Advanced Image Processing

Part II: Image Representation

S. Voloshynovskiy

Course Outline

- Recall of Linear Algebra.
- Introduction. Human Visual System.
- Image Representation: pyramids and wavelets.
- Random Signals.
- Image Modeling.
- Image Sensor Models. Noise Models.
- Image Denoising.
- Image Restoration.
- Image Compression.
- Video Modeling and Compression.
- Digital Data Hiding.

Recommended books

- G. Strang and T. Nguen, Wavelets and Filter Banks, Wellesley, MA, Wellesley-Cambrodge, 1996.
- M. Vetterli and J. Kovacevic, Wavelets and Subband Coding, Prentice-Hall, 1995.
- S. Mallat, A Wavelet Tour of Signal Processing, San Diego, Academic Press, 1998.
- A. Bovik, Handbook of Image & Video Processing, Academic Press, 2000.
- A. K. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, 1989.
- M. Rabbani and P. Jones, Digital Compression Techniques, SPIE, 1991.

Recommended books

Key Paper:

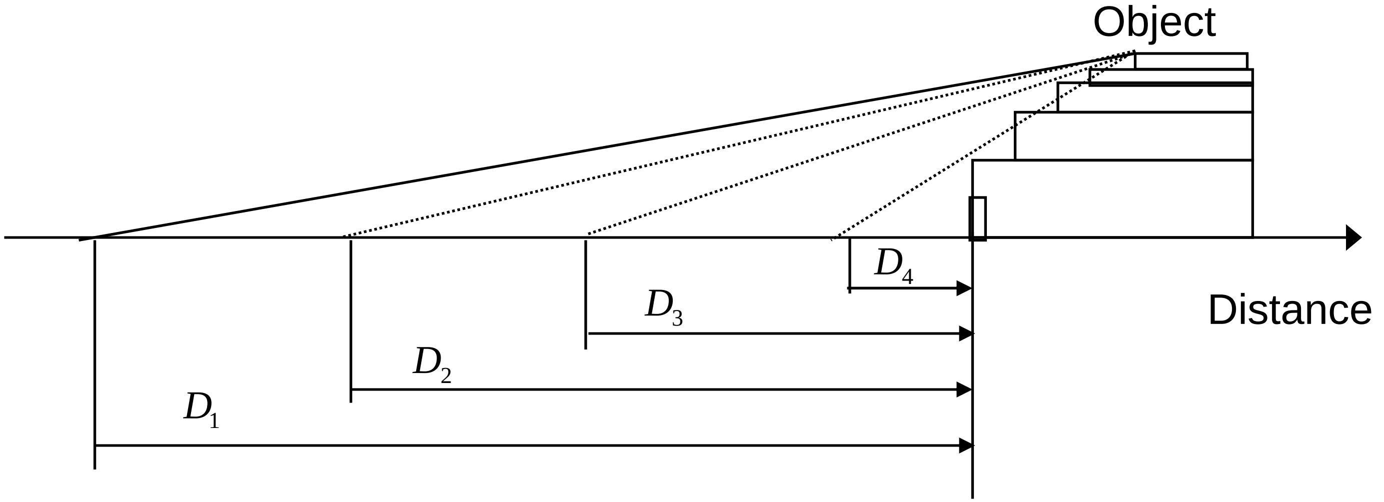
- S.G. Mallat, A theory of multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. PAMI-11, No 7, pp. 674-693, 1989.

Roadmap:

1. Multiscale Image Decomposition
2. Pyramid decomposition (Gaussian and Laplacian Pyramids)
3. Wavelets

1. Multiscale Image Decomposition

Intuitive interpretation of resolution or scale.



What can you see from different distances?

1. Multiscale Image Decomposition

Intuitive advantages:

Lead to hierarchical image representation that can be efficiently used for reduced-complexity algorithms in:

- texture analysis
- compression
- restoration;
- segmentation;
- edge detection;
- motion analysis;
- image understanding in Computer Vision.

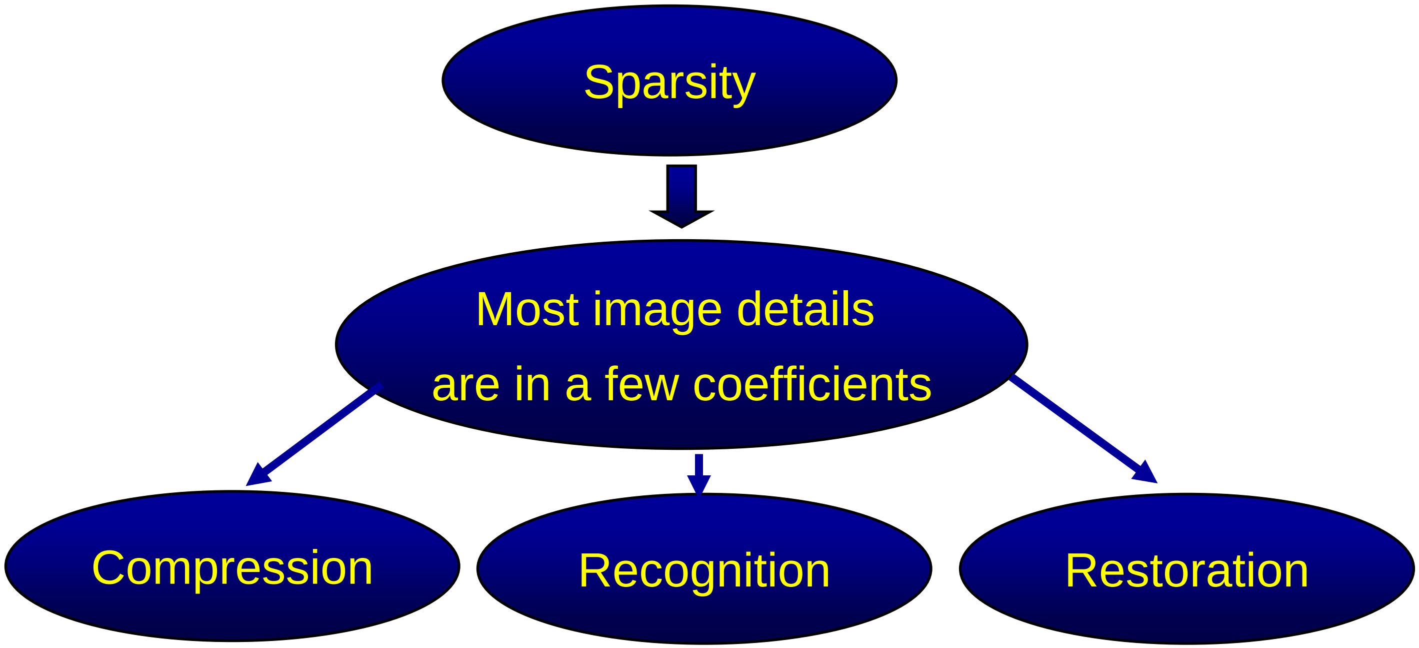
1. Multiscale Image Decomposition

Two main methods of multiscale image decompositions:

- Gaussian-Laplacian pyramids
- Wavelets

1. Multiscale Image Decomposition

<u>Main features of Laplacian pyramids and wavelet coefficients</u>:

