A Tutorial on
Wavelets and their Applications

Martin J. Mohlenkamp

University of Colorado at Boulder
Department of Applied Mathematics
`mjm@colorado.edu`

This tutorial is designed for people with little or no experience with wavelets. We will cover the basic concepts and language of wavelets and computational harmonic analysis, with emphasis on the applications to numerical analysis. The goal is to enable those who are unfamiliar with the area to interact more productively with the specialists.

Thanks to Gregory Beylkin, Willy Hereman, and Lucas Monzón for help with this tutorial.

Goals

- Enable the wavelet novice to interact more productively with the specialists, by
  - Introducing the basic concepts and language
  - Doing some physics related examples
  - Explaining why people like them

Not Goals

- Give the history and assign credit

- Convince you that wavelets are better than any particular technique for any particular problem

Outline

- Multiresolution Analysis
  - wavelets (traditional)
  - properties
  - fast algorithms
- Connections with Fourier Analysis
  - Local Cosine and the phase plane
  - 1D Schrödinger example
  - wavelet packets
- Operators in Wavelet Coordinates
  - density matrix example
- Review and Questions

Multiresolution Analysis

A multiresolution analysis is a decomposition of $\mathbf{L}^2(\mathbf{R})$, into a chain of closed subspaces

$$\cdots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0 \subset \mathbf{V}_{-1} \subset \mathbf{V}_{-2} \subset \cdots \subset \mathbf{L}^2(\mathbf{R})$$

such that

1. $\bigcap_{j\in\mathbf{Z}} \mathbf{V}_j = \{0\}$ and
   $\bigcup_{j\in\mathbf{Z}} \mathbf{V}_j$ is dense in $\mathbf{L}^2(\mathbf{R})$

2. $f(x) \in \mathbf{V}_j$ if and only if $f(2x) \in \mathbf{V}_{j-1}$

3. $f(x) \in \mathbf{V}_0$ if and only if $f(x-k) \in \mathbf{V}_0$ for any $k \in \mathbf{Z}$.

4. There exists a *scaling function* $\varphi \in \mathbf{V}_0$ such that $\{\varphi(x-k)\}_{k\in\mathbf{Z}}$ is an orthonormal basis of $\mathbf{V}_0$.

Multiresolution Analysis

Let $\mathbf{W}_j$ be the orthogonal complement of $\mathbf{V}_j$ in $\mathbf{V}_{j-1}$:

$$\mathbf{V}_{j-1} = \mathbf{V}_j \oplus \mathbf{W}_j,$$

so that

$$\mathbf{L}^2(\mathbf{R}) = \bigoplus_{j \in \mathbf{Z}} \mathbf{W}_j.$$

Selecting a coarsest scale $\mathbf{V}_n$ and finest scale $\mathbf{V}_0$, we truncate the chain to

$$\mathbf{V}_n \subset \cdots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0$$

and obtain

$$\mathbf{V}_0 = \mathbf{V}_n \bigoplus_{j=1}^{n} \mathbf{W}_j.$$

From the scaling function φ we can define the *wavelet* ψ, such that $\{\psi(x-k)\}_{k \in \mathbf{Z}}$ is an orthonormal basis of $\mathbf{W}_0$.

Multiresolution Analysis

Example: Haar wavelets

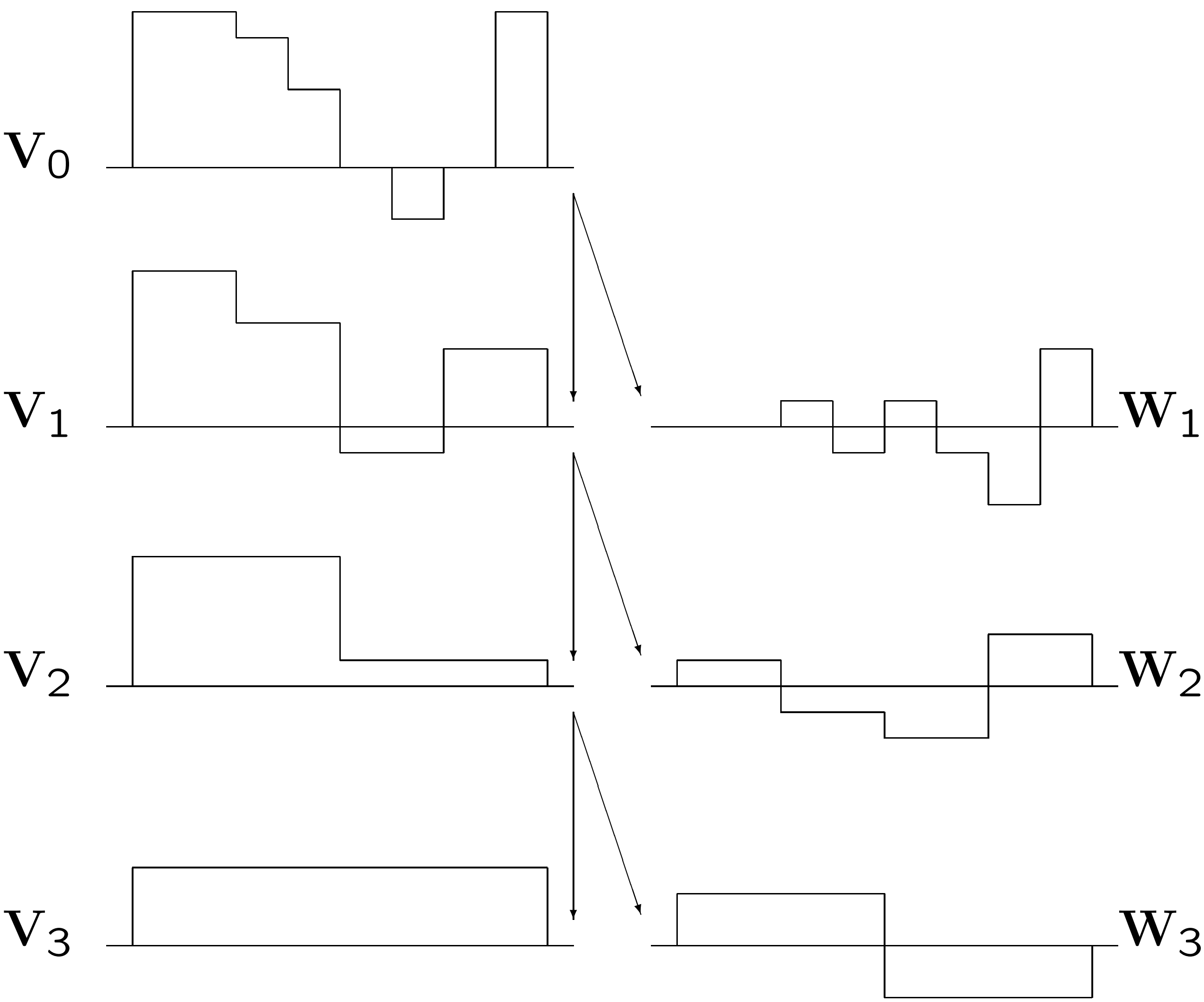
$$\varphi(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

$\mathbf{V}_0 = \text{span}(\{\varphi(x-k)\}_{k\in\mathbf{Z}})$ are piecewise constant functions with jumps only at integers.

$$\psi(x) = \begin{cases} 1 & \text{for } 0 < x < 1/2 \\ -1 & \text{for } 1/2 \le x < 1 \\ 0 & \text{elsewhere.} \end{cases},$$

$\mathbf{W}_0 = \text{span}(\{\psi(x-k)\}_{k\in\mathbf{Z}})$ are piecewise constant functions with jumps only at half-integers, and average 0 between integers.

