Wavelets

Kurt Bryan

May 19, 2011

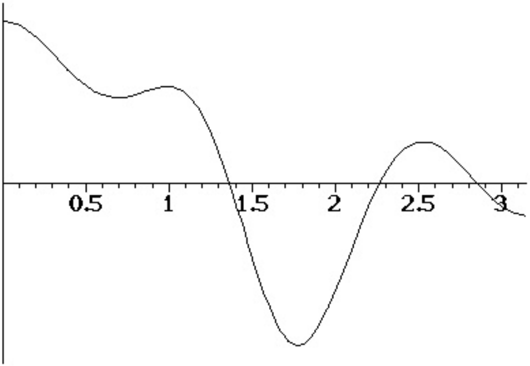
Fourier Cosine Series

Any reasonable function $f(t)$ on $0 \leq t \leq \pi$ can be approximated with a *Fourier cosine series*

$$f(t) \approx a_0 + a_1 \cos(t) + a_2 \cos(2t) + \cdots + a_N \cos(Nt)$$

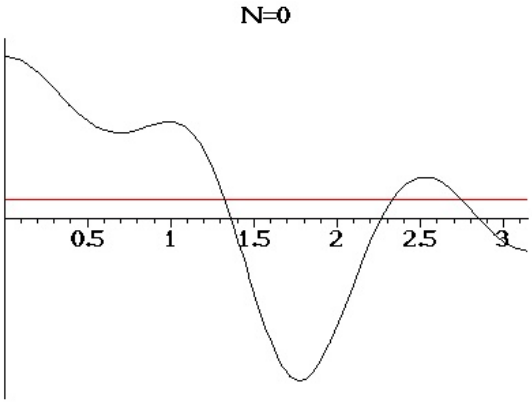
if we pick the a_k correctly (and take N large enough).