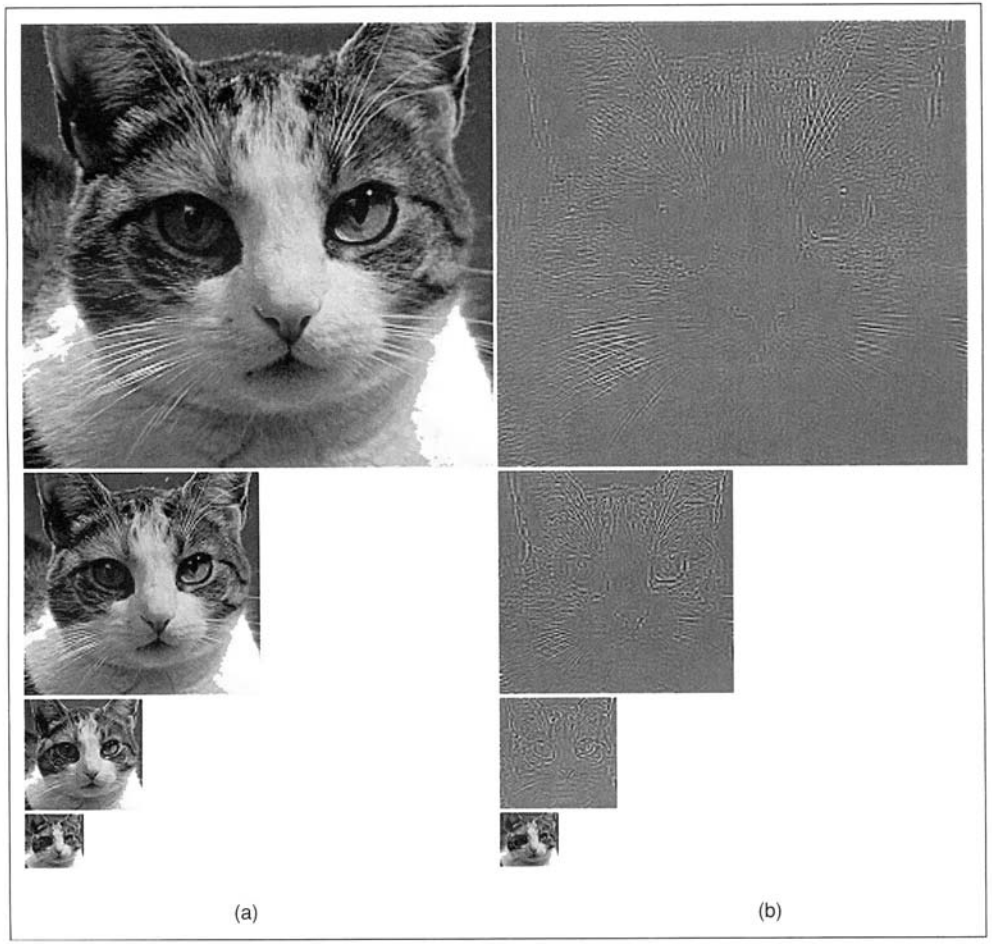


Bits are allocated only for significant coefficients

Easy to search in the space of significant features

Image and noise possess rather distinct properties in transform domain

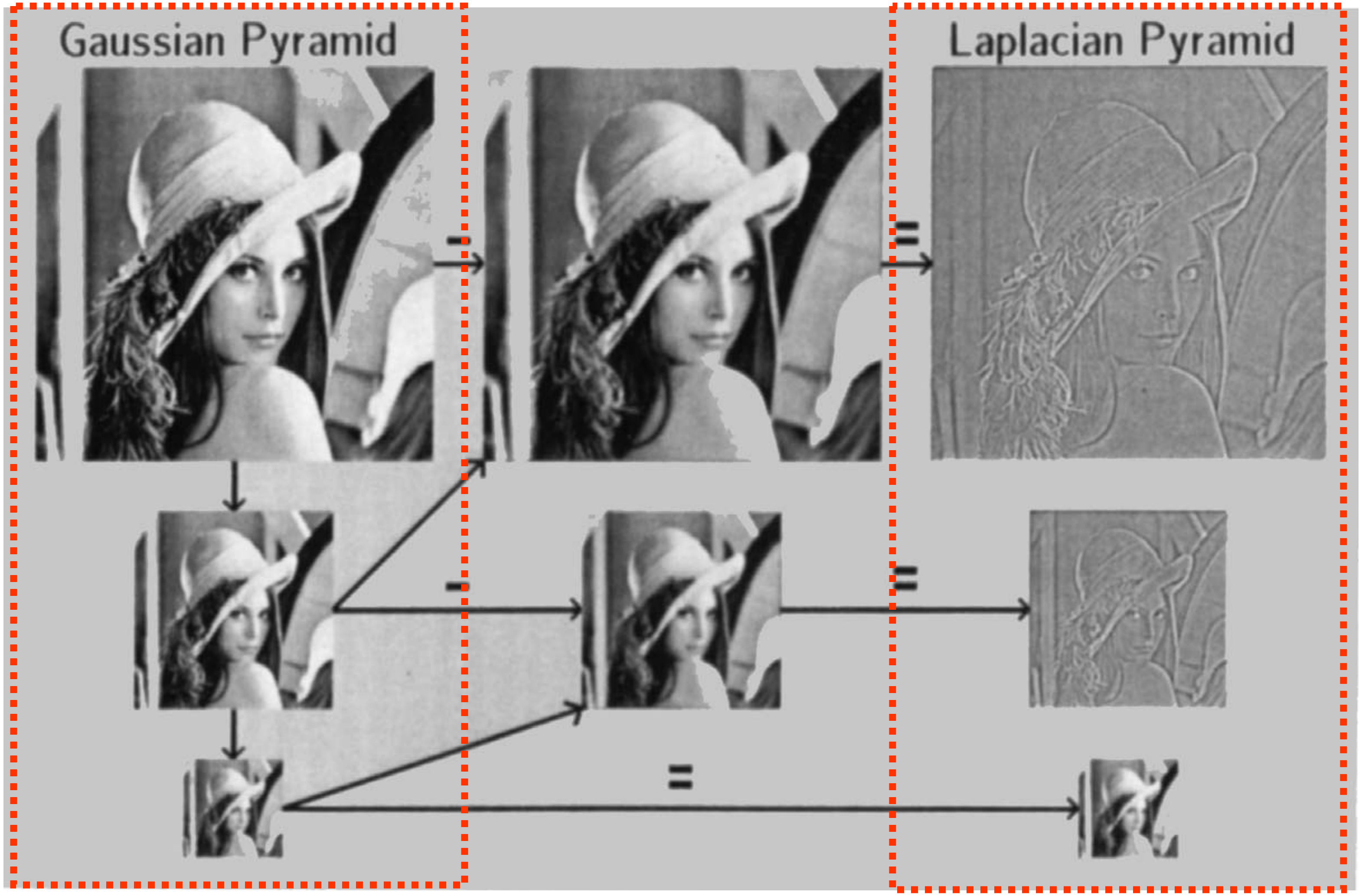
2. Pyramid decomposition (IEEE Com Mag, Unser, Nov 1999)



Gaussian Laplacian

1983: Burt and Adelson proposed pyramid decomposition.

2. Pyramid decomposition mechanism (Rabbani, p.200)

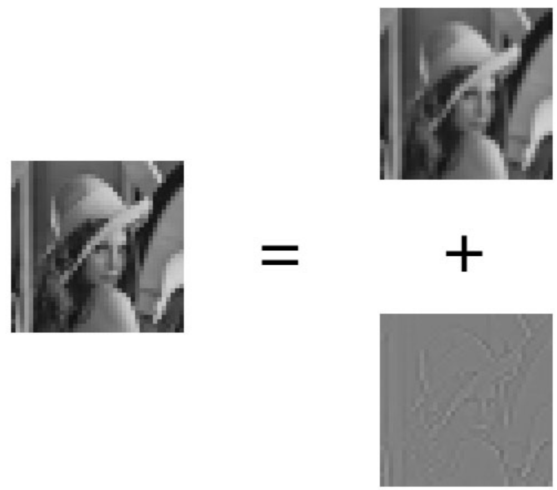


2. Pyramid decomposition

Main idea:

Every image can be represented as a linear combination of low-pass and high-pass components.

$$x = x_{LP} + x_{HP}$$



Low-pass

High-pass

Main concept: “**separate and rule**”

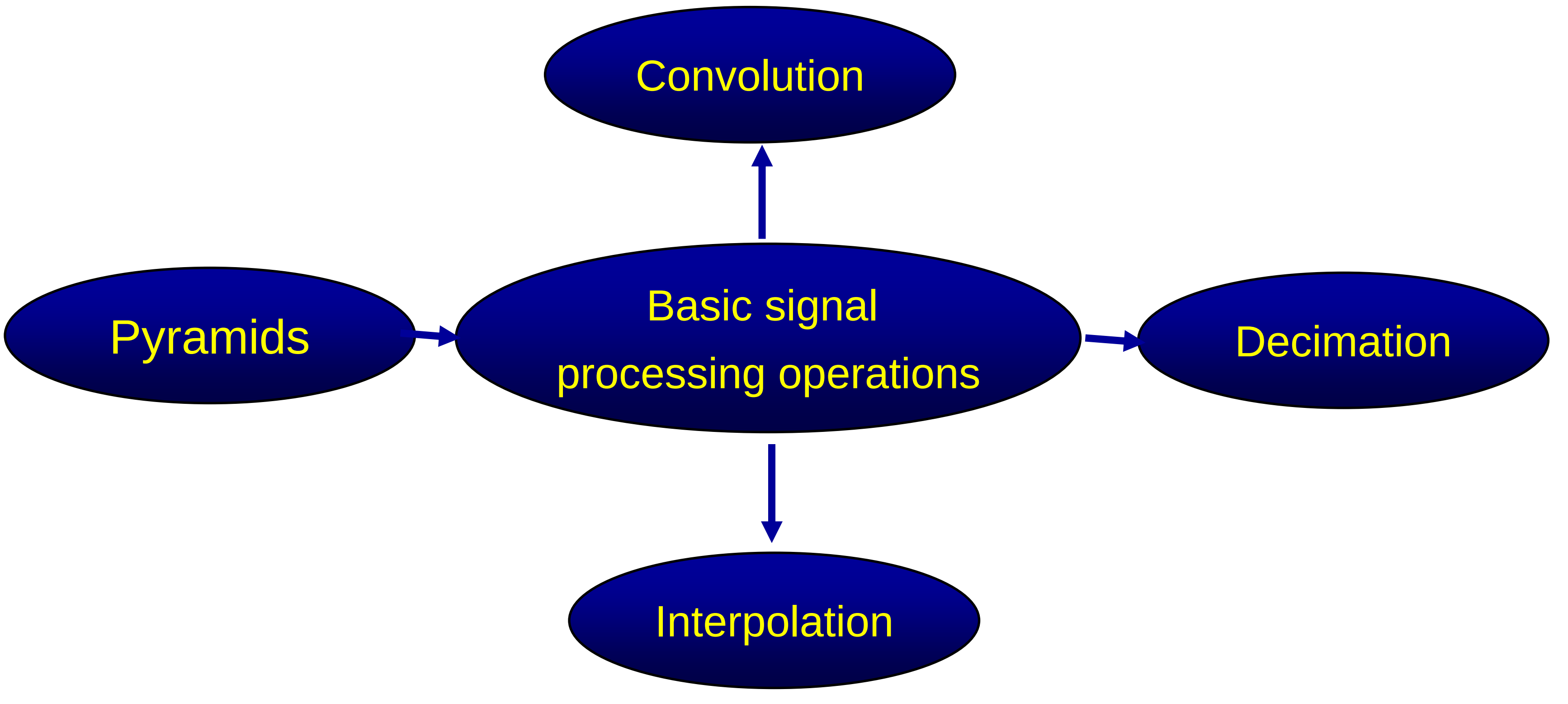
2. Pyramid decomposition

For a given image, the low-pass component can be computed passing original image through the low-pass filter (local mean computation).