Multiresolution Analysis



$$\mathbf{V}_0 = \mathbf{V}_3 \oplus \mathbf{W}_3 \oplus \mathbf{W}_2 \oplus \mathbf{W}_1$$

Where's the Wavelet?

Since $\mathbf{W}_j$ is a dilation of $\mathbf{W}_0$, we can define

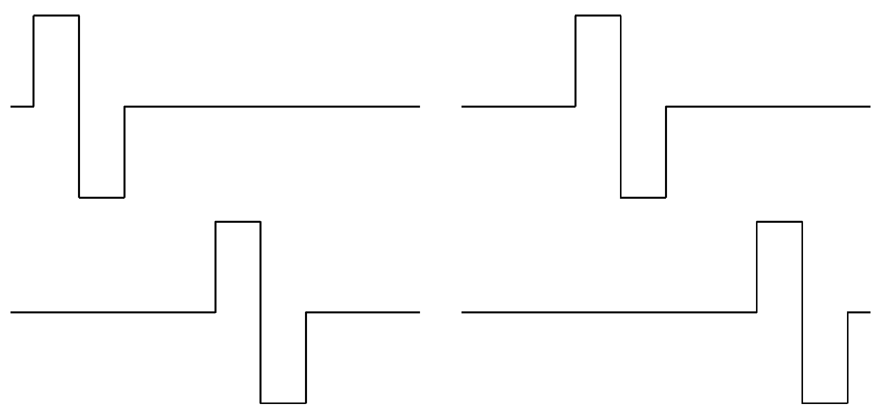
$$\psi_{j,k} = 2^{-j/2}\psi(2^{-j}x - k)$$

and have

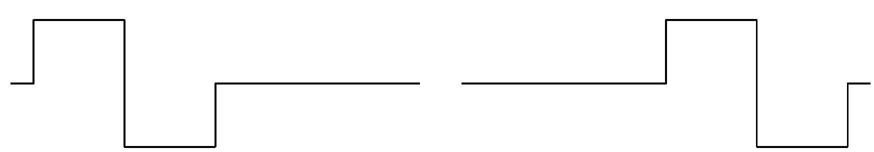
$$\mathbf{W}_j = \text{span}(\{\psi_{j,k}(x)\}_{k \in \mathbf{Z}}).$$

In this example,

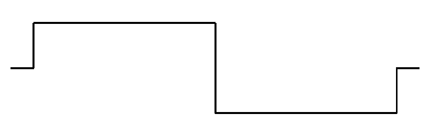
$\mathbf{W}_1 = \text{span}$



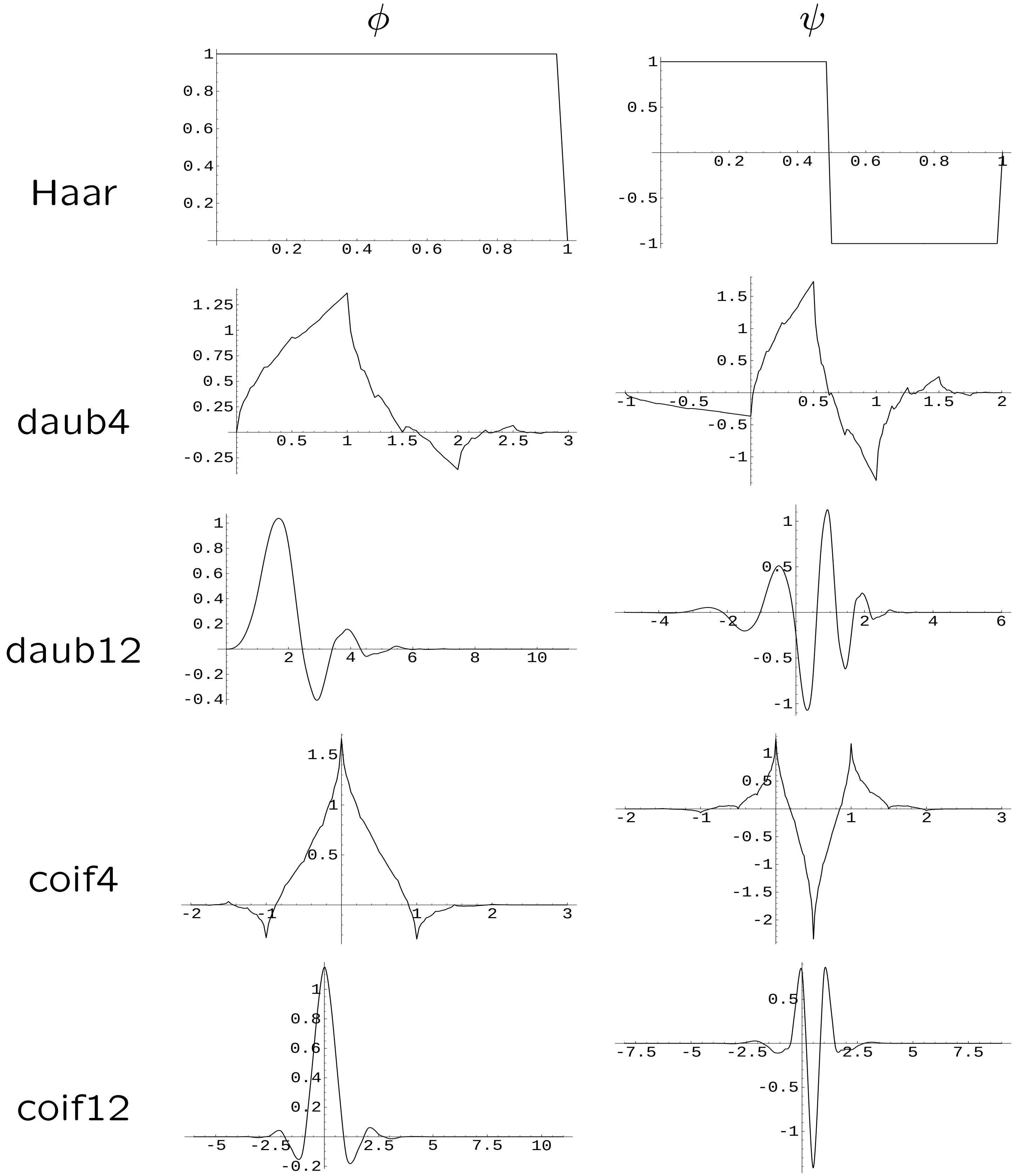
$\mathbf{W}_2 = \text{span}$



$\mathbf{W}_3 = \text{span}$



The Wavelet Zoo



Vanishing Moments

Wavelets are usually designed with vanishing moments:

$$\int_{-\infty}^{+\infty} \psi(x) x^m dx = 0, \qquad m = 0, \ldots, M-1,$$

which makes them orthogonal to the low degree polynomials, and so tend to compress non-oscillatory functions.