A Function to Approximate



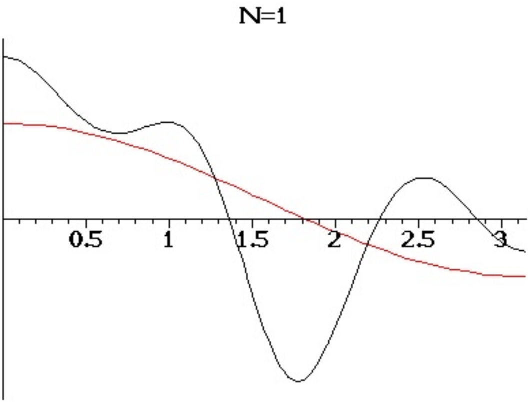
Cosine Series Example

$f(t) \approx 4.70$



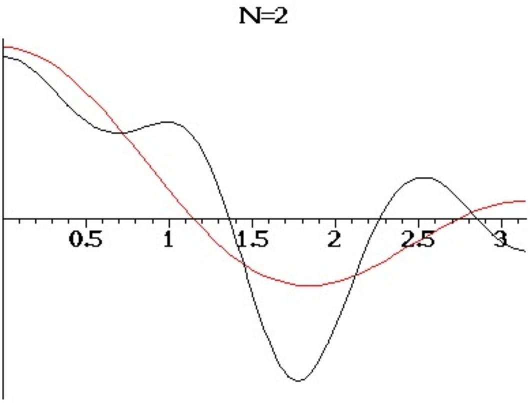
Cosine Series Example

$$f(t) \approx 4.70 + 19.1\cos(t)$$



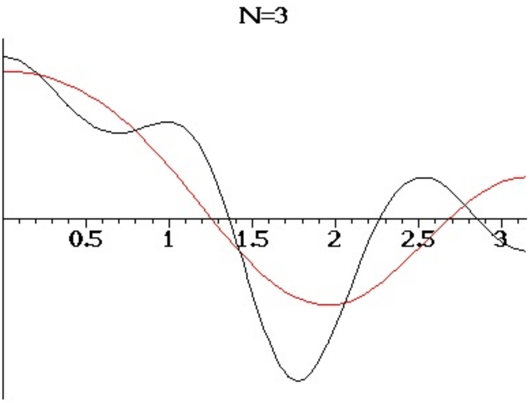
Cosine Series Example

$$f(t) \approx 4.70 + 19.1\cos(t) + 19.0\cos(2t)$$



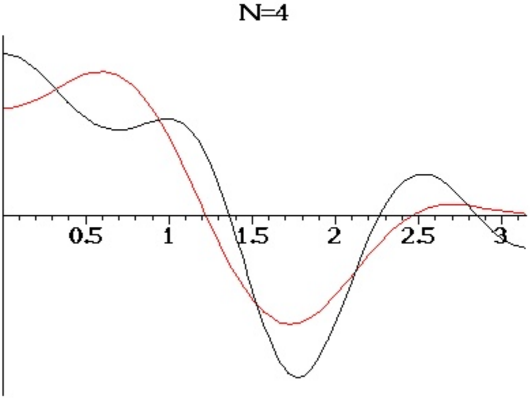
The Cosine Series

$$f(t) \approx 5.97 + 19.1\cos(t) + 19.0\cos(2t) - 5.88\cos(3t)$$



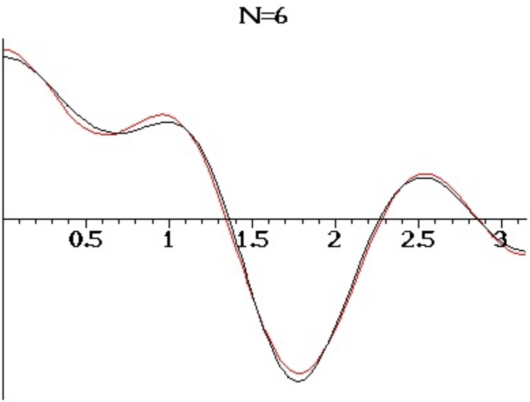
The Cosine Series

$$f(t) \approx 5.97 + 19.1\cos(t) + 19.0\cos(2t) - 5.88\cos(3t) - 9.92\cos(4t)$$



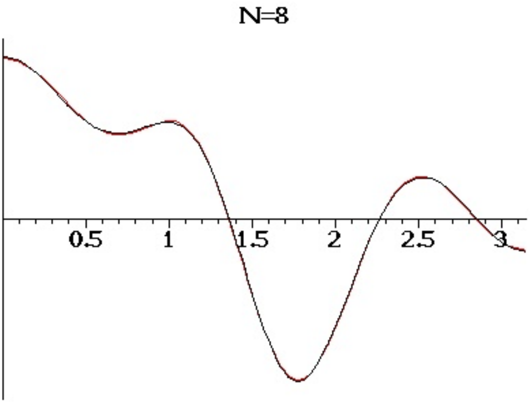
The Cosine Series

$$+\cdots+12.4\cos(5t)+2.97\cos(6t)$$



The Cosine Series

$$+\cdots - 1.70\cos(7t) - 0.53\cos(8t)$$



The Cosine Coefficients

Any "nice" function $f(t)$ defined on $[0, \pi]$ can be approximated

$$f(t) \approx \frac{a_0}{2} + a_1 \cos(t) + a_2 \cos(2t) + \cdots + a_N \cos(Nt)$$

where

$$a_k = \frac{2}{\pi} \int_0^{\pi} f(t) \cos(kt)\, dt$$

The more terms you take, the better it gets.

General Theory

Suppose $\phi_0(t), \phi_1(t), \phi_2(t), \ldots$ are a family of functions on interval $[a, b]$ such that any reasonable $f(t)$ can be written

$$f(t) = c_0\phi_0(t) + c_1\phi_1(t) + c_2\phi_2(t) + \cdots$$

General Theory

Suppose $\phi_0(t), \phi_1(t), \phi_2(t), \ldots$ are a family of functions on interval $[a, b]$ such that any reasonable $f(t)$ can be written

$$f(t) = c_0\phi_0(t) + c_1\phi_1(t) + c_2\phi_2(t) + \cdots$$

Suppose also that the family is orthogonal, i.e., the inner product

$$(\phi_j, \phi_k) := \int_a^b \phi_j(t)\phi_k(t)\, dt$$

is zero when $j \neq k$. Then

General Theory

To find the coefficients c_k, start with

$$f = c_0\phi_0 + c_1\phi_1 + c_2\phi_2 + \cdots$$

General Theory

To find the coefficients c_k, start with

$$f = c_0\phi_0 + c_1\phi_1 + c_2\phi_2 + \cdots$$

Take the inner product of each side with ϕ_k:

$$(f, \phi_k) = c_0(\phi_0, \phi_k) + c_1(\phi_1, \phi_k) + c_2(\phi_2, \phi_k) + \cdots$$

General Theory

To find the coefficients c_k, start with

$$f = c_0\phi_0 + c_1\phi_1 + c_2\phi_2 + \cdots$$

Take the inner product of each side with ϕ_k:

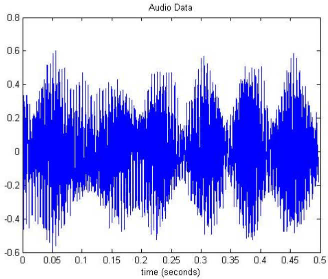
$$(f, \phi_k) = c_0(\phi_0, \phi_k) + c_1(\phi_1, \phi_k) + c_2(\phi_2, \phi_k) + \cdots$$

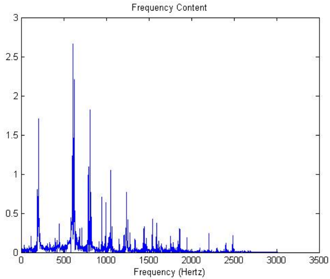
All the inner products on the right are zero except for $c_k(\phi_k, \phi_k)$ which leads to $(f, \phi_k) = c_k(\phi_k, \phi_k)$, so

$$c_k = (f, \phi_k)/(\phi_k, \phi_k).$$

Graphical Fourier Analysis

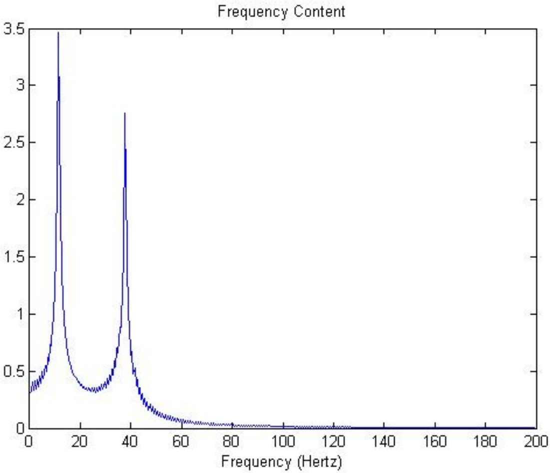
Audio signal and Fourier cosine coefficient magnitudes:





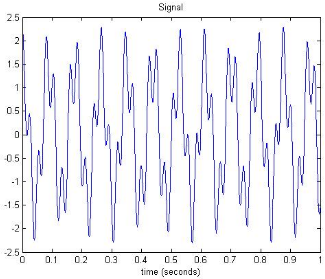
Fourier Shortcomings

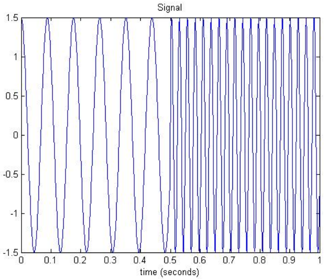
Here's a plot of the Fourier cosine coefficients for some signal:



Fourier Shortcomings

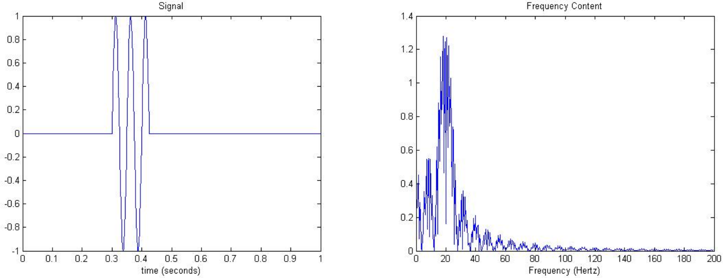
Which signal was it?





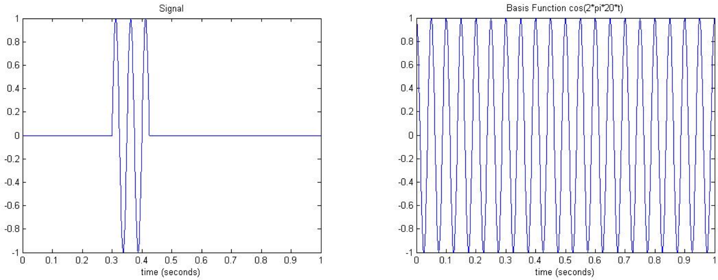
Fourier Shortcomings

The problem: a short stretch of signal at frequency "k" ANYWHERE in the signal excites the corresponding Fourier frequency.



Fourier Shortcomings

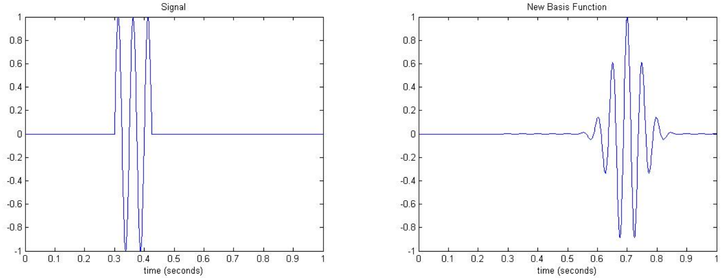
The basis function overlaps the short signal, no matter where the signal is supported.



The integral $\int_0^1 f(t) \cos(2\pi(20)t)\, dt$ doesn't much depend on the location of f.

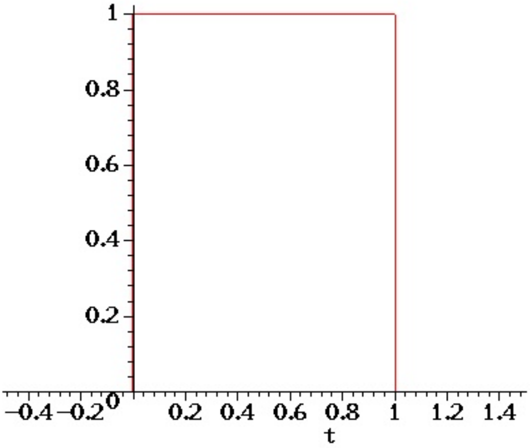
Fourier Shortcomings

What we'd really like is to replace "globally supported" cosines with something that has small support (but still encodes frequency information):



Haar Scaling Function

The Haar scaling function $\phi_0(t)$ (on $[0, 1]$) looks like



Level 0 Approximation

A typical function $f(t)$ can be approximated as

$$f(t) \approx c_0 \phi_0(t)$$

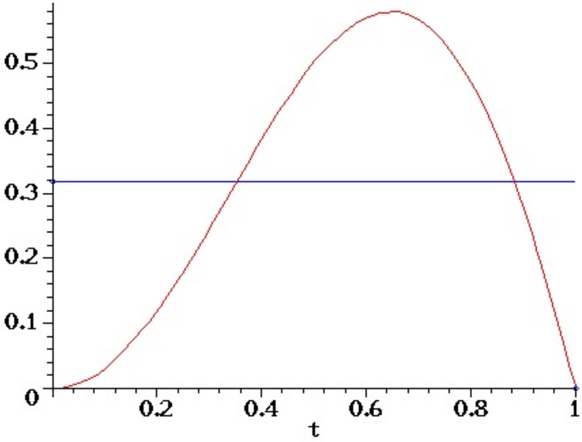
with

$$c_0 = \frac{(f, \phi_0)}{(\phi_0, \phi_0)} = \int_0^1 f(t)\,dt.$$

That is, c_0 is just the average value of f.

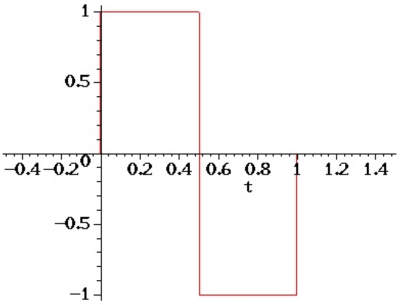
Level 0 Approximation

The result:



Mother Haar Wavelet

The mother Haar wavelet is the function $\psi_0(t)$



Note $(\phi_0, \psi_0) = 0$.