Then, the high-pass component is easily found as:

$$\boxed{x_{HP} = x - x_{LP} = x - \bar{x}}$$

2. Basic operations of Pyramid Decomposition



2.1. Interpolation (increase of image size)

64x64

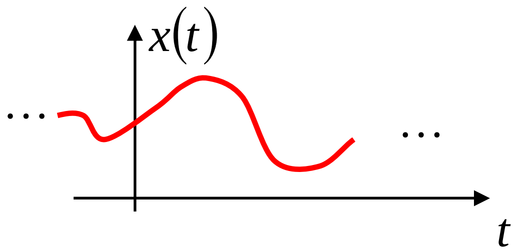
128x128

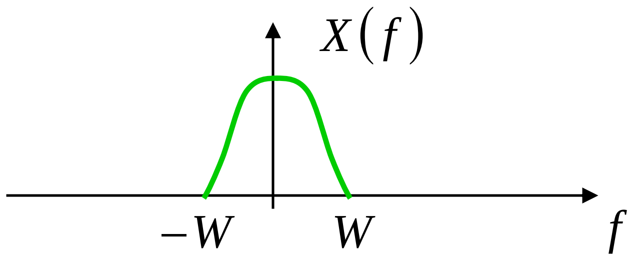
256x256

2.1. Sampling Theory and Fourier Transform: Recall

Continuous Fourier Transform

$$X(f) = \Im\{x(t)\} = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi ft} dt \qquad x(t) = \Im^{-1}\{X(f)\} = \int_{-\infty}^{+\infty} X(f) e^{j2\pi ft} df$$





2.1. Discrete-Space Fourier Transform (DSFT)

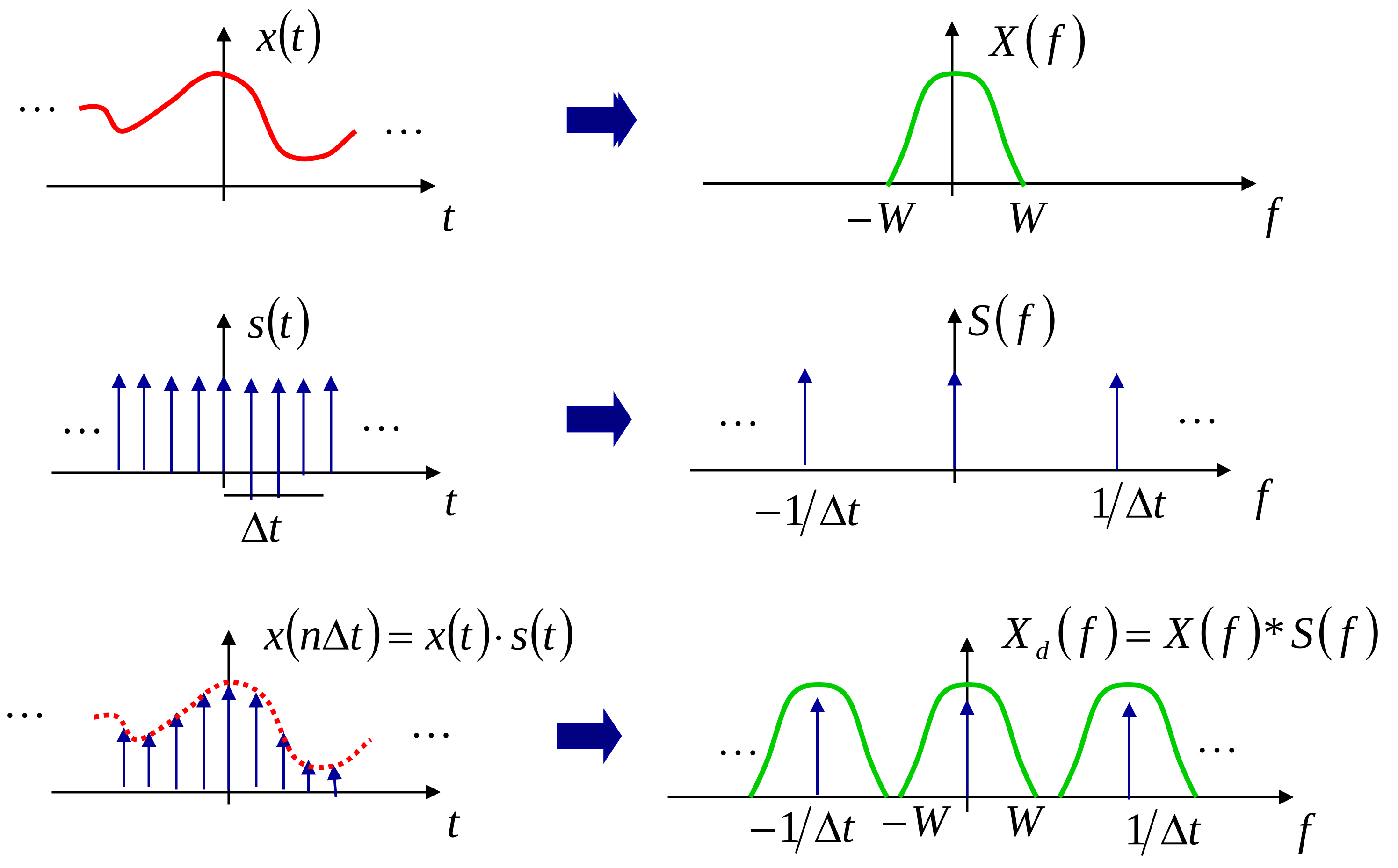
- The spectrum of the discrete signal is a copy of the continuous signal spectrum repeated with the period $1/\Delta t$

$$X_d(f) = \frac{1}{\Delta t}\sum_{m=-\infty}^{+\infty} X\left(f - \frac{m}{\Delta t}\right), \text{ for all } f$$

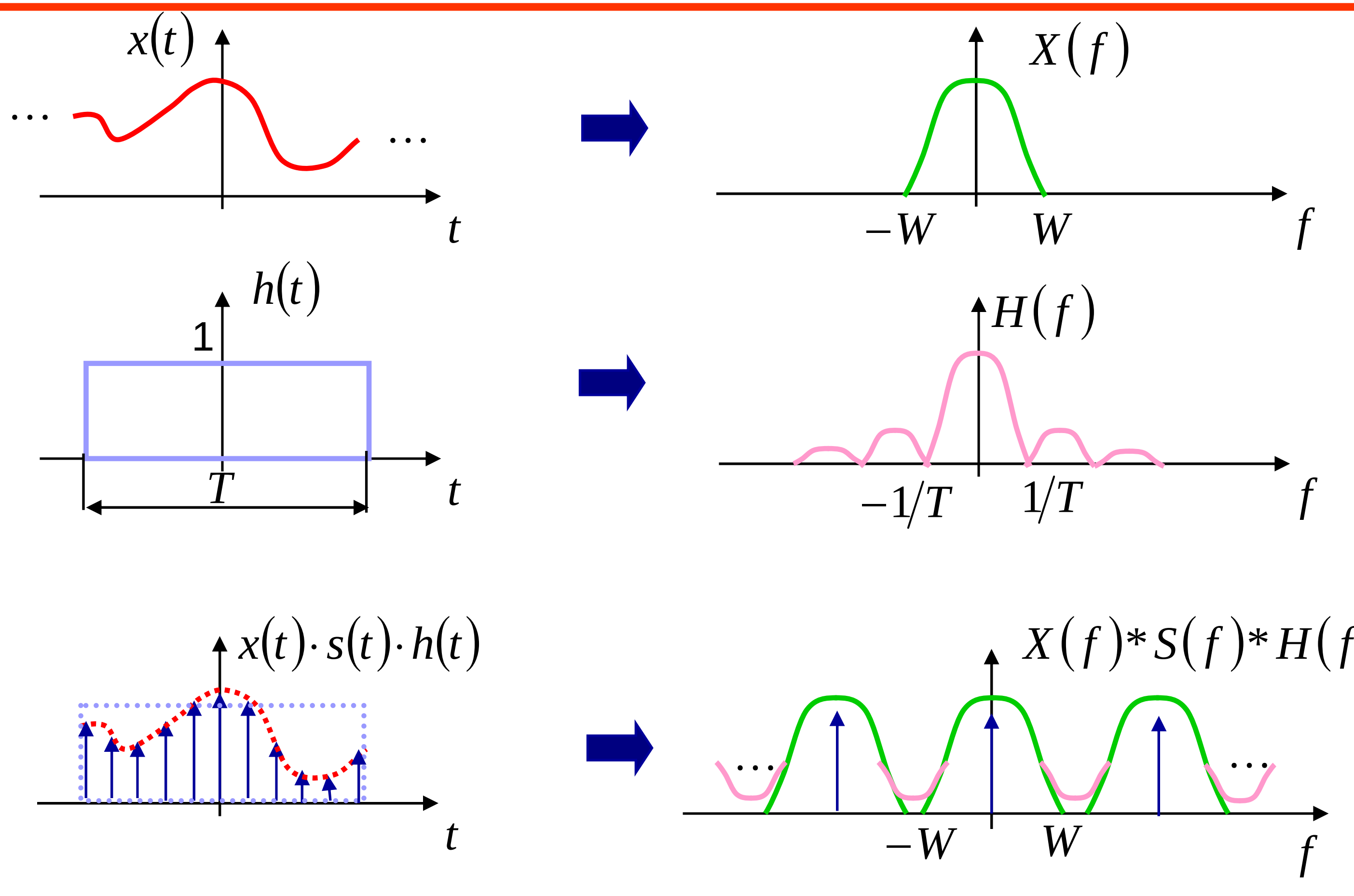
$$X_d(f) = \frac{1}{\Delta t} X(f), \text{ for } |f| < W$$

$$\boxed{X_d(f) = \sum_{n=-\infty}^{+\infty} x(n) e^{-j2\pi f(n\Delta t)} \qquad x(n) = \Delta t \int_{-\infty}^{\infty} X_d(f) e^{j2\pi f(n\Delta t)} df}$$

