify that the theorem is still true if \( \tilde{\psi} \) is multiplied by any \( 2\pi \) periodic function which has absolute value one. Multiplying the Fourier transform by a \( 2\pi \) periodic function allows us to get linear combinations of integer translates of the original \( \psi \). Suppose
\[ \tilde{\psi}^*(\xi) = \alpha(\xi) \tilde{\psi}(\xi) \]
with \( \alpha \) a \( 2\pi \) periodic function and with \( \tilde{\psi} \) defined as in equation 7. Then \( \alpha \) can be written as a Fourier series so that
\[ \tilde{\psi}^*(\xi) = \sum_n a_n e^{i n \xi} \tilde{\psi}(\xi). \]

Inverting the transform, we have
\[ \psi^*(x) = \sum_n a_n \psi(x - n). \]

This flexibility allows the obvious simple translations of the wavelet and the multiplications by a constant of absolute value one. It also permits more exotic adjustments. For example, take \( \alpha(\xi) = e^{i \sin \xi} \). Then the Fourier coefficients are \( a_n = J_n(1) \) where \( J_n \) is the Bessel function of the first kind. (See [4], page 135 for the details.) From this, we see that
\[ \psi^*(x) = \sum_n J_n(1) \psi(x - n) \] (10)
is another choice for the wavelet.

**Example**

Let us return to the scaling function introduced in the first example (see figure 1):
\[ \phi(x) = Z_0(x) = \begin{cases} 1 & x \in [0, 1] \\ 0 & \text{otherwise} \end{cases} \] (11)
To compute the corresponding wavelet, we must use the $h_n$ which are obvious from inspection: $h_{1,0} = 1/\sqrt{2}$ and $h_{1,1} = 1/\sqrt{2}$. Now using equation 9, we get

$$\psi(x) = \phi_{1,0} - \phi_{1,1} = \sqrt{2}(\phi(2x) - \phi(2x - 1)).$$

This is the generator for the well-known Haar basis (see figure 2). If the transformation in equation 10 is applied to this wavelet, then the function displayed in figure 3 results.

The Haar basis is the classical example of a wavelet basis, but it is not even continuous. With a little more work, we can construct multiresolution analyses with $C^k$ scaling functions for any desired $k$.

### Battle-Lemarié Wavelets

The Haar basis can be regarded as merely the first in a sequence of wavelet bases, each one smoother than the last. The basic idea is to use

$$Z_n = Z_0 \ast Z_0 \cdots \ast Z_0$$

$n + 1$ times

where $Z_0$ is defined by equation 11. Then the function $Z_n$ is $C^{n-1}$ and it agrees with a degree $n$ polynomial on each interval of the form $[k, k + 1]$ where $k$ is an integer. These functions are all zero outside the interval $[0, n + 1]$. Except for $Z_0$, they are not orthogonal so they must be orthogonalized.

In the Fourier domain,

$$\hat{Z}_n = \frac{1}{\sqrt{2\pi}} e^{i(n+1)\xi/2} \left( \frac{\sin \xi/2}{\xi/2} \right)^{n+1}.$$

The piecewise quadratic example ($n = 2$) is (see figure 4)
Figure 3: An alternate choice for the Haar wavelet is obtained by using $\psi^*(x) = \sum_n J_n(1)\psi(x - n)$ where $\psi$ is the usual Haar wavelet.

Figure 4: A choice for a piecewise quadratic scaling function
Figure 5: The piecewise quadratic scaling function after being orthogonalized.