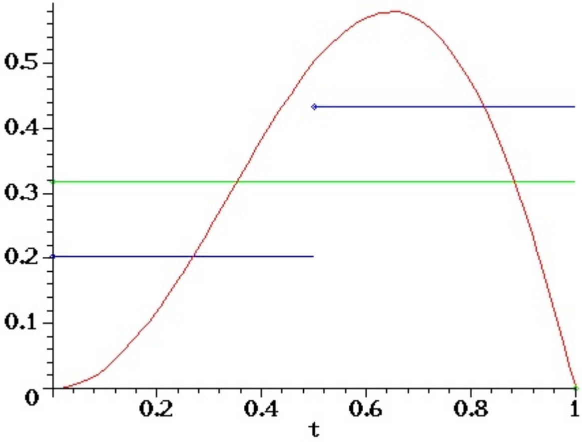
Level 1 Approximation

We can approximate $f(t) = c_0\phi_0(t) + d_0\psi_0(t)$ with c_0 as before and

$$d_0 = \frac{(f, \psi_0)}{(\psi_0, \psi_0)} = \int_0^1 f(t)\psi_0(t)\,dt$$
$$= \int_0^{1/2} f(t)\,dt - \int_{1/2}^1 f(t)\,dt$$

Level 1 Approximation

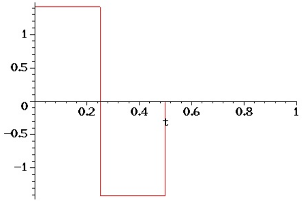
The result:

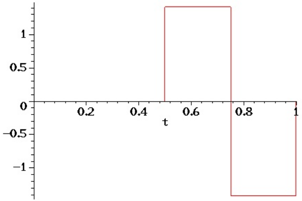


Level 2 Approximation

To improve the approximation we toss in functions

$$\psi_{1,0}(t) := \psi(2t) \text{ and } \psi_{1,1}(t) := \psi(2t-1)$$

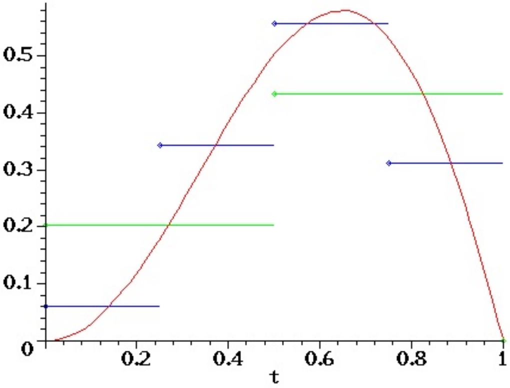




Both are orthogonal to each other and ϕ_0, ψ_0.

Level 2 Approximation

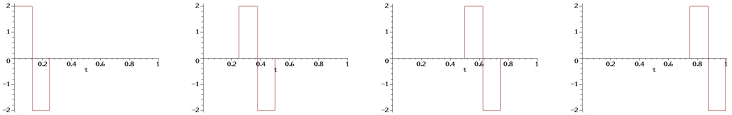
The approximation $f \approx c_0\phi_0 + d_0\psi_0 + d_{1,0}\psi_{1,0} + d_{1,1}\psi_{1,1}$ looks like



Level 3 Approximation

To improve the approximation further we toss in 4 new functions

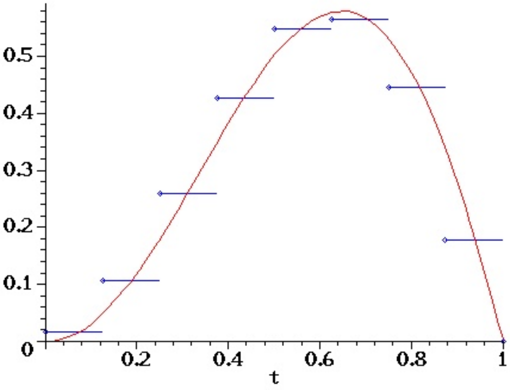
$$\begin{aligned} \psi_{2,0}(t) &:= \psi(4t),\ \psi_{2,1}(t) := \psi(4t-1), \\ \psi_{2,2}(t) &:= \psi(4t-2),\ \psi_{2,3}(t) := \psi(4t-3) \end{aligned}$$



All are orthogonal to each other and the previous functions.

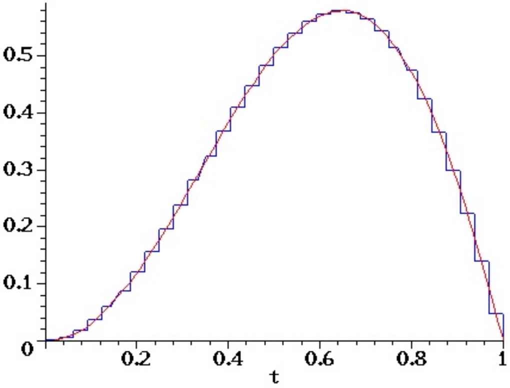
Level 3 Approximation

The approximation to f now looks like



Level 5 Approximation

If we toss if everything up to $\psi_{4,15}$ it looks like



Haar Summary

We have

- The Haar scaling function ϕ_0 (constant)

Haar Summary

We have

- The Haar scaling function ϕ_0 (constant)
- The "mother Haar wavelet" ψ_0

Haar Summary

We have

- The Haar scaling function ϕ_0 (constant)
- The "mother Haar wavelet" ψ_0
- The family of wavelets $\psi_{k,n}(t) = \psi(2^k t - n)$, translates and dilations of the mother Haar wavelet.

Haar Summary

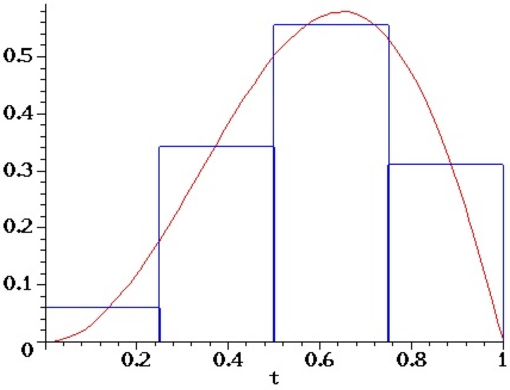
We have

- The Haar scaling function ϕ_0 (constant)
- The "mother Haar wavelet" ψ_0
- The family of wavelets $\psi_{k,n}(t) = \psi(2^k t - n)$, translates and dilations of the mother Haar wavelet.