2.1. Discrete-Space Fourier Transform (DSFT)



Space Limited Discrete Fourier Transform



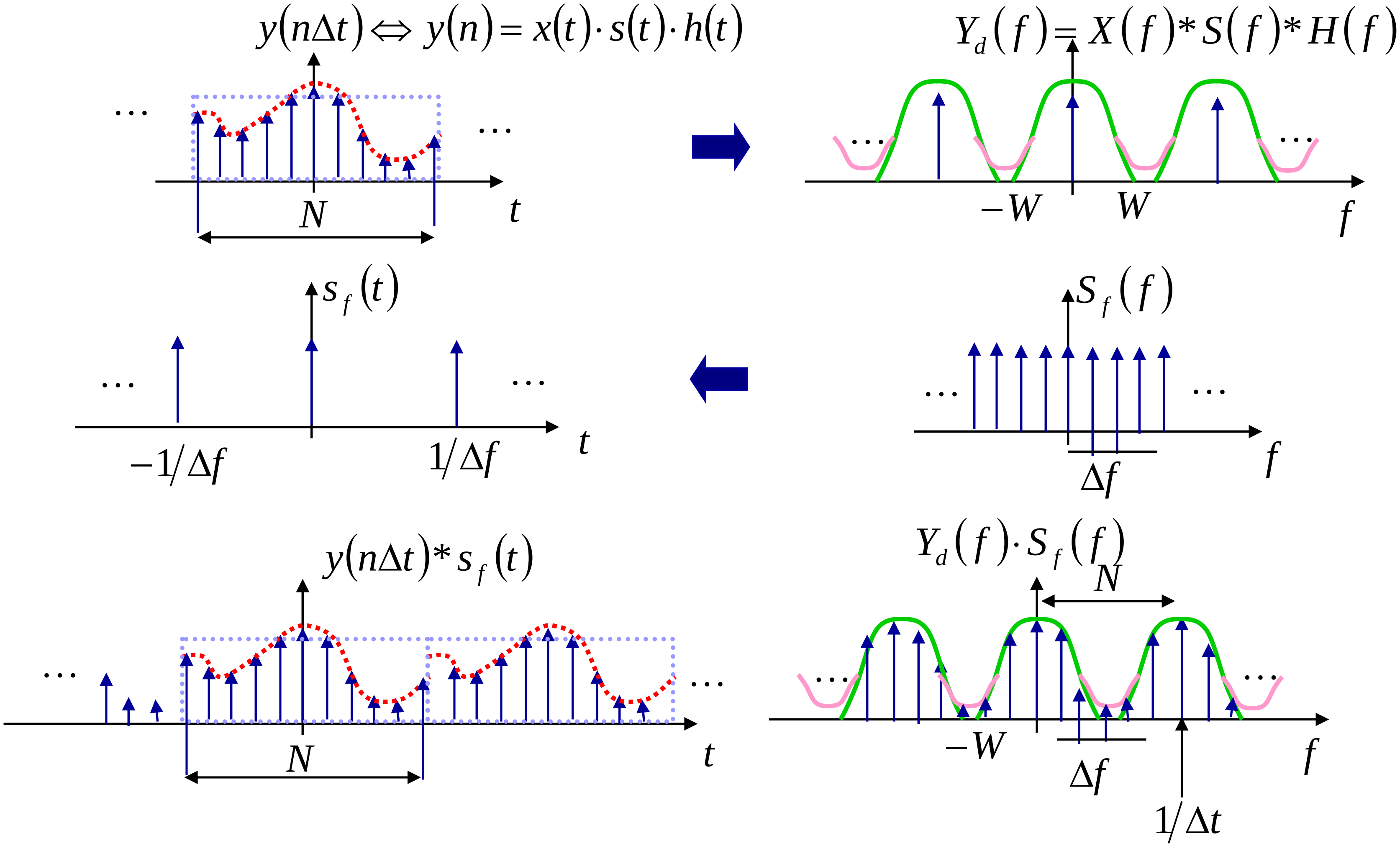
2.1. Discrete Fourier Transform (DFT)

- Signal is also sampled in the frequency domain with step Δf .
- Therefore, the signal will be periodic in the space domain with the period $T = \frac{1}{\Delta f}$.
- Thus, $f = k\Delta f,\ k = 0, \pm 1, \pm 2, \ldots,\quad f = k\frac{1}{T}$ and $T = N\Delta t$

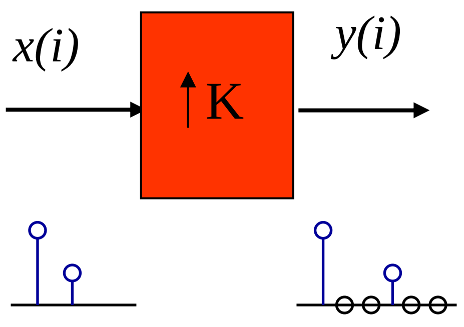
$$\boxed{f = k\frac{1}{N\Delta t}}$$

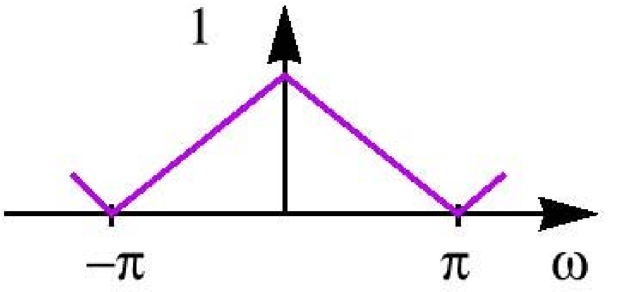
$$\boxed{X(k) = X_d(k\Delta f) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi kn}{N}\frac{\Delta t}{\Delta t}} = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi kn}{N}} \qquad x(n) = \frac{1}{N}\sum_{k=0}^{N-1} X(k) e^{j\frac{2\pi kn}{N}}}$$