Figure 6: The wavelet for the piecewise quadratic scaling function.

\[
\phi = Z_2 = \begin{cases} 
\frac{1}{2} x^2 & 0 \leq x \leq 1 \\
\frac{1}{2} - (x - \frac{3}{2})^2 & 1 \leq x \leq 2 \\
\frac{1}{2} (x - 3)^2 & 2 \leq x \leq 3 \\
0 & \text{otherwise}
\end{cases}
\]

Now \( \phi \) satisfies

\[
\phi(x) = \frac{1}{4} \phi(2x) + \frac{3}{4} \phi(2x - 1) + \frac{3}{4} \phi(2x - 2) + \frac{1}{4} \phi(2x - 3)
\]

so the conditions for the construction are satisfied. The orthogonalized scaling function and corresponding wavelet appear in figures 5 and 6. In general, the orthogonalized scaling function and the wavelet obtained from this construction have infinite support but exponential decay. (See [2] for a proof.) They can be made \( C^k \) for any finite \( k \), but not \( C^\infty \). An example of a \( C^\infty \) wavelet is the Meyer wavelet which is described in [2] (p. 137). The Meyer wavelets decay faster than \( x^{-k} \) for any \( k \), but they do not decay exponentially. Both of these examples are infinitely supported. Smooth compactly supported wavelets are constructed in [2].
Calculation with Filter Banks

Multiresolution Analysis with wavelets was originally proposed by Mallat as a means to analyze real world signals. In such applications, it is important to be able to calculate the various projections efficiently.

If $f$ is a signal, then the signal at the $i$th resolution, $S_if$ is just the projection of $f$ onto $V_i$, so $S_if = P_if$. This is characterized by the coefficients of the scaling function basis. Let $s_i$ be the coefficients so $s_if(k) = \langle f, \phi_{i,k} \rangle$.

Similarly, let the detail signal $D_if$ be the projection of $f$ onto $O_i$. This projection is characterized by the coefficients of its expansion in the wavelet basis, so let $d_i$ be the coefficients so $d_if(k) = \langle f, \psi_{i,k} \rangle$.

Suppose we have a signal $s_if$ and we would like to calculate from it $s_{i-1}f$ and $d_{i-1}f$. Note that if $k < j$ then $S_kS_j = S_k$ so it is reasonable to calculate each signal from the next higher one. To prove that $S_kS_j = S_k$, we begin by noting that

$$f = S_jf + (f - S_jf)$$

where $S_jf \in V_j$ and $(f - S_jf)$ is normal to $V_j$ and hence also normal to $V_k \subset V_j$. Then by applying $S_k$, we have

$$S_kf = S_kS_jf + S_k(f - S_jf).$$

The second term is zero since it is normal to $V_k$, so the desired result follows.

Since $V_i \subset V_{i+1}$, I can write $\phi_{i,k}$ in terms of $\phi_{i+1,k}$:

$$\phi_{i,k} = \sum_j \langle \phi_{i,k}, \phi_{i+1,j} \rangle \phi_{i+1,j}$$

$$\langle \phi_{i,k}, \phi_{i+1,j} \rangle = \sqrt{2^{2i+1}} \int \phi(2^i u - k)\overline{\phi(2^{i+1} u - j)} \, du.$$  

Now let $v = 2^{i+1}u - 2k$. Then $dv = 2^{i+1}du$, and the expression becomes

$$\int \frac{1}{\sqrt{2}} \phi(v + 2k - j)\overline{\phi(v)} \, dv = \int \phi_{-1}(v)\phi(v + 2k - j) \, dv.$$  

So

$$\langle f(u), \phi_{i,n}(u) \rangle = \sum_j \langle \phi(u + 2k - j), \phi_{-1}(u) \rangle \langle f(u), \phi_{i+1,j} \rangle.$$  

With $h(n) = \langle \phi(u + n), \phi_{-1}(u) \rangle$ this becomes

$$s_if(k) = \sum_j h(2k - j)s_{i+1}f(j).$$

Thus $s_if(k) = (h * s_{i+1})(2k)$ where $*$ denotes convolution.
A similar result holds for obtaining \( d_i \). Since \( O_i \in V_{i+1} \), we have

\[
\psi_{i,k} = \sum_j \langle \psi_{i,k}, \phi_{i+1,j} \rangle \phi_{i+1,j}.
\]