For example, we can expand in a Taylor series

$$f(x) = f(0) + f'(0)x + \cdots + f^{(M-1)}(0)\frac{x^{M-1}}{(M-1)!} + f^{(M)}(\xi(x))\frac{x^M}{M!}$$

and conclude

$$|\langle f, \psi \rangle| \leq \max_x \left| f^{(M)}(\xi(x)) \frac{x^M}{M!} \right|.$$

Haar has $M = 1$.

Quadrature Mirror Filters

Wavelets are designed through properties of their "quadrature mirror filter" $\{H, G\}$.

$$\begin{aligned}\text{Haar} \leftrightarrow H &= \frac{1}{\sqrt{2}}[1, 1]\\ G &= \frac{1}{\sqrt{2}}[1, -1]\\ \text{daub4} \leftrightarrow H &= \frac{1}{4\sqrt{2}}[1+\sqrt{3}, 3+\sqrt{3}, 3-\sqrt{3}, 1-\sqrt{3}]\\ G &= [H(3), -H(2), H(1), -H(0)]\}\end{aligned}$$

(The values are usually not in closed form.)

For instance, vanishing moments

$$\int_{-\infty}^{+\infty} \psi(x) x^m dx = 0, \qquad m = 0, \ldots, M-1,$$

are a consequence of

$$\sum_i G(i) i^m = 0, \qquad m = 0, \ldots, M-1.$$

Trade-offs

You can get

- higher M
- more derivatives
- closer to symmetric
- closer to interpolating (coiflets)

if you pay by increasing the filter length, which causes

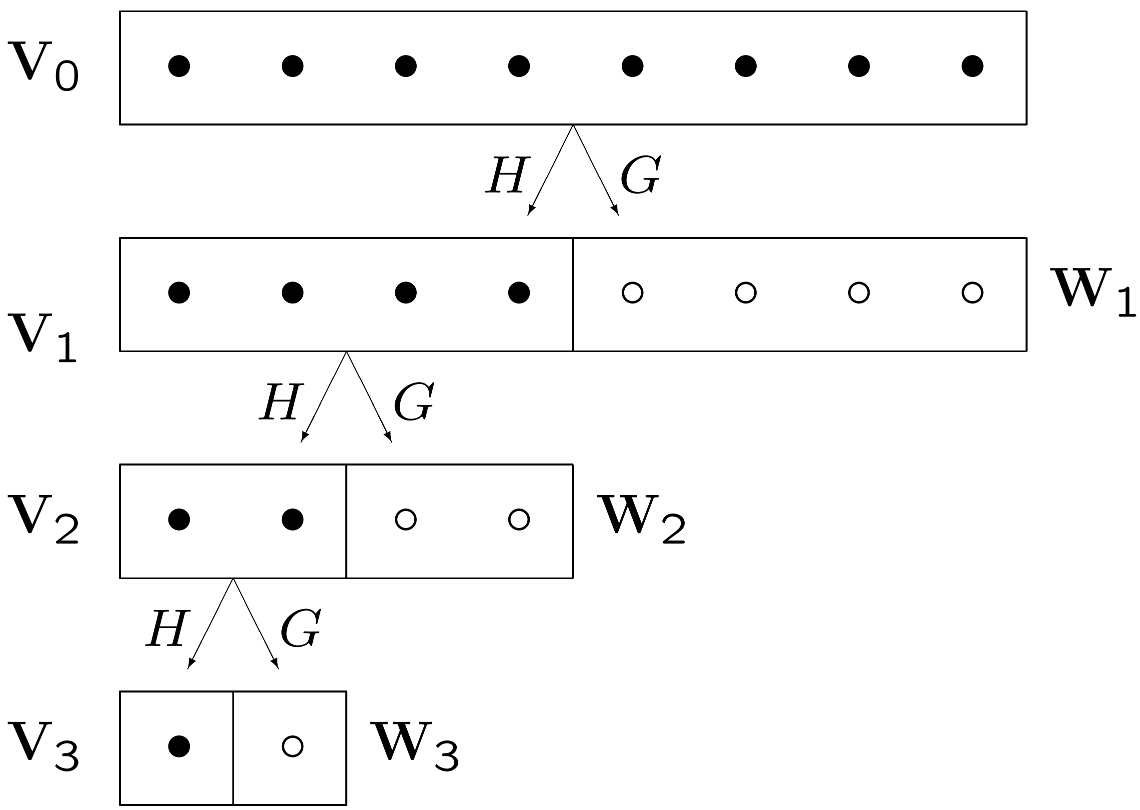
- longer (overlapping) supports, and so worse localization.

- slower transforms

The cost is linear (in M etc.).

Fast Wavelet Transform

Sample onto the finest resolution and then apply the "quadrature mirror filter" $\{H, G\}$.



The total cost of this cascade is $2 \cdot N \cdot L$, where L is the length of the filter.

There are many, many 'lets

By loosening the definitions, you can get

- symmetric
- interpolating
- 2D properties (brushlets)
- ...

Rules of thumb:

- Use a special purpose wavelet if and only if you have a special need.

- Use one vanishing moment per digit desired (truncation level).

- Do not use more derivatives than your function typically has.

Multiwavelets

(Polynomial version on $[0,1]$)

Fix $k \in \mathbf{N}$, and let $\mathbf{V}_n$ be the space of functions that are polynomials of degree less than k on the intervals $(2^n j, 2^n(j+1))$ for $j = 1, \ldots, 2^{-n} - 1$, and 0 elsewhere.

$\mathbf{V}_0$ is spanned by k scaling functions.

$\mathbf{W}_0$ is spanned by k multiwavelets.

By construction, the wavelets have k vanishing moments, and so the same sparsity properties as ordinary wavelets.

The wavelets are not even continuous, but this allows weak formulations of the derivative, which allows better treatment of boundaries.

Connections with Fourier Analysis

Fourier analysis gives an understanding of frequency, but "non-stationary" signals beg for space (time) localization.

This need motivates Computational Harmonic Analysis and its tools, such as wavelets and local cosine.

The theory and intuition are still based on Fourier analysis.

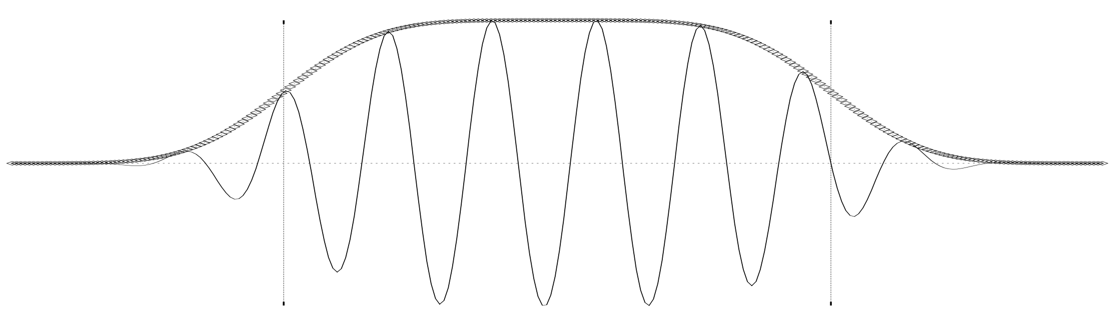
Local Cosine Basis

Partition the line (interval, circle) with
$\cdots a_i < a_{i+1} \cdots$, $I_i = [a_i, a_{i+1}]$.