Discrete Fourier Transform



2.1. Interpolation: upsampling



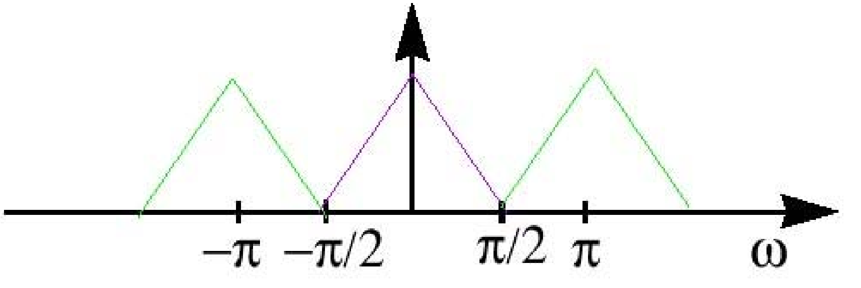


Factor of K=2

$$\boxed{Y_d\left(e^{i\omega}\right) = X_d\left(e^{i2\omega}\right)}$$

Factor of K

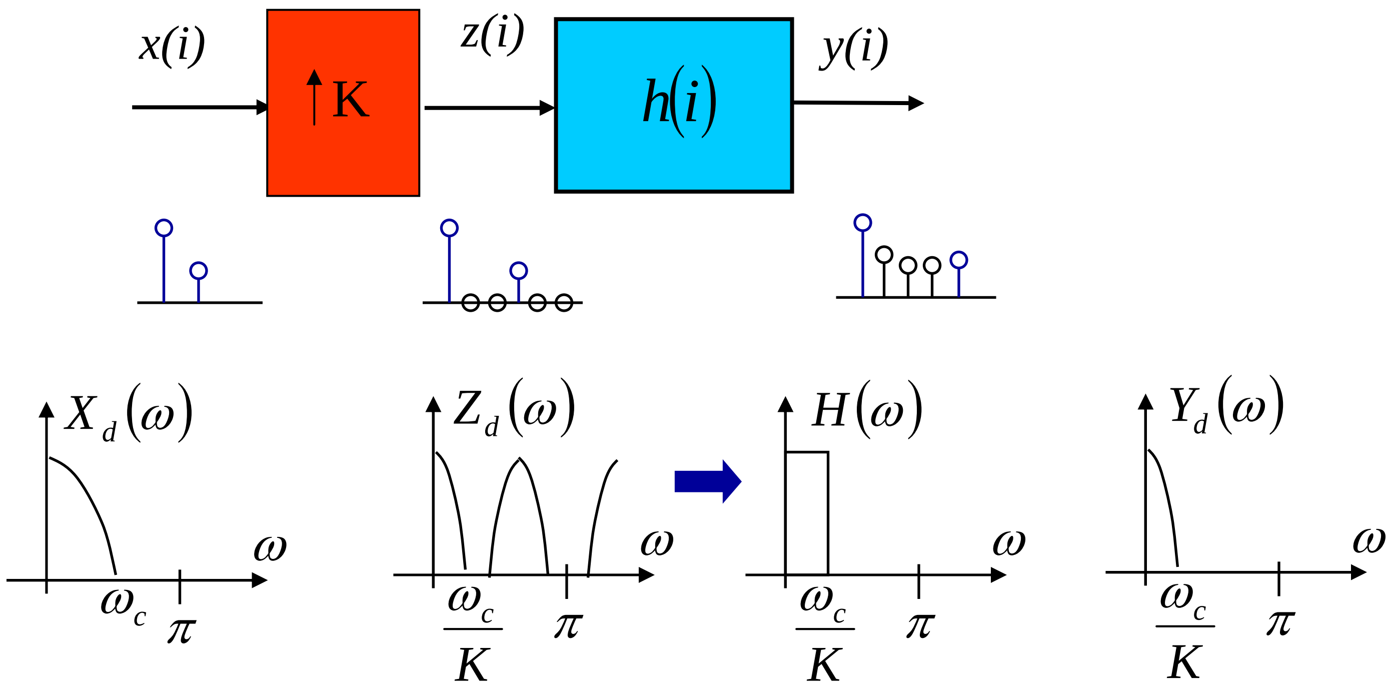
$$\boxed{Y_d\left(e^{i\omega}\right) = X_d\left(e^{iK\omega}\right)}$$



2.1. Interpolation: fundamentals

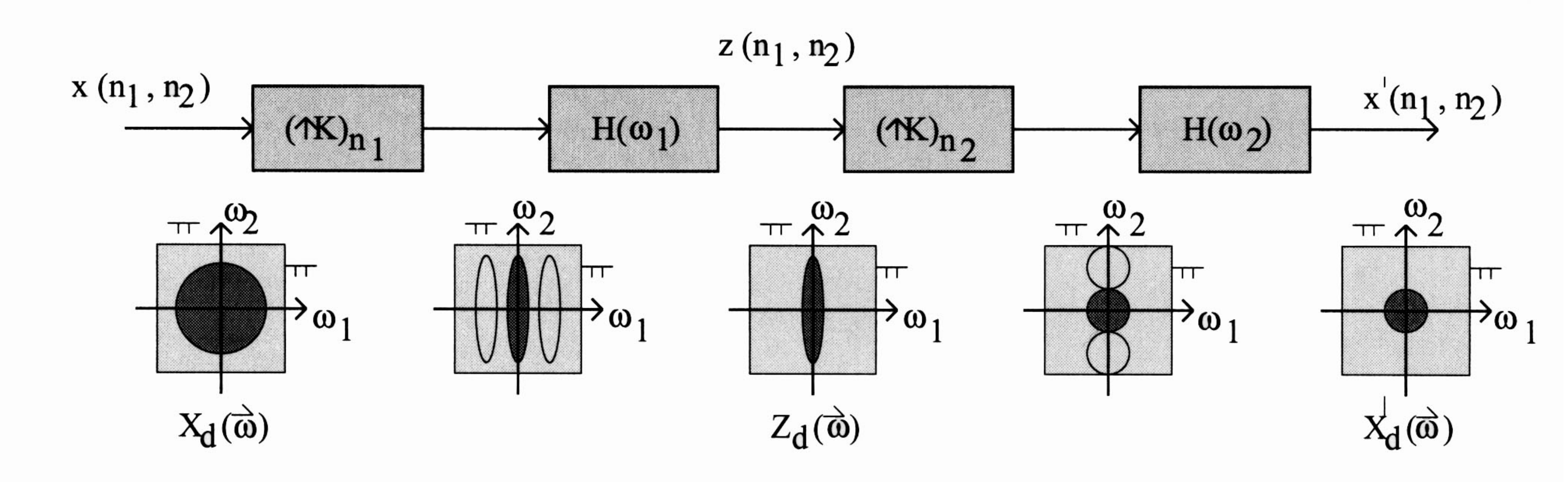
Ideal 1-D interpolator is the cascade of an upsampler and an ideal low-pass filter.

Interpolator inside

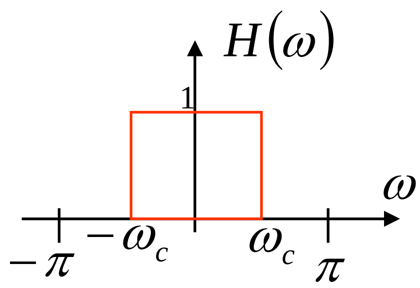


2.1. Interpolation: 2-D case

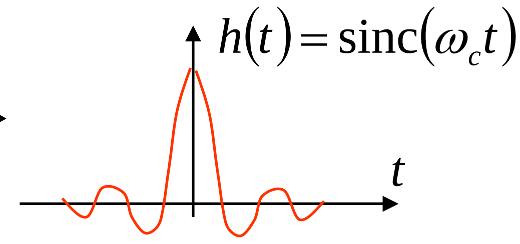
Ideal 2-D interpolator can be implemented using separable sequential 1-D interpolation.



2.1. Interpolation: ideal low-pass filter



Impulse response



Problem: infinite oscillating function

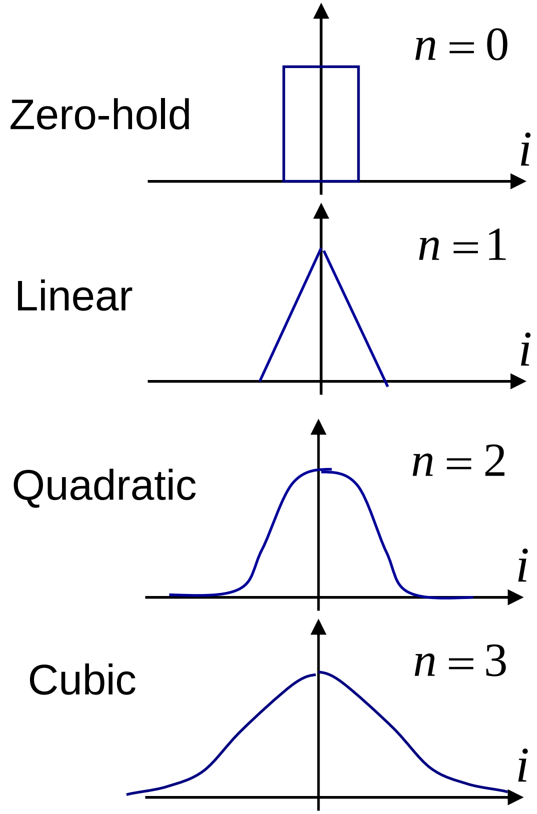
Solution: finite support functions

Splines

Polynomials

2.1. Interpolation: low-pass filtering - splines

We will concentrate on B-splines of degree n.



M. Unser's Group, EPFL:

http://bigwww.epfl.ch/

2.1. Interpolation: low-pass filtering - splines

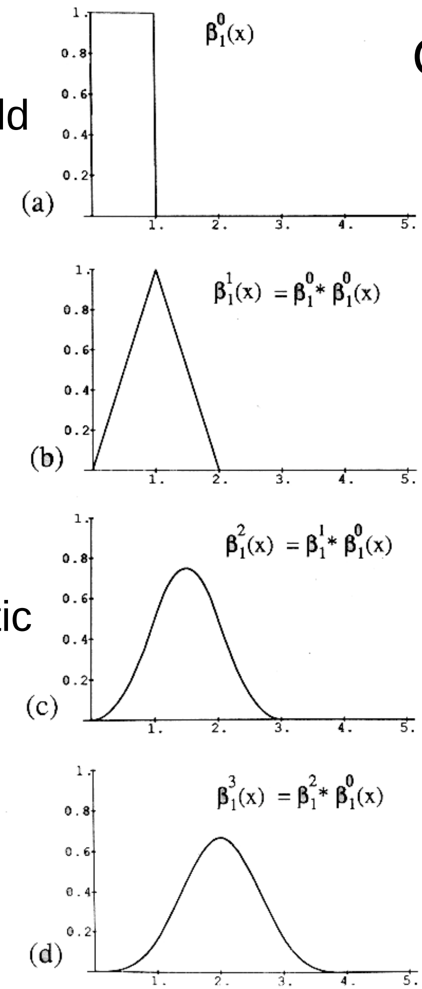
Convolution properties of continuous B -splines