Now

\[
\langle \psi_{i,k}, \phi_{i+1,j} \rangle = \sqrt{2^{i+1}} \int \psi(2^i u - k) \phi(2^{i+1} u - j) \, du
\]

\[
= \int \frac{1}{\sqrt{2}} \psi(v + 2k - j) \phi(2v) \, dv
\]

\[
= \int \psi(v + 2k - j) \phi_{-1}(v) \, dv.
\]

With \( g(n) = \langle \psi(u + n), \phi_{-1}(u) \rangle \) we have

\[
d_i \tilde{f}(k) = \sum_j g(2k - j) s_{i+1}(j) = (g * s_{i+1})(2k).
\]

This shows that the discrete signal information \( s_i f \) and \( d_i f \) can be rapidly calculated by a simple algorithm. In practice, we will begin with a real world signal which has finite resolution. If we start with \( s_0 f \), we can repeatedly apply the algorithm to decompose \( s_0 f \) into \( d_0 f, d_{-1} f, d_{-2} f, \ldots, d_{-n}, s_{-n-1} \).

Data Compression

The localization of wavelets in the time domains makes them suitable for lossy compression of signals. When data is compressed so that the original data can be exactly restored, one is using lossless compression. Lossy compression only recovers an approximation to the original data. The general technique is to begin with the function of interest \( f \) and express it as the coefficients \( d_i f \) and \( s_i f \). To reduce the space needed to store the data, the coefficients are truncated in some way. The simplest scheme is to simply discard all of the coefficients below a threshold. We hope to reconstruct a good approximation of \( f \) from the truncated coefficients. Let \( \tilde{f} \) be the approximated function.

Suppose we measure error using the \( L^2 \) norm so we want to minimize

\[
\| f - \tilde{f} \|
\]

subject to some restriction on the size of the compressed representation. If we use the simple method of simply throwing out coefficients, then we will always want to throw out the smaller coefficients first.

Suppose we write \( f = \sum_n f_n \gamma_n \) where \( \gamma_n \) is an orthonormal basis and \( \{ f_n \} \in \ell^2 \). We then construct an approximation \( \tilde{f} = \sum_n \tilde{f}_n \gamma_n \). The error in our
The approximation is then
\[ E^2 = \left\| f - \tilde{f} \right\|^2 = \left\| \sum_n (f_n - \tilde{f}_n) \gamma_n \right\|^2 = \left( \sum_n (f_n - \tilde{f}_n) \gamma_n \right) \left( \sum_n (f_n - \tilde{f}_n) \gamma_n \right) = \sum_n (f_n - \tilde{f}_n)^2 \gamma_n \gamma_n + \sum_n \sum_{m \neq n} (f_n - \tilde{f}_n)(f_m - \tilde{f}_m) \langle \gamma_n, \gamma_m \rangle = \sum_n (f_n - \tilde{f}_n)^2. \]

If we want to minimize the error by dropping \( N \) coefficients, clearly we should drop the \( N \) smallest ones. When this compression technique is applied to image compression using the usual Fourier basis, the results are poor. The difficulty results from the lack of spatial localization of sines and cosines. The errors in reconstructed images appear as textures which cover the whole image. Using a wavelet basis produces much better results. Because the wavelets are spatially localized, the approximated image is free of global artifacts. (See [8] for example images.) Rather than simply discarding coefficients, one can develop more sophisticated algorithms for expressing less important coefficients with coarser quantization.

It is natural to choose \( L^2 \) as the metric for error because of these nice results for \( L^2 \) norms, but we have no reason for believing that the \( L^2 \) norm is the best choice. In a recent paper [3], DeVore, Jawerth and Lucier analyze wavelet approximations with respect to the general \( L^p \) norm. They also argue that the \( L^1 \) norm is the appropriate choice for measuring errors in images intended for human viewing.

Another application of this sort of data compression is in computation. Some operators can be rapidly computed on sparse data sets. By representing the data in the wavelet basis and truncating small coefficients, one may be able to obtain a sparse representation for the data of interest. This has further applications in equation solving because it is easier to solve equations which have sparse operators. (See [1] for the details.)