Construct a set of bells: $\{b_i(x)\}$
with $\sum_i b_i^2(x) = 1$, $b_i(x) b_{i-1}(x)$ even about a_i,
and $b_i(x) b_j(x) = 0$ if $j \neq i \pm 1$.

Construct the cosines which are even on the left and odd on the right

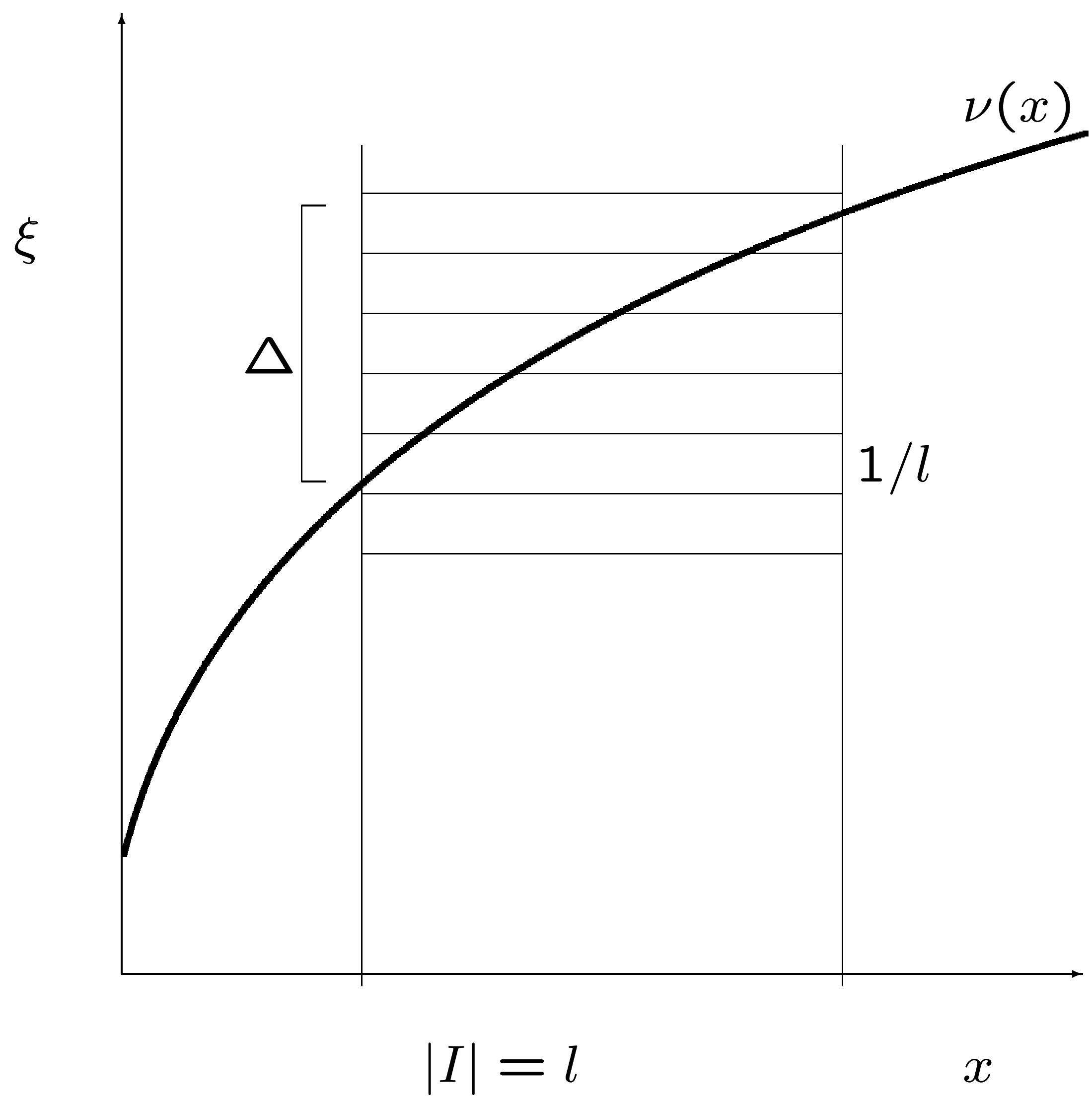
$$\left\{ c_i^j(x) = \sqrt{\frac{2}{a_{i+1} - a_i}} \cos\left(\frac{(j + 1/2)\pi (x - a_i)}{a_{i+1} - a_i} \right) \right\}.$$



$\{b_i(x) c_i^j(x)\}$ forms an orthonormal basis with fast transform based on the FFT.

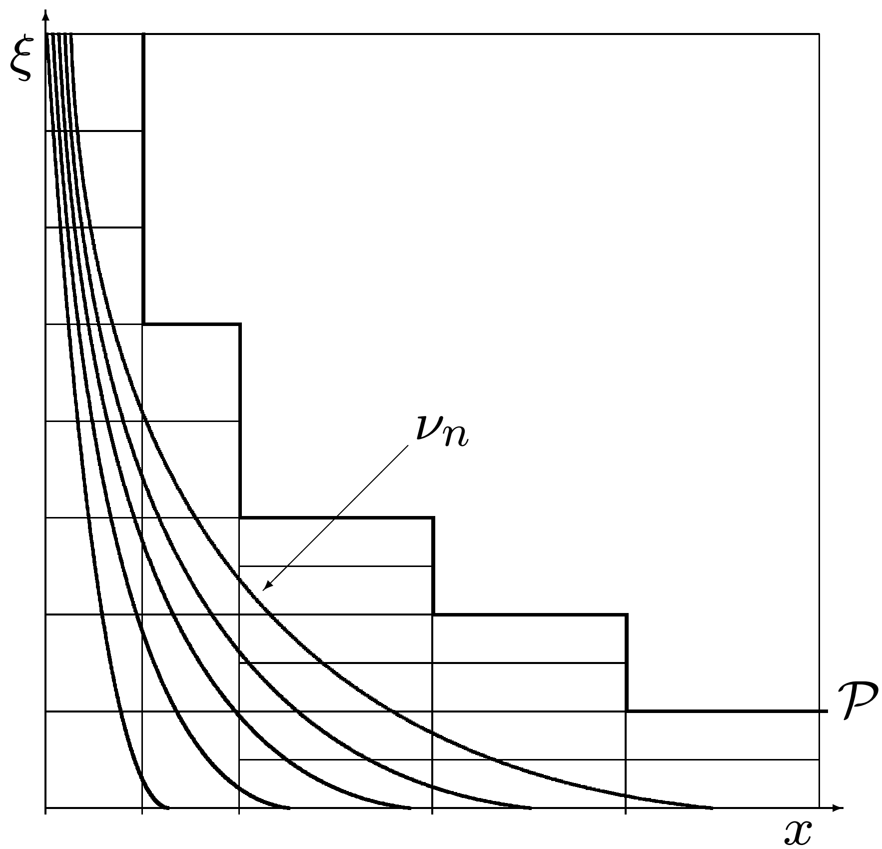
Phase Plane Intuition

If a function has 'instantaneous frequency' $\nu(x)$, it should be represented by those Local Cosine basis elements whose rectangles intersect $\nu(x)$. There are $\max\{\Delta l, 1\}$ of these.