The entire family is orthogonal and can be used to approximate any continuous function to arbitrary accuracy.

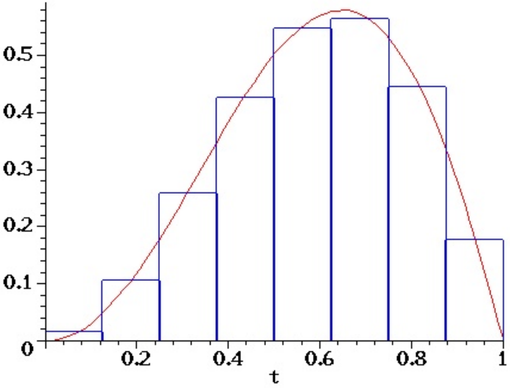
A Variation

Note: we could forget the wavelets and use just scalings/translates of the scaling function ϕ_0 to build f:



A Variation

If we want to boost resolution to the next level, throw out the $1/4$ wide basis functions, use $1/8$ wide functions.



Why the Wavelets?

- With scaling function at level 2 we use

$$\{\phi(4t), \phi(4t-1), \phi(4t-2), \phi(4t-3)\}.$$

Why the Wavelets?

- With scaling function at level 2 we use

$$\{\phi(4t), \phi(4t-1), \phi(4t-2), \phi(4t-3)\}.$$

- To go to level 3 we toss all these out and use

$$\{\phi(8t), \phi(8t-1), \ldots, \phi(8t-7)\}.$$

Why the Wavelets?

- With scaling function at level 2 we use

$$\{\phi(4t), \phi(4t-1), \phi(4t-2), \phi(4t-3)\}.$$

- To go to level 3 we toss all these out and use

$$\{\phi(8t), \phi(8t-1), \ldots, \phi(8t-7)\}.$$

- With wavelets, level 2 to level 3 lets us reuse previous basis functions

$$\underbrace{\{\phi_0, \psi_0, \psi_{1,0}, \psi_{1,1}\}}_{\text{level 2}} \cup \underbrace{\{\psi_{2,0}, \psi_{2,1}, \psi_{2,2}, \psi_{2,3}\}}_{\text{add for level 3}}$$

Generalizing

Can this be generalized? Specifically, are there other scaling functions $\phi(t)$ and wavelets $\psi(t)$ so that

- The set $\phi(t), \psi(t)$, and the wavelets $\psi_{k,n}$ are orthogonal,

Generalizing

Can this be generalized? Specifically, are there other scaling functions $\phi(t)$ and wavelets $\psi(t)$ so that

- The set $\phi(t), \psi(t)$, and the wavelets $\psi_{k,n}$ are orthogonal,
- Linear combinations can approximate any function to any desired accuracy,

Generalizing

Can this be generalized? Specifically, are there other scaling functions $\phi(t)$ and wavelets $\psi(t)$ so that

- The set $\phi(t), \psi(t)$, and the wavelets $\psi_{k,n}$ are orthogonal,
- Linear combinations can approximate any function to any desired accuracy,
- The functions have local support,

Generalizing

Can this be generalized? Specifically, are there other scaling functions $\phi(t)$ and wavelets $\psi(t)$ so that

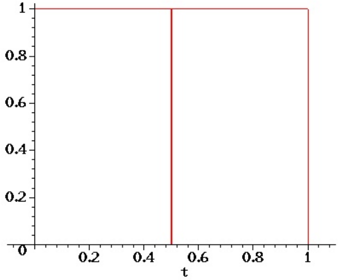
- The set $\phi(t), \psi(t)$, and the wavelets $\psi_{k,n}$ are orthogonal,
- Linear combinations can approximate any function to any desired accuracy,
- The functions have local support,
- The function are "easy" to compute?

More General Wavelets

Forget the wavelets for a minute. The essential ingredient in the Haar scheme is the scaling function. Note

$$\phi_0(t) = c_0\phi_0(2t) + c_1\phi_0(2t - 1)$$

with $c_0 = c_1 = 1$:



More General Wavelets

To generalize, seek a scaling function $\phi(t)$ with the property that $\phi(t)$ can itself be built from a linear combination of half-width translated versions of itself (the "dilation equation"):

$$\phi(t) = \sum_{m=0}^{M} c_m \phi(2t - m)$$

for some coefficients $c_0, \ldots, c_m$.

More General Wavelets

To generalize, seek a scaling function $\phi(t)$ with the property that $\phi(t)$ can itself be built from a linear combination of half-width translated versions of itself (the "dilation equation"):

$$\phi(t) = \sum_{m=0}^{M} c_m \phi(2t - m)$$

for some coefficients $c_0, \ldots, c_m$.

What should we use for the c_m? And if we know those, how would we find ϕ?

Finding ϕ

Pretend we know some suitable choices for the c_m. We can try fixed point iteration to compute ϕ:

1. Make an initial guess $\phi(t) = \phi_0(t)$.

Finding ϕ

Pretend we know some suitable choices for the c_m. We can try fixed point iteration to compute ϕ:

1. Make an initial guess $\phi(t) = \phi_0(t)$.
2. Iterate

$$\phi_{k+1}(t) = \sum_{m=0}^{M} c_m \phi_k(2t - m)$$

Finding ϕ

Pretend we know some suitable choices for the c_m. We can try fixed point iteration to compute ϕ:

1. Make an initial guess $\phi(t) = \phi_0(t)$.
2. Iterate
$$\phi_{k+1}(t) = \sum_{m=0}^{M} c_m \phi_k(2t - m)$$
3. Repeat to convergence.