Zero-hold

Linear

Quadratic

Cubic

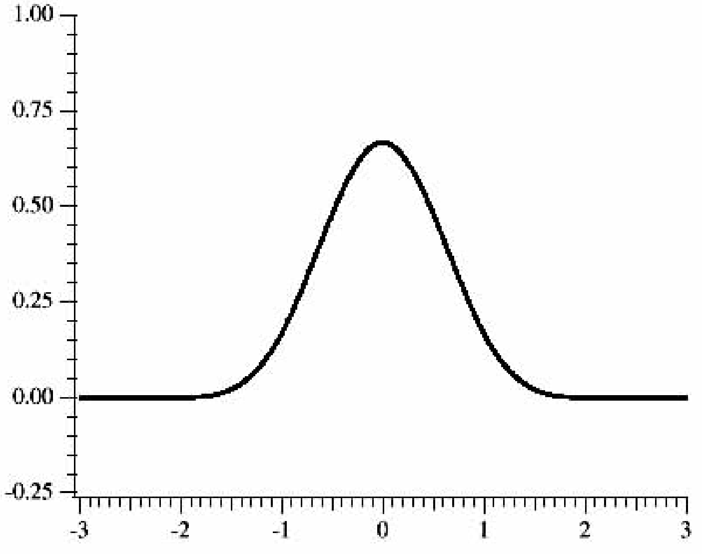


Note:

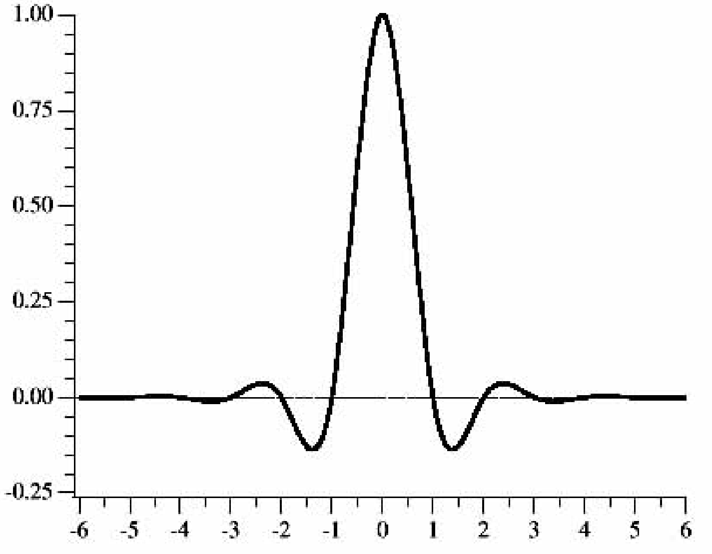
- Only zero-hold and linear (first order) splines can be directly applied as interpolators.
- The rest splines should be first converted to cardinal interpolation kernel.

M. Unser, A. Aldroubi, M. Eden, IEEE PAMI, Vol13, No3, 1991.

2.1. Interpolation: Cubic-interpolants

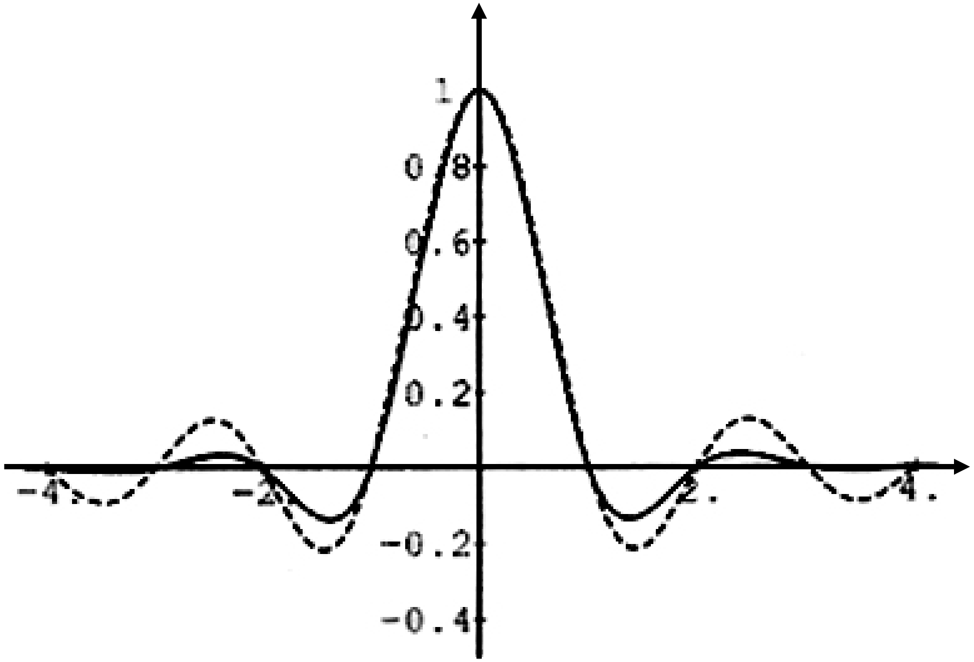


Cubic B-spline (function shape)



Equivalent Interpolant

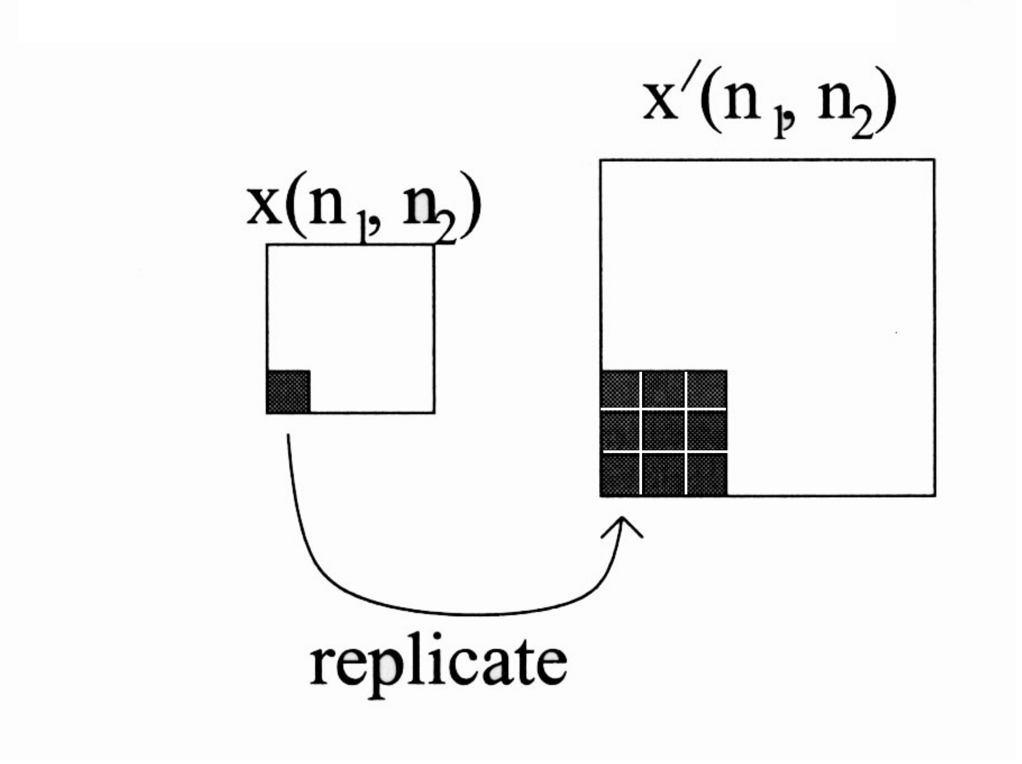
2.1. Interpolation: Sinc and Cubic-interpolants



Comparison of the cardinal cubic spline function (solid line) and ideal sinc interpolation kernel (dashed line).

2.1. Interpolation: ideal low-pass filter

The simplest solution: replication.



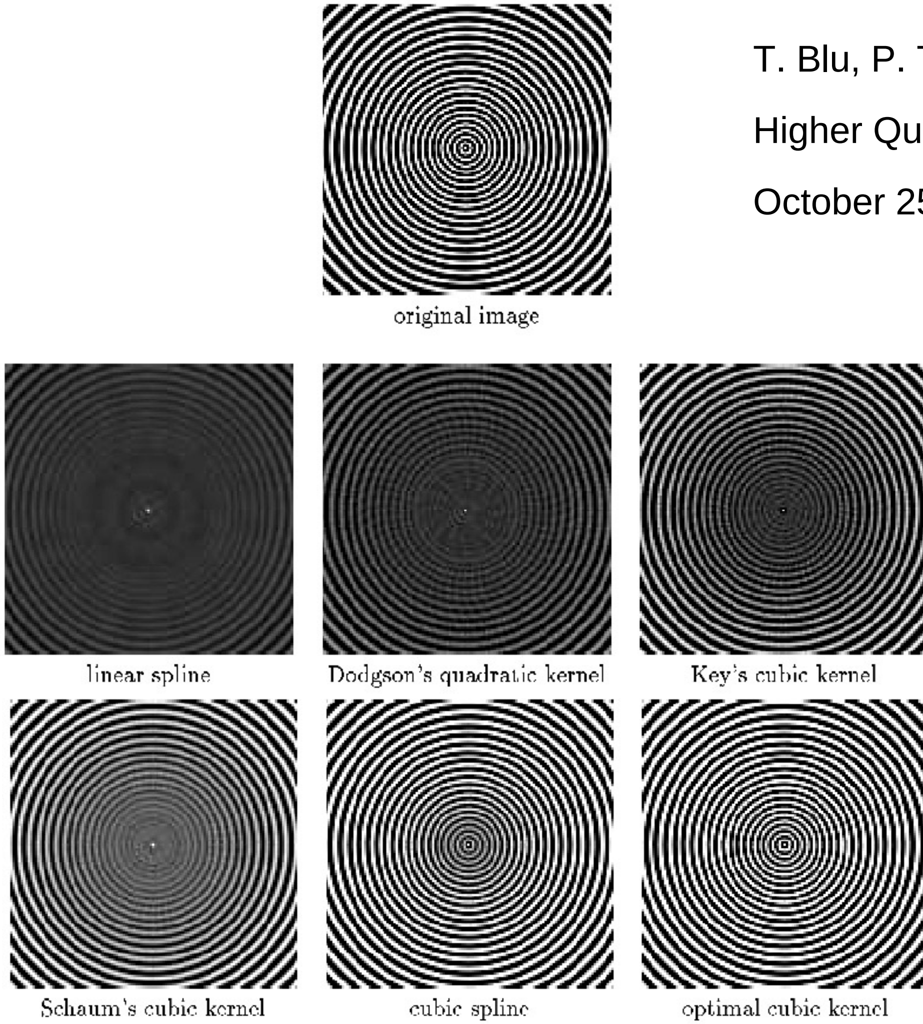
2.1. Interpolation: Different systems

Interpolation families:

- B-splines
- Dodgson’s function
- Keys’s function
- o-Moms
- Schaum’s function
- and many others

2.1. Interpolation: Different systems

T. Blu, P. Thévenaz, M. Unser, Generalized Interpolation: Higher Quality at no Additional Cost, ICIP'99, Kobe, Japan, October 25-28, 1999, vol. III, pp. 667-671.



2.1. Interpolation: Different systems


original image


linear spline


Dodgson's quadratic kernel


Key's cubic kernel


Schaum's cubic kernel


cubic spline


optimal cubic kernel

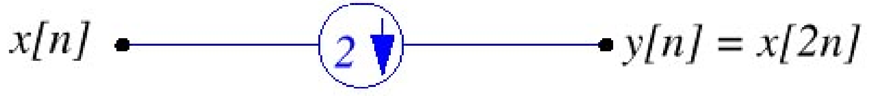
2.2. Decimation (decrease of size)

256x256

128x128

64x64

2.2. Downsampling: coordinate domain

- 

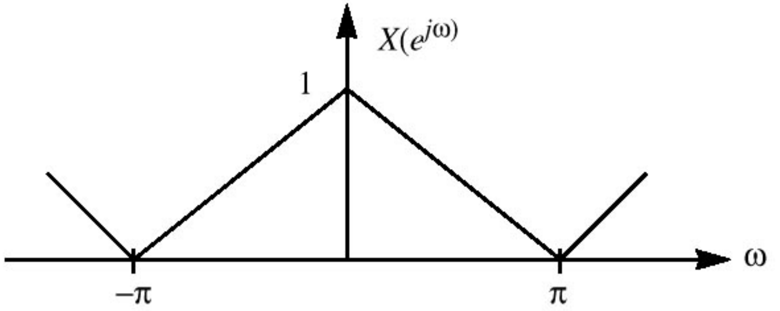
$$\ldots x[0]x[1]x[2]\ldots \qquad \ldots x[0]x[2]x[4]\ldots$$

- $y = D_2 \cdot x$

$$\begin{bmatrix} \cdots \\ y(-1) \\ y(0) \\ y(1) \\ \cdots \end{bmatrix} = \begin{bmatrix} . & . & . & . & . & . \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ . & . & . & . & . & . \end{bmatrix} \cdot \begin{bmatrix} \cdots \\ x(-1) \\ x(0) \\ x(1) \\ x(2) \\ \cdots \end{bmatrix}$$

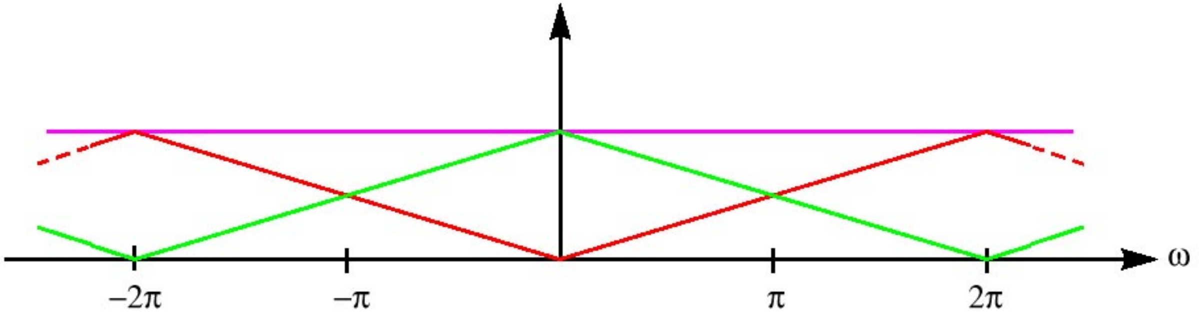
2.2. Decimation: Fourier domain

$$\boxed{Y\left(e^{i\omega}\right)=\frac{1}{2}\left[X\left(e^{i\frac{\omega}{2}}\right)+X\left(e^{i\frac{\omega-2\pi}{2}}\right)\right]\Rightarrow\frac{1}{2}\left[X\left(z^{\frac{1}{2}}\right)+X\left(-z^{\frac{1}{2}}\right)\right]=Y(z)}$$



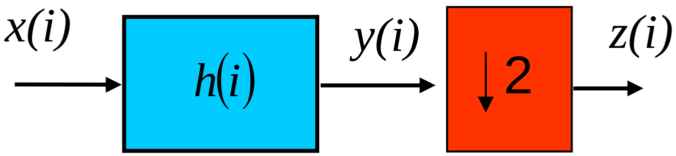
$$X(z)=\sum_{n=-\infty}^{+\infty}x(n)z^{-n}$$

$$z=e^{j\omega}$$



2.2. Decimation (reduction of size)

Decimation of signal by a factor of 2.



Convolution: $$y(i) = x(i) * h(i) = \sum_k x(k) h(i-k)$$

Downsampler: $$z(i) = y(2i)$$

Combining: $$\boxed{z(i) = \sum_k x(k) h(2i-k) = \sum_k x(2i-k) h(k)}$$

2.2. Decimation

Properties of filter $h(i)$:

- low-pass filter;
- typically symmetric (efficient computation);
- length of support between 3 and 20;
- for 0-frequency (DC) has response equal to 1:

$$\sum_i h(i) = 1$$

<u>Examples:</u>

- 3-tap FIR filter $h(i) = (1/4, 1/2, 1/4)$
- (2L+1)-tap truncated Gaussian filter $h(i) = Ce^{-i^2/(2\sigma^2)}, |i| \le L$

Normalizing constant

2.2. Pyramid decomposition: conclusions

Advantage: sparse representation in Laplacian pyramid.

Drawbacks of pyramid approach:

- Overcompleteness: the number of image samples after transform is larger than the number before transform.

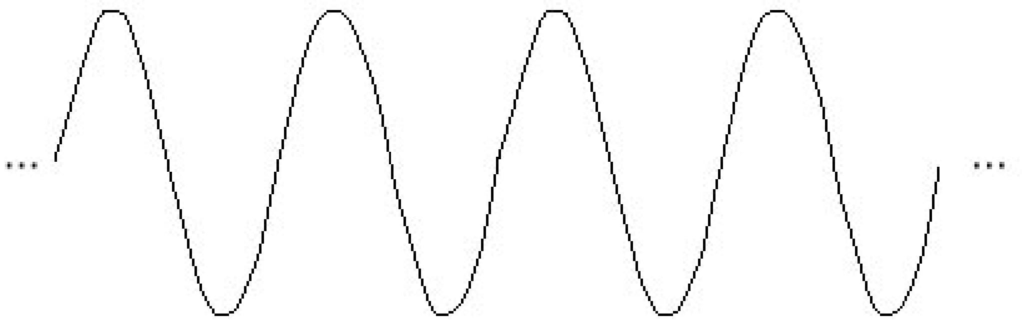
- The pyramid decomposition captures the multiresolution image representation. However, orientation is not included.

3. Wavelets

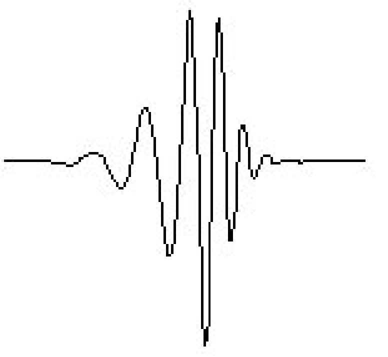
Advantages:

- multiresolution and sparse image representation;
- complete representation;
- capture different spatial orientations;
- relate to the Human Visual System.