Local Cosine Phase Plane

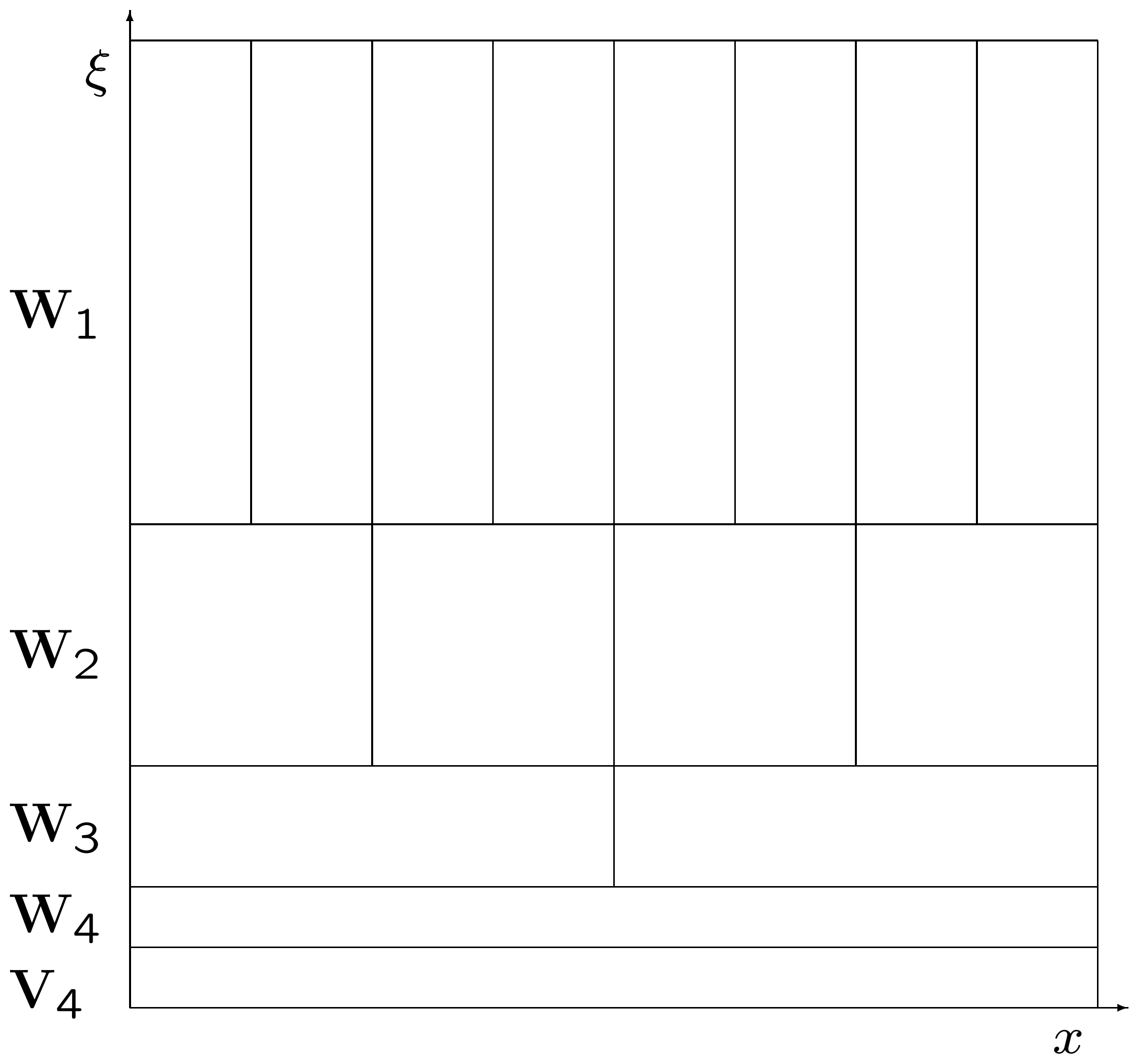
Consider the eigenfunctions of $-\Delta - C/x$, which have instantaneous frequency $\nu_n = \sqrt{C/x + \lambda_n}$.



We can get an efficient representation (with proof) by adapting, but this is not very flexible, especially in higher dimensions.

Wavelet Phase Plane

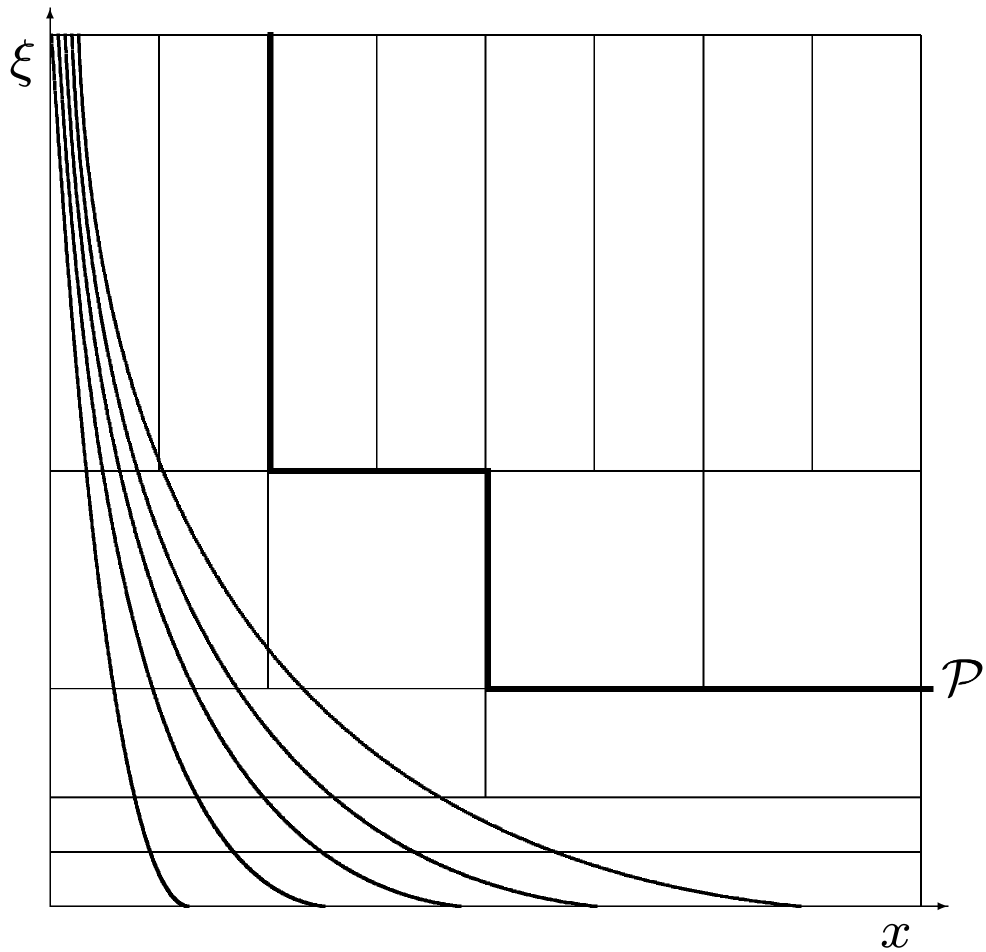
The multiresolution analysis divides the phase plane differently:



For wavelets this is intuitive but not rigorous.