**Local Properties of Functions**

For many applications, it is useful to detect the singularities in a function. Because of the localization in time and frequency, wavelets are suited to this task. A measurement which is often used to characterize local regularity of a function is the Lipschitz exponent.

Let \( \alpha \) be a positive real number and let \( n \) be the greatest integer such that \( n \leq \alpha \). Then a function \( f(x) \) is said to be Lipschitz \( \alpha \) at \( x_0 \) if and only if there
exist two constants $A$ and $h_0 > 0$ and a polynomial $P_n(x)$ of order $n$ such that for $|h| < h_0$

$$|f(x_0 + h) - P_n(h)| \leq A|h|^\alpha.$$  \hspace{1cm} (12)

The function $f(x)$ is uniformly Lipschitz $\alpha$ over the interval $(a, b)$ if and only if\n
equation 12 is satisfied whenever $x_0 + h \in (a, b)$ where the choice of $A$ is fixed,\n
but the polynomial can depend on $x_0$.

Wavelets can be used to evaluate regularity by using this result. If $f(x) \in L^2$ and\n
$0 < \alpha < 1$ then $f(x)$ is Lipschitz $\alpha$ over $(a, b)$ if and only if for any $\epsilon > 0$\n
there exists a constant $A_\epsilon$ such that for all $x \in (a + \epsilon, b - \epsilon)$ and $s > 0$

$$|f \ast \frac{1}{s} \psi(x/s)| \leq A_\epsilon s^\alpha.$$ \hspace{1cm} (13)

The convolution expression which appears in this bound is generally called the\n
continuous wavelet transform:

$$Wf(s, x) = f \ast \frac{1}{s} \psi(x/s).$$

Its properties are discussed at length in [2]. We will simply note that $Wf(s, x)$ is bounded for $s > s_0$:

$$|Wf(s, x)| = \left| \int f(w) \frac{1}{s} \psi(x-w/s) \, dw \right| \leq \sqrt{\int |f(w)|^2 \, dw} \sqrt{\int \frac{1}{s^2} \left| \psi(x-w/s) \right|^2 \, dw} = \|f\| \cdot \|\frac{1}{s} \psi(x/s)\|.$$ \hspace{1cm} (13)

Therefore, the decay condition amounts to a restriction on the behavior of\n
$Wf(s, x)$ as $s \to 0$. It is reasonable that the local regularity of a function should\n
be related to this, because as $s \to 0$ the dilated wavelet becomes increasingly\n
localized.

This condition can be translated into a restriction on the wavelet coefficients by considering $s = 2^{-j}$ as $j \to \infty$. If $f(x)$ is Lipschitz $\alpha$ at $x_0$ then obviously $f(-x)$ is Lipschitz $\alpha$ at $-x_0$. So the condition for regularity over $[a, b]$ (equation 13) can be written

$$\left| \int f(-w) \frac{1}{s} \psi(x-w/s) \, dw \right| < A_\epsilon s^\alpha.$$ \hspace{1cm} (13)

where $x$ is in $(-b + \epsilon, -a - \epsilon)$. Now making the variable substitutions $s = 2^{-j}$ and $n = 2^j x$, we obtain

$$\left| \int f(-w) 2^j \psi(m - 2^j w) \, dw \right| < A_\epsilon 2^{-j \alpha}$$

where this expression holds for $2^{-j}m \in (-b + \epsilon, -a - \epsilon)$. If we require $m$ to be an\n
integer, then this expression is weaker than the original one. Note, however, that
if we wished to evaluate equation 13 for \( x_0 \), that we can make \( 2^{-j} m \) arbitrarily close to \( -x_0 \) by letting \( j \) go to infinity, and taking appropriate values for \( m \). Therefore, it is not surprising that this weaker result can replace the stronger condition, even though this proof only works in one direction. Returning to the expression, we have

\[
|\int f(-w) 2^{j/2} \psi_{j,-m}(-w) \, dw| = |\int f(y) 2^{j/2} \psi_{j,-m}(y) \, dy| = |2^{j/2} (f, \psi_{j,-m})|.
\]

Finally, we can let \( n = -m \) and conclude that

\[
(f, \psi_{j,n}) < A 2^{-j(\alpha+1/2)}
\]

where \( 2^{-j} n \in (a + c, b - c) \).