Convergence

Under certain conditions on the c_m (algebraic, messy)

- The iteration converges to a function $\phi(t)$.

Convergence

Under certain conditions on the c_m (algebraic, messy)

- The iteration converges to a function $\phi(t)$.
- The function ϕ satisfies the dilation equation, and

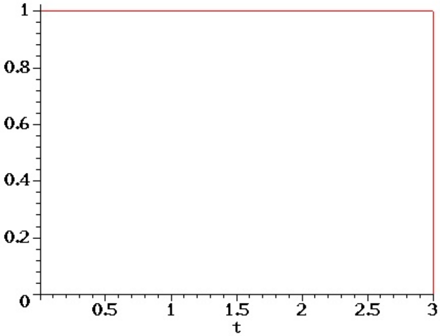
Convergence

Under certain conditions on the c_m (algebraic, messy)

- The iteration converges to a function $\phi(t)$.
- The function ϕ satisfies the dilation equation, and
- The set $\{\phi(2^N t - n); 0 \leq n \leq 2^N - 1\}$ can be used to approximate functions to arbitrary accuracy by taking N large.

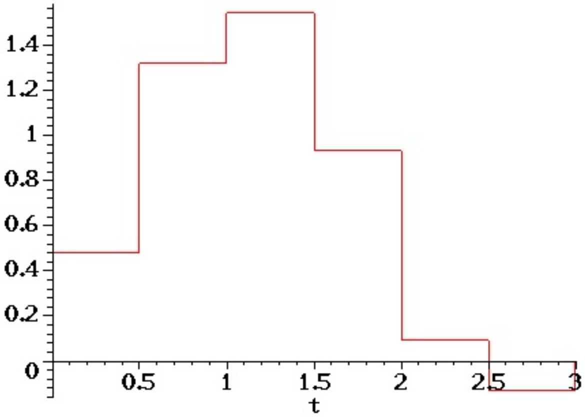
Example

Take $c_0 = (1+\sqrt{3})/4\sqrt{2}$, $c_1 = (3+\sqrt{3})/4\sqrt{2}$, $c_2 = (3-\sqrt{3})/4\sqrt{2}$, $c_3 = (1-\sqrt{3})/4\sqrt{2}$. Start with $\phi_0(t) = 1$ on $[0, 3]$:



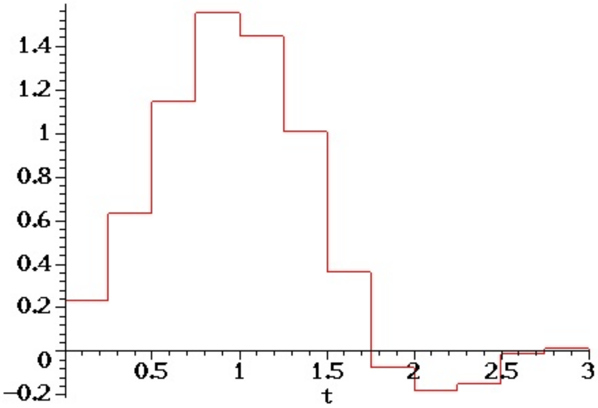
Example

First iteration: $\phi_1(t) = \sum_{m=0}^{3} c_m \phi_0(2t - m)$



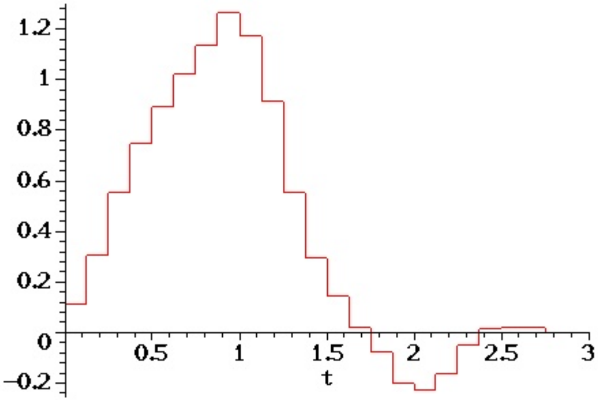
Example

Second iteration: $\phi_2(t) = \sum_{m=0}^{3} c_m \phi_1(2t - m)$



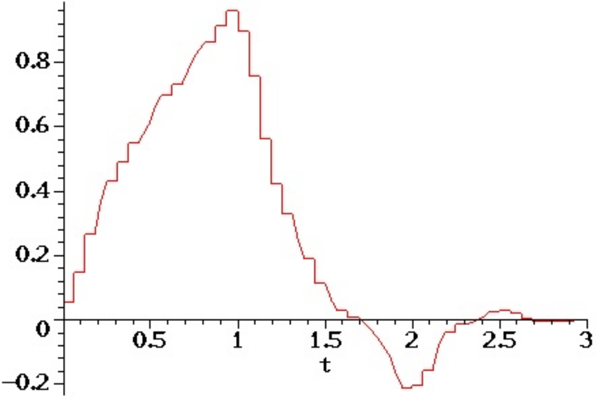
Example

Third iteration: $\phi_3(t) = \sum_{m=0}^{3} c_m \phi_2(2t - m)$



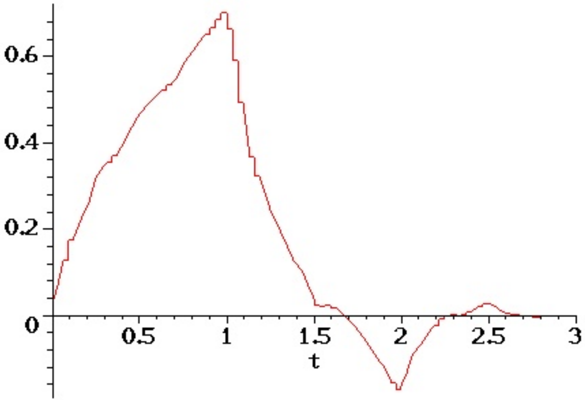
Example

Fourth iteration: $\phi_4(t) = \sum_{m=0}^{3} c_m \phi_3(2t - m)$



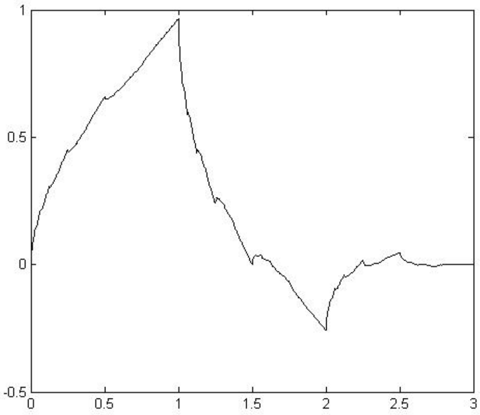
Example

Fifth iteration: $\phi_5(t) = \sum_{m=0}^{3} c_m \phi_4(2t - m)$



Example

The Daubechies D4 scaling function



Computing the Wavelet

If we find a scaling function that satisfies the dilation equation

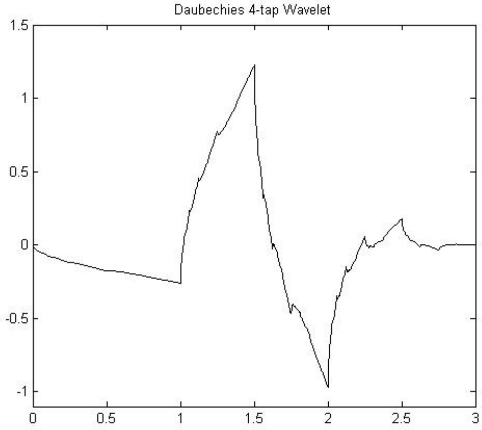
$$\phi(t) = \sum_{m=0}^{M} c_m \phi(2t - m)$$

then the mother wavelet ψ can be computed from

$$\psi(t) = \sum_{m=0}^{M} (-1)^m c_{M-m} \phi(2t - m)$$

Example

The Daubechies D4 mother wavelet



The D4 Wavelet Family

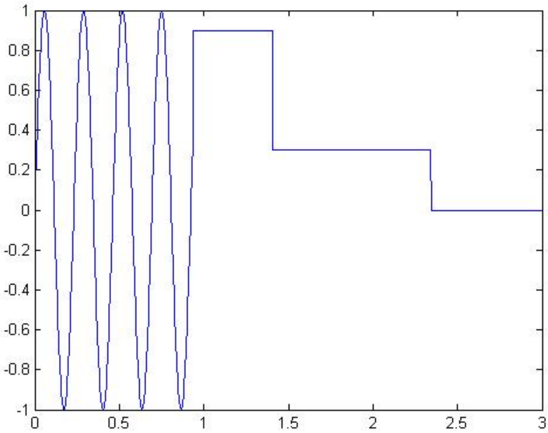
The D4 scaling function $\phi(t)$, the mother wavelet $\psi(t)$, and the translates/scalings

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

with $0 \leq n \leq 2^k - 1$ form an orthogonal basis for the space of (square-integrable) functions on $[0, 3]$.

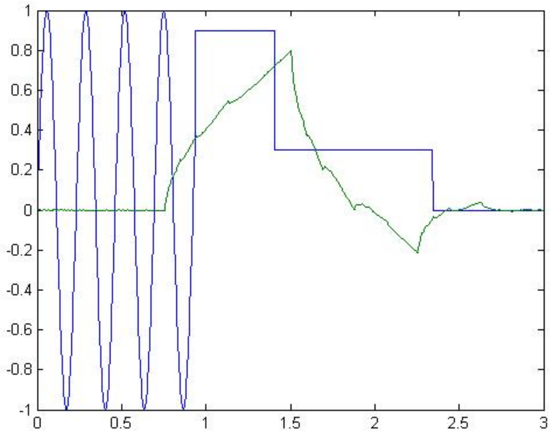
Example

A function on $[0, 3]$.



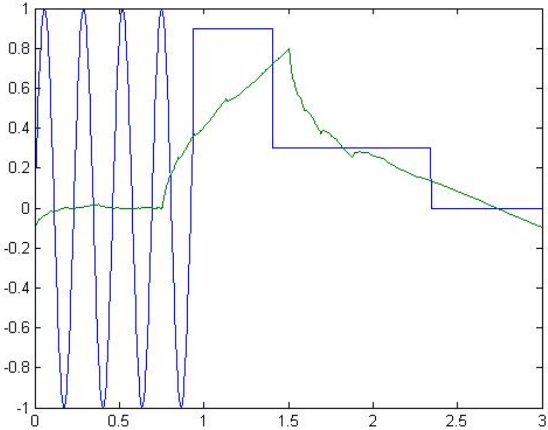
Example

Approximation from just scaling function $\phi(t)$:



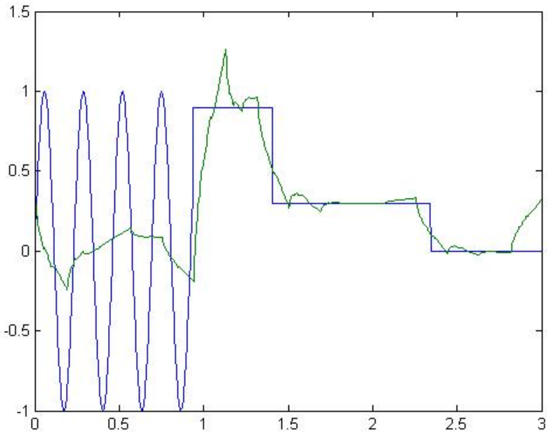
Example

Approximation from ϕ and mother wavelet ψ.