Wavelet Phase Plane

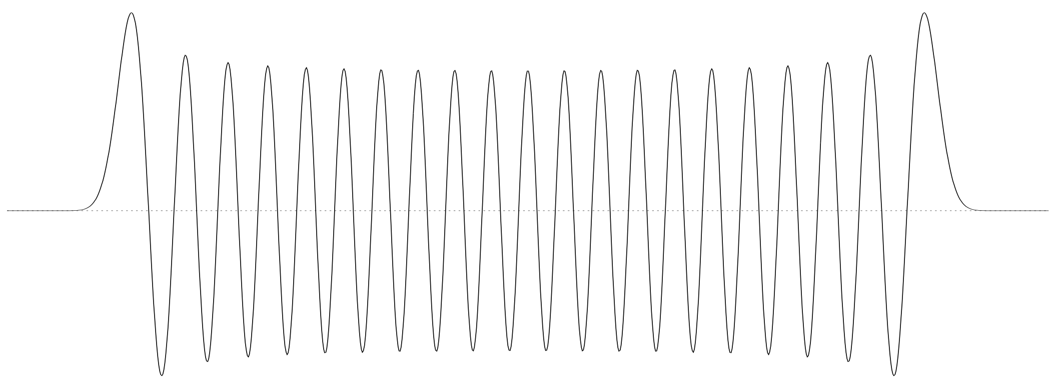
Consider again the instantaneous frequency $\nu_n = \sqrt{C/x + \lambda_n}$.



We get an efficient representation without adapting, so the location of the discontinuity is not important.

Tones

Sustained high frequencies, such as those in spherical harmonics

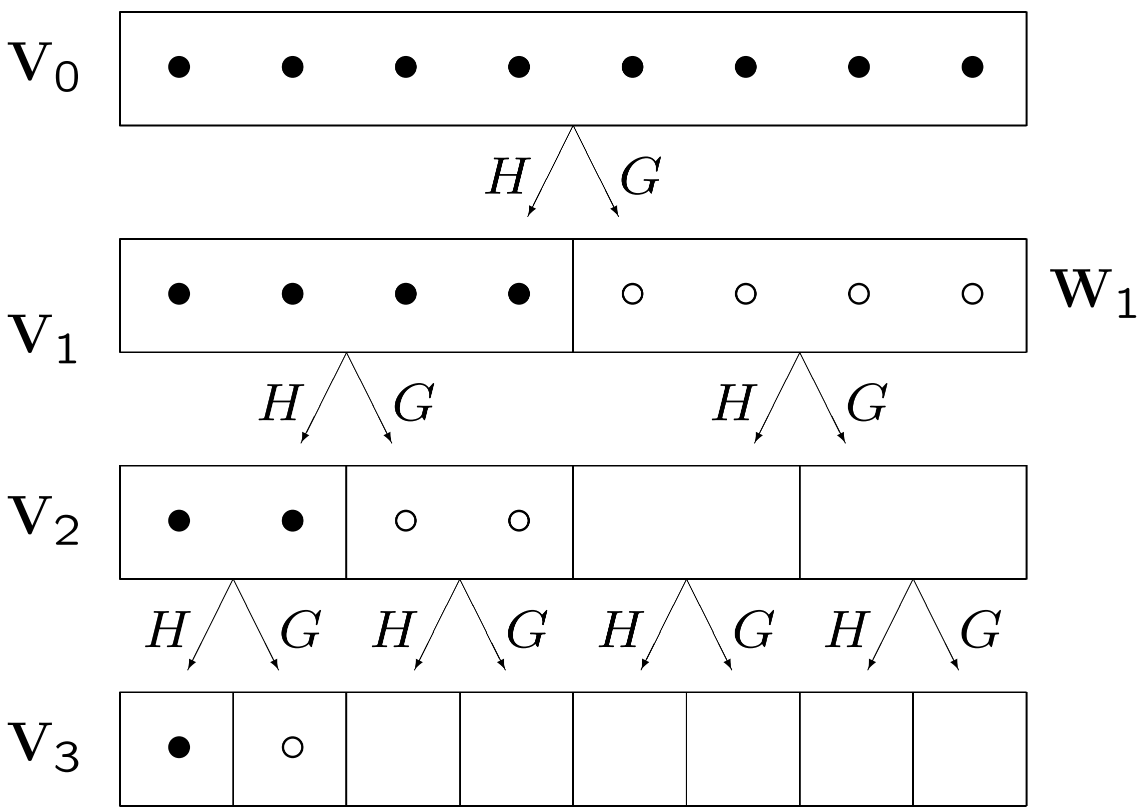


are a problem for Wavelets, but not for local cosine.

To enable wavelets to handle such functions, “wavelet packets” were developed.

Wavelet Packets and Best Basis Searches

Idea: by filtering the wavelet spaces, we can partition phase space in different ways:

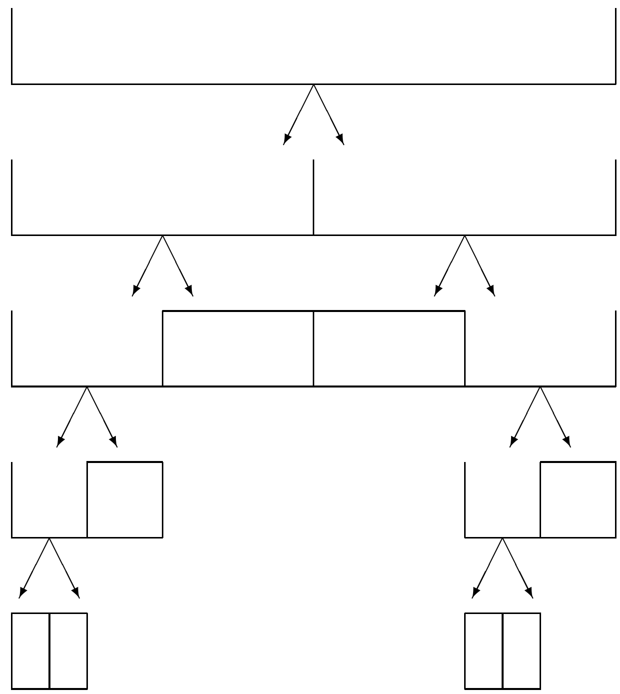


Any choice of decompositions gives a wavelet packet representation.