By assuming stronger conditions on \( \psi \), it is possible to extend this test for regularity to \( \alpha > 1 \). To avoid messy details, we will prove only a special case of this regularity result. A more general proof can be found in [5].

**Theorem 4** If \( f \) is a bounded \( L^2 \) function, and \( 0 < \alpha < 1 \), then \( f(x) \) is uniformly Lipschitz \( \alpha \) for all \( x \in \mathbb{R} \) if and only if

\[
(f, \psi_{j,n}) < A 2^{-j(\alpha+1/2)}
\]

for all integers \( n \) and \( j \) where \( \psi \) is a \( C^1 \) wavelet constructed through theorems 2 and 3 which satisfies

\[
\psi(x) \leq \frac{B}{(1 + |x|)^{1+\beta}}
\]

and

\[
\psi'(x) \leq \frac{C}{(1 + |x|)^{1+\gamma}}.
\]

**Proof:** Assume that \( f \) is Lipschitz \( \alpha \). This means that

\[
|f(x+y) - f(x)| \leq E|y|^\alpha.
\]

We will show that this implies equation 13, which in turn implies the desired expression. First, we need to show that \( \int \psi(x) \, dx = 0 \). We know that \( \hat{\phi}(0) = \frac{1}{\sqrt{2\pi}} \) because Theorem 2 required this. Now we use equation 6 which says \( \hat{\phi}(\xi) = H(\xi/2) \hat{\phi}(\xi/2) \). Inserting \( \xi = 0 \) into this equation produces

\[
\frac{1}{\sqrt{2\pi}} = H(0) \frac{1}{\sqrt{2\pi}}
\]

which implies \( H(0) = 1 \). From the relation \( |H(\zeta)|^2 + |H(\zeta + \pi)|^2 = 1 \) (equation 8), we see that \( H(\pi) = 0 \). Putting this into the definition of the wavelet \( \psi \)
(equation 7) for \( \xi = 0 \) we obtain \( \hat{\psi}(0) = 0 \) from which we can conclude that \( \psi \) has mean zero.

Returning to the problem at hand,

\[
|f * \frac{1}{x} \psi(x/s)| = \left| \int f(x-y) \frac{1}{x} \psi(y/s) \, dy \right|.
\]

Now \( \int f(x) \frac{1}{x} \psi(y/s) \, dy = 0 \) because \( \psi \) has mean zero and \( f(x) \) is constant with respect to the integration. Therefore,

\[
|f * \frac{1}{x} \psi(x/s)| = \left| \int (f(x-y)-f(x)) \frac{1}{x} \psi(y/s) \, dy \right|
\leq \int C|y|^\alpha \frac{1}{x} \psi(y/s) \, dy
\leq Es^\alpha \int |z|^\alpha |\psi(z)| \, dz.
\]

The integral is constant, so we have the desired bound.

In the other direction, suppose

\[
\{f, \psi_{j,n}\} \leq C2^{-j(\alpha+1/2)}.
\]

If \( y \geq 1 \) then since \( f \) is bounded,

\[
|f(x+y) - f(x)| \leq 2 \sup(y) |y|^\alpha.
\]

Therefore, we only need to consider the case \( y < 1 \). Writing \( f(x+y) - f(x) \) in the wavelet basis gives

\[
f(x+y) - f(x) = \sum_j \sum_n \{f, \psi_{j,n}\} (\psi_{j,n}(x+y) - \psi_{j,n}(x)).
\]

Let \( J \) be an integer chosen so that \( J \approx -\log_2 |y| \). First consider the sum over \( j \geq J \).