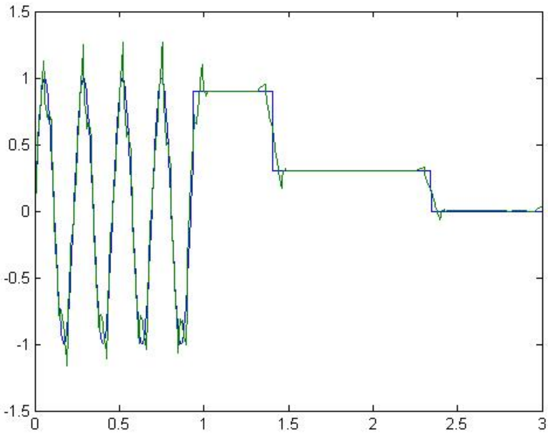
Example

Approximation from $\phi, \psi, \psi_{1,0}, \ldots, \psi_{3,7}$.



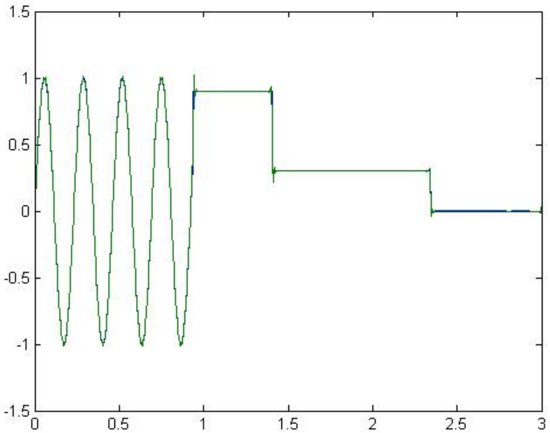
Example

Approximation from $\phi, \psi, \psi_{1,0}, \ldots, \psi_{5,31}$.



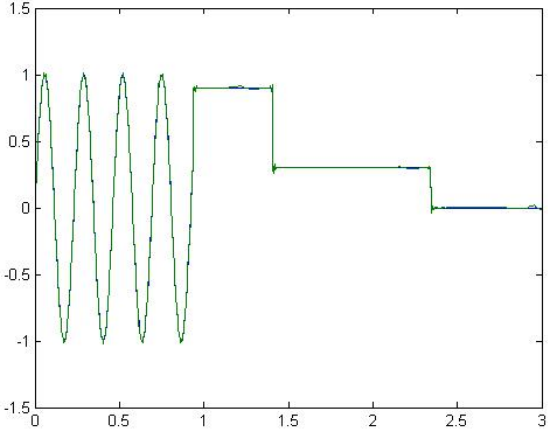
Example

Approximation from $\phi, \psi, \psi_{1,0}, \ldots, \psi_{7,127}$.



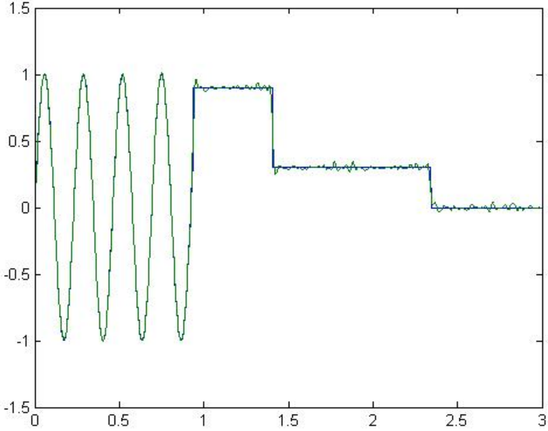
Compression Example: D4 Wavelets

Compute “all” coefficients $c_{j,k} = (f, \psi_{j,k})$ keep only 100 largest, reconstruct:



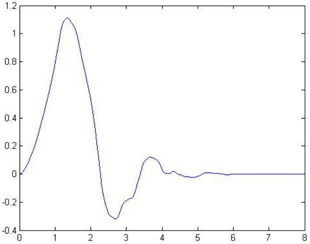
Compression Example: Cosine Basis

Compute “all” coefficients $c_k = (f, cos(k\pi t/3)$ keep only 100 largest, reconstruct:



Other Wavelet Families

There are MANY of other types of wavelets that have been constructed. The D8 scaling function and wavelet:



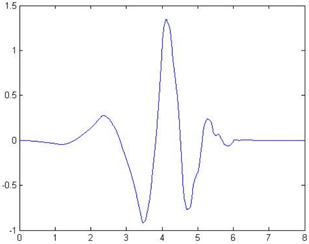


Image Compression Example, LeGall 5/3 Wavelets

An image (left) and wavelet compressed version (right, 75 percent compression).

Image Compression Example, LeGall 5/3 Wavelets

Wavelet compressed images at 94 percent (left) and 98.6 percent (right)

Conclusion

Wavelets have found many uses in mathematics and engineering:

- The JPEG 2000 compression standard is based on wavelets (the LeGall 5/3 and Daubechies 9/7 wavelets).
- The FBI compresses fingerprint records using a wavelet-based algorithm.
- Wavelets are used in signal processing/analysis (to localize frequency analysis).
- Wavelets are even useful in "pure" mathematics, as a tool in functional analysis.