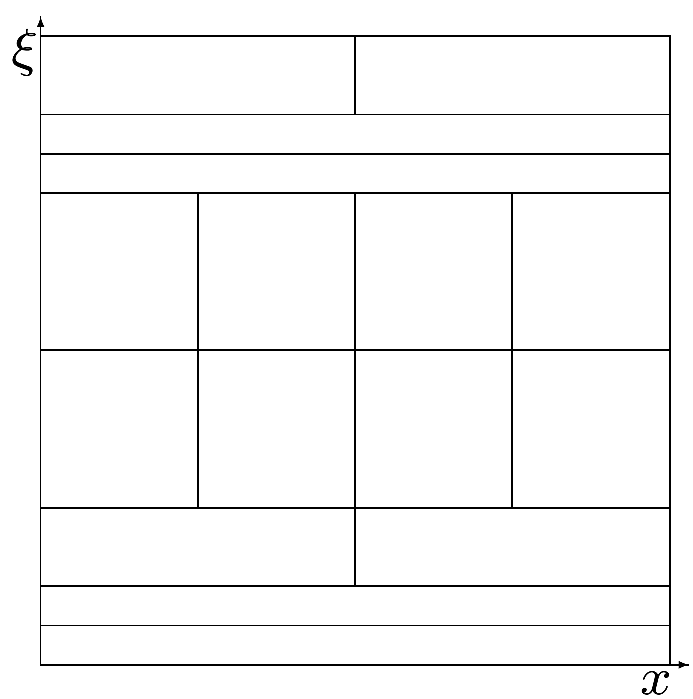
A fast tree search can find the "best basis".

Wavelet Packets Phase Plane

If we choose:



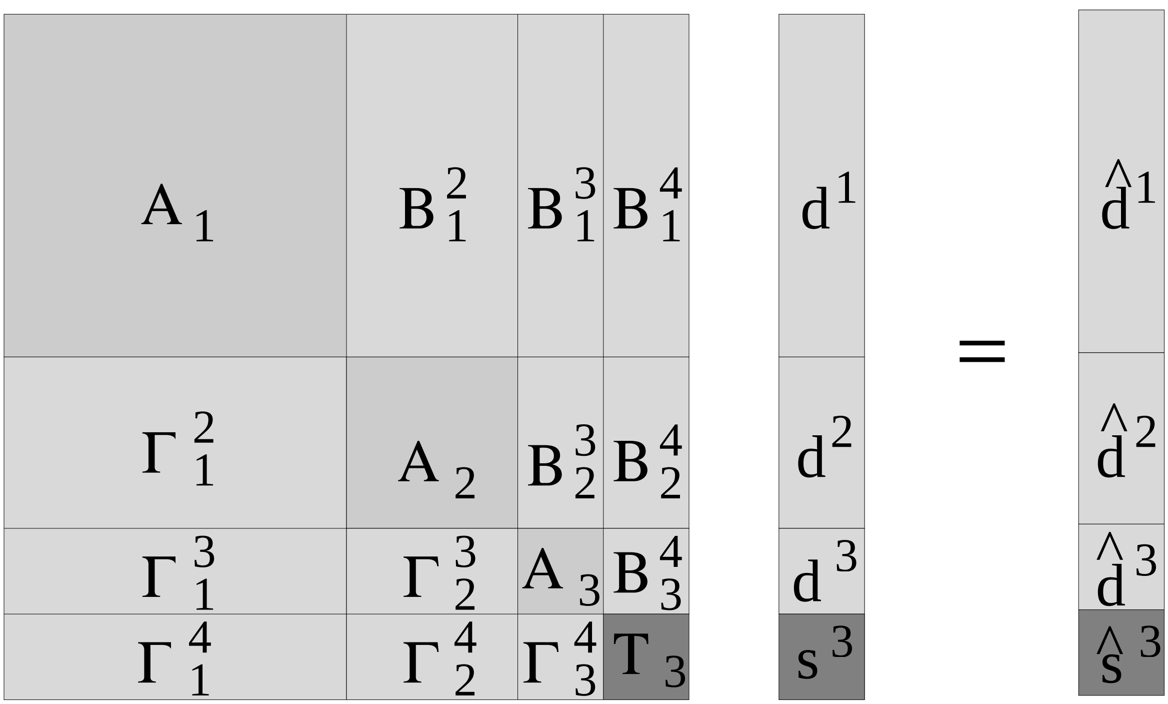
then our phase plane looks like:



Operators

There are many competing, adaptive ways to represent functions.

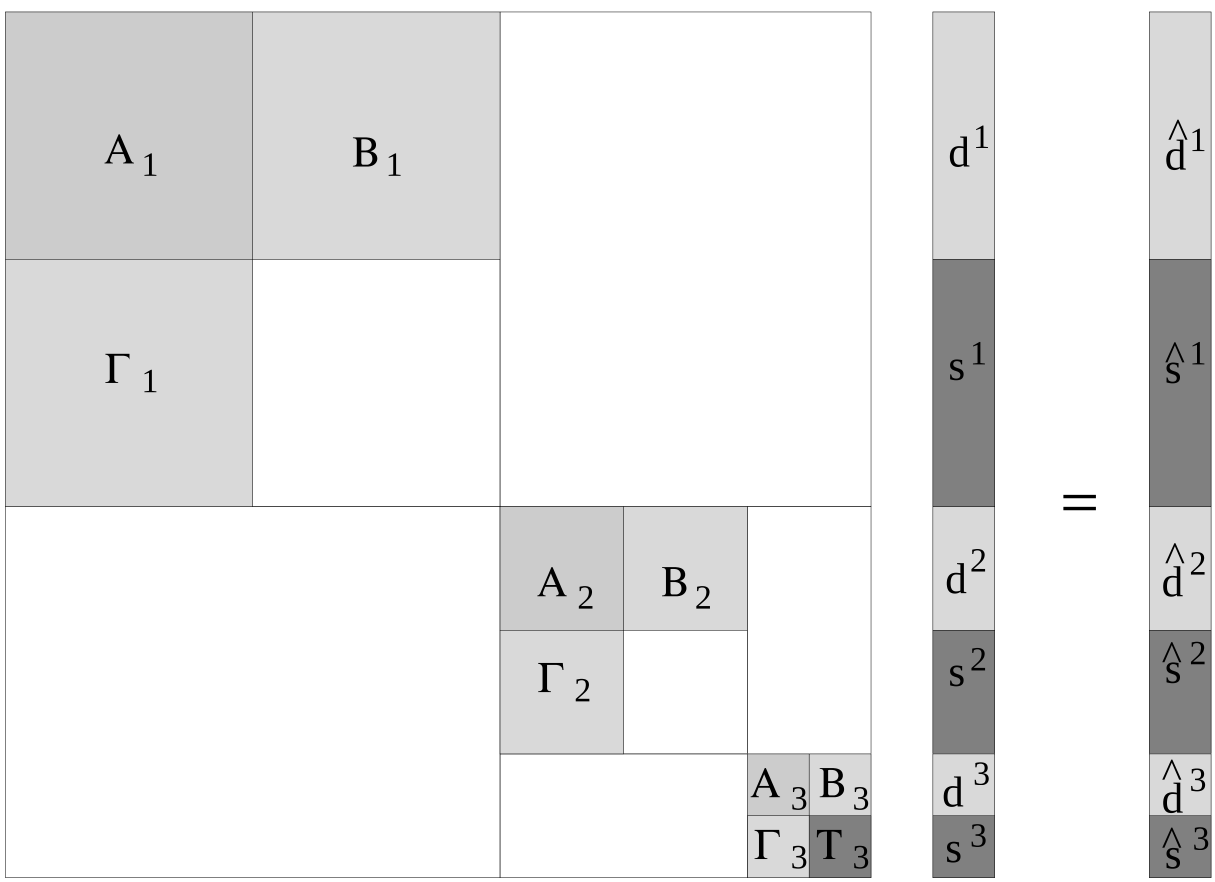
It is more interesting to consider operators and develop operator calculus.