\[
\left| \sum_{j \geq J} \sum_n \{f, \psi_{j,n}\} (\psi_{j,n}(x+y) - \psi_{j,n}(x)) \right| \leq \sum_{j \geq J} C2^{-j(\alpha+1/2)} \left( \sum_n |\psi_{j,n}(x+y)| + \sum_n |\psi_{j,n}(x)| \right)
\]

We would like to bound the inner summations. Consider

\[
\sum_n |\psi_{j,n}(x_0)|
\]

where \( x_0 \) is a fixed value. Using the definition of \( \psi_{j,n} \) we have

\[
\sum_n |\psi_{j,n}(x_0)| = 2^{j/2} \sum_n |\psi(2^j x_0 - n)|
\leq 2^{j/2} \sum_n \frac{B}{(1 + |n|)^{1+\beta}}
\]

25
The final summation is constant with respect to $x_0$, so we can write
\[
\left| \sum_{j \geq J} \sum_n (f, \psi_{j,n}) (\psi_{j,n}(x+y) - \psi_{j,n}(x)) \right| \leq D \sum_{j \geq J} 2^{-j(\alpha+1/2)} 2^{j/2} \\
= D \sum_{j=J}^{\infty} 2^{-j\alpha} \\
= D 2^{-J\alpha} \frac{1}{1 - 2^{-\alpha}} \\
\approx \frac{D}{1 - 2^{-\alpha}} |y|^\alpha
\]

This proves the desired bound for the summation over $j \geq J$. Now we must tackle the remaining terms
\[
\left| \sum_{j < J} \sum_n (f, \psi_{j,n}) (\psi_{j,n}(x+y) - \psi_{j,n}(x)) \right| \leq \sum_{j < J} C 2^{-j(\alpha+1/2)} \sum_n |\psi_{j,n}(x+y) - \psi_{j,n}(x)|.
\]
Applying the mean value theorem,
\[
|\psi_{j,n}(x+y) - \psi_{j,n}(x)| = |\psi_{j,n}'(c_n) y| = 2^{3j/2} |\psi'(2^j c_n - n) y|
\]
for some $c_n$ in $[x-y, x+y]$. Returning to the sum over $n$,
\[
\sum_n |\psi_{j,n}(x+y) - \psi_{j,n}(x)| \leq 2^{3j/2} |y| \sum_n |\psi'(2^j c_n - n)|.
\]

Now $c_n$ depends on $n$, but is confined to an interval of length less than 2. Therefore we can use the bound on $\psi'$ to bound the sum:
\[
\sum_n |\psi'(2^j c_n - n)| < \frac{C}{(1 + |n|)^{1+\gamma}}.
\]
As before, the summation is finite, so for the sum over $j < J$ we have
\[
\left| \sum_{j < J} \sum_n (f, \psi_{j,n}) (\psi_{j,n}(x+y) - \psi_{j,n}(x)) \right| < F |y| \sum_{j < J} 2^{-j(\alpha+1/2)} 2^{3j/2} \\
\approx F 2^{-J} \sum_{j = J-1}^{\infty} 2^{-j(\alpha+1/2)} 2^{3j/2} \\
= F 2^{-J} \sum_{j = 1-J}^{\infty} 2^{j(\alpha-1)}.
\]
The summation is convergent since $\alpha < 1$, so
\[
\left| \sum_{j < J} \sum_n (f, \psi_{j,n}) (\psi_{j,n}(x+y) - \psi_{j,n}(x)) \right| < 2^{-J} 2^{(1-J)(\alpha-1)} \frac{F}{1 - 2^{\alpha-1}} \\
= \frac{F 2^{\alpha-1}}{1 - 2^{\alpha-1}} 2^{-J \alpha} \\
\approx C |y|^\alpha
\]
Therefore, combining the results from the two summations over $j$, we conclude that

$$|f(x+y) - f(x)| < C|y|^{\alpha},$$

so $f$ is Lipschitz $\alpha$.

Acknowledgements

Much of the material for this paper was taken from Daubechies book [2] and from discussions with Gerald Folland.

References