Many operators are sparse in Wavelet bases.

Operators in the Nonstandard Form

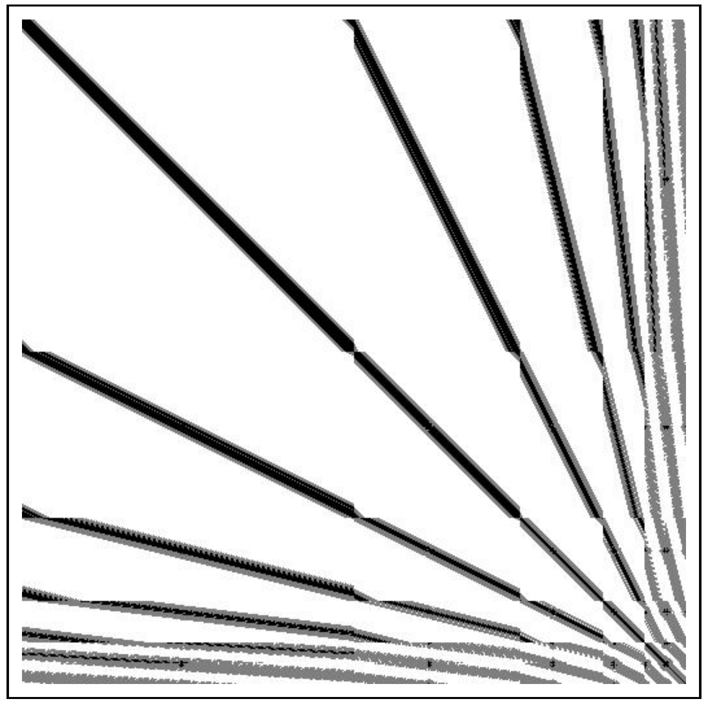
The nonstandard form gives a more isotropic, and often more sparse, representation.



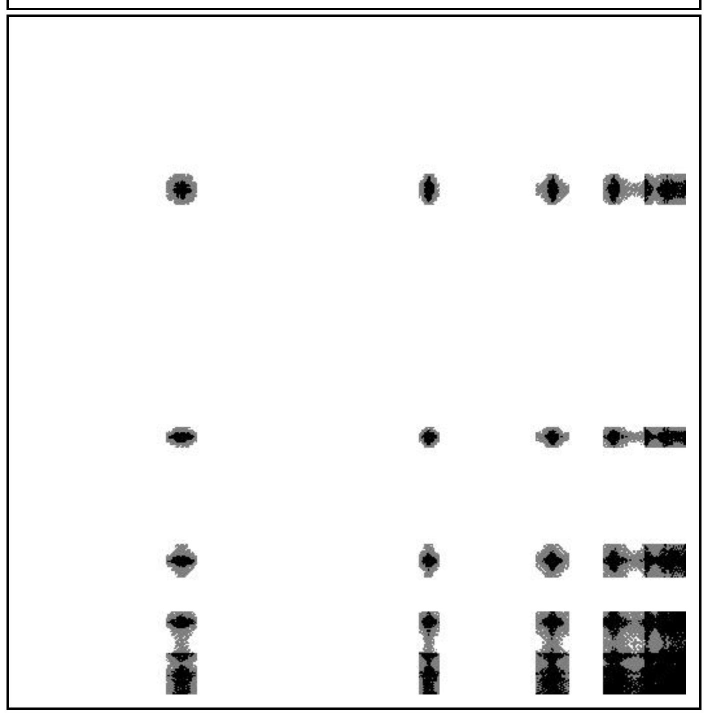
Operators in Wavelets

Hamiltonian

$$-\Delta - \frac{300}{|x|}$$

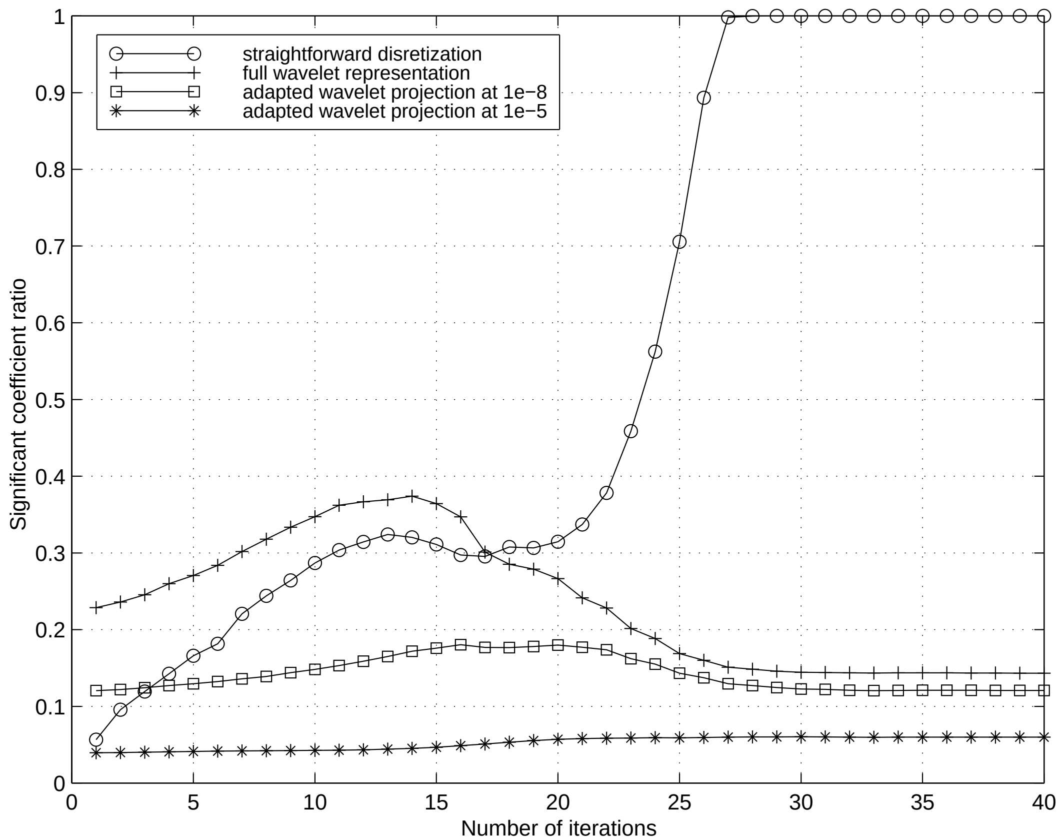


Density Matrix with 15 eigenfunctions



Operators in Wavelets

Some operators can be computed rapidly.



Here the density matrix is computed via the sign iteration

$$\begin{aligned} T_0 &= T/||T||_2 \\ T_{k+1} &= (3T_k - T_k^3)/2, \quad k = 0, 1, \ldots \end{aligned}$$

Philosophical Review

Multiscale assumption: Efficient when high frequencies (sharp features) happen for a short amount of time/space.

Cleanly adaptive: Refine or coarsen the "grid" by adding or deleting basis functions.

Automatically adaptive: Simply truncate small coefficients.

Tunable: Decide on the needed precision and properties first, then choose which wavelet.