3. Wavelets: basis functions



Sine Wave

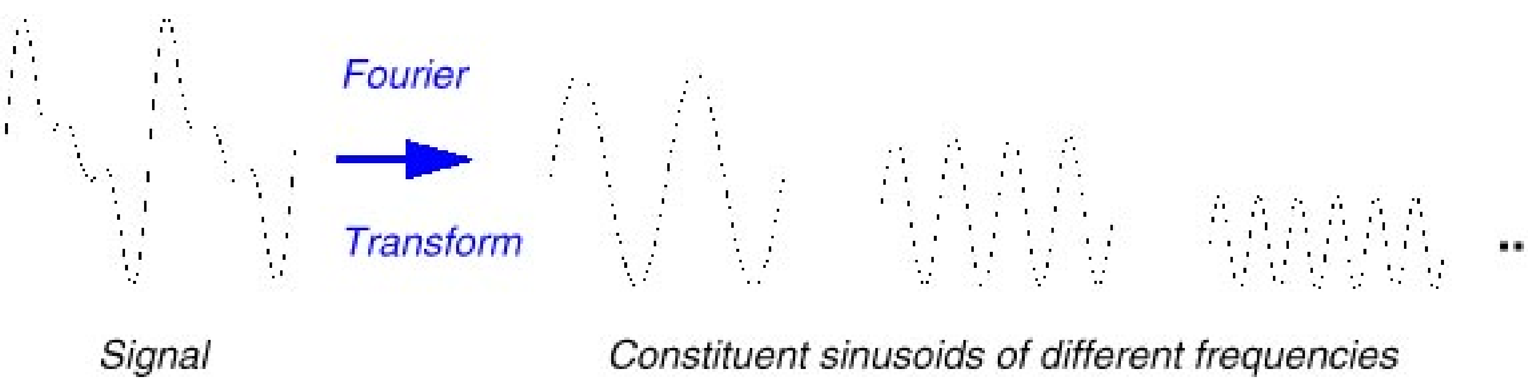


Wavelet (db10)

3. Wavelets: basic concept

<u>Fourier Transform</u>:

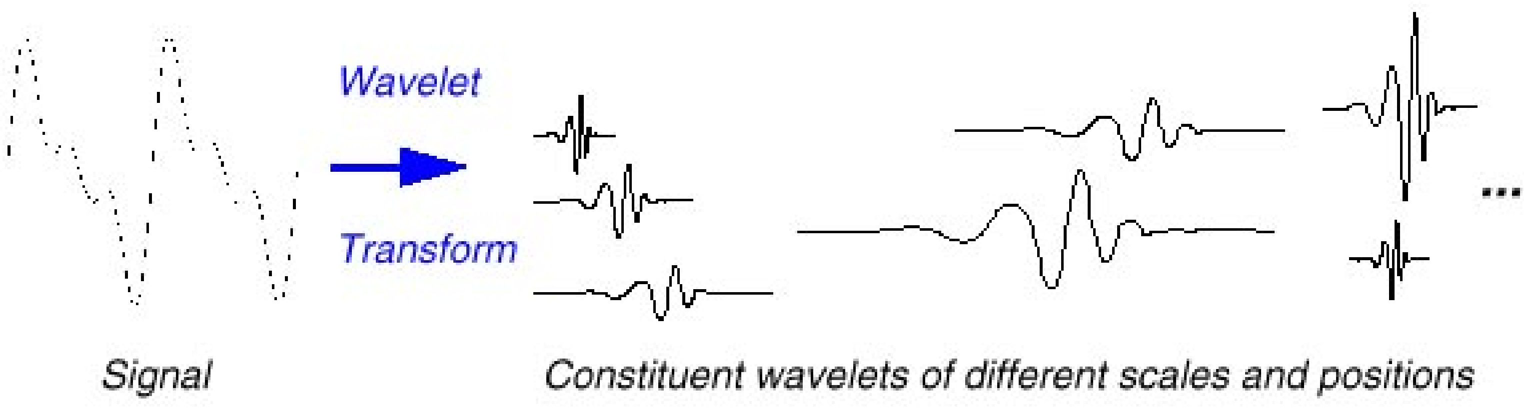
$$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt$$



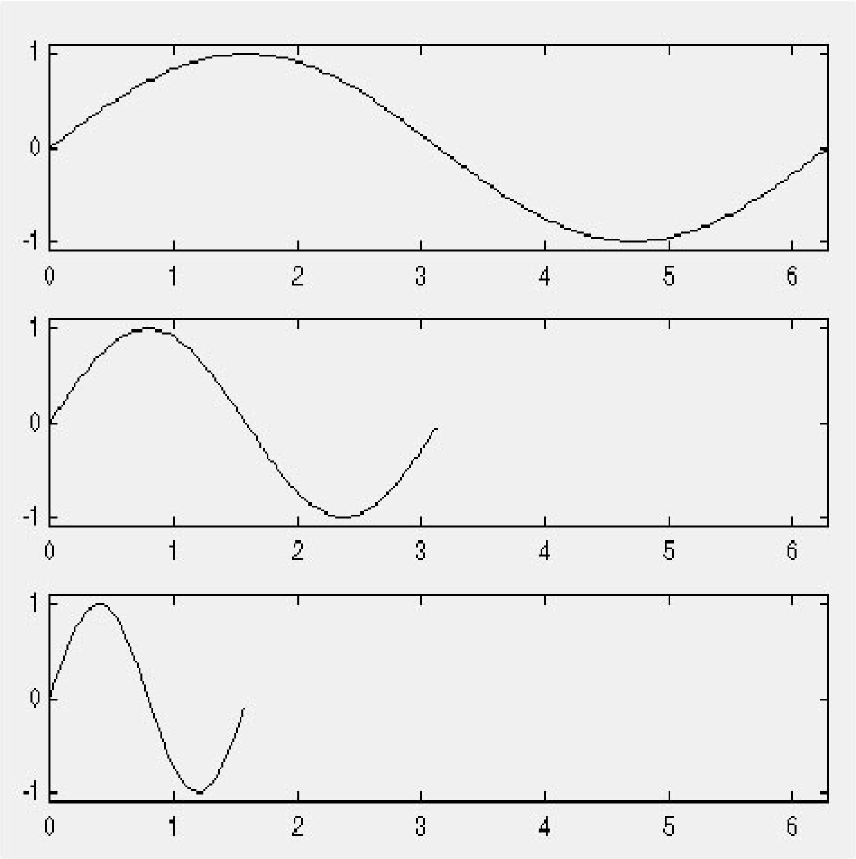
3. Wavelets: basic concept

Continuous Wavelet Transform:

$$C(scale, position) = \int_{-\infty}^{\infty} x(t)\psi(scale, position)dt$$



3. Wavelets: scaling



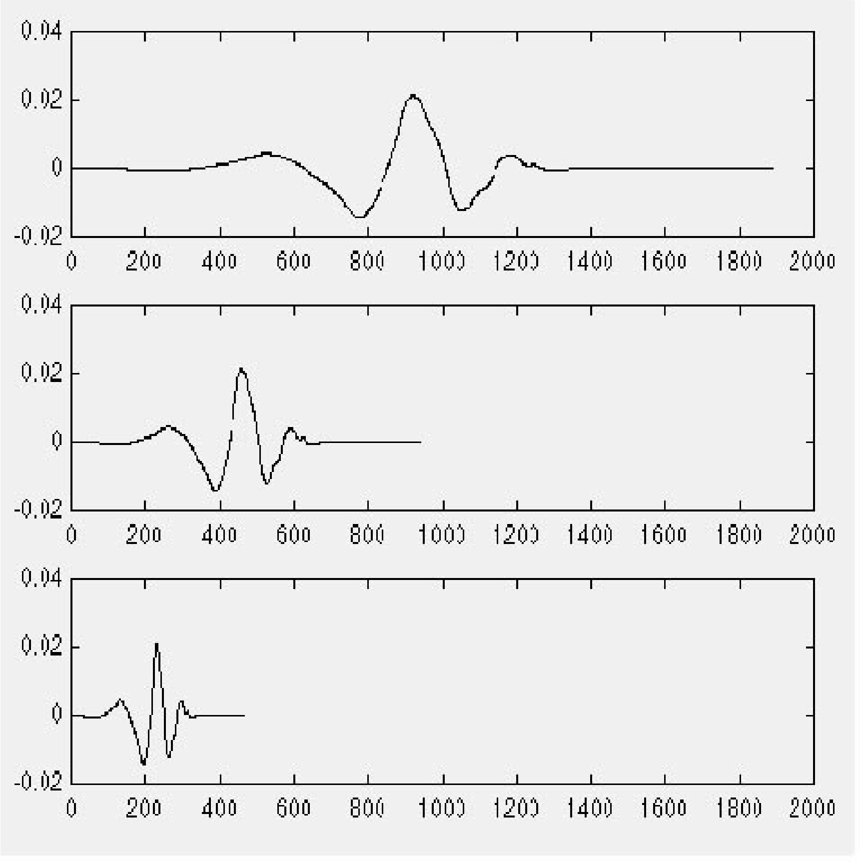
a: scaling factor

$$f(t) = \sin(t); \quad a = 1$$

$$f(t) = \sin(2t); \quad a = \frac{1}{2}$$

$$f(t) = \sin(4t); \quad a = \frac{1}{4}$$

3. Wavelets: scaling



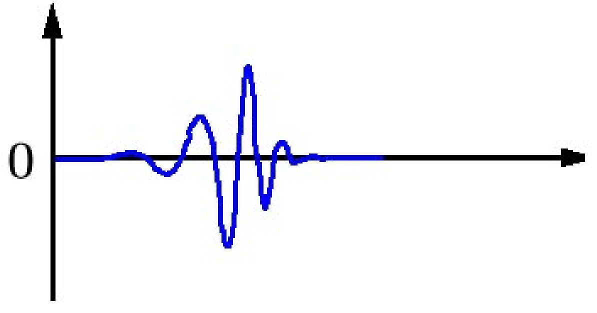
$$f(t) = \psi(t) \quad ; \quad a = 1$$

$$f(t) = \psi(2t) \quad ; \quad a = \frac{1}{2}$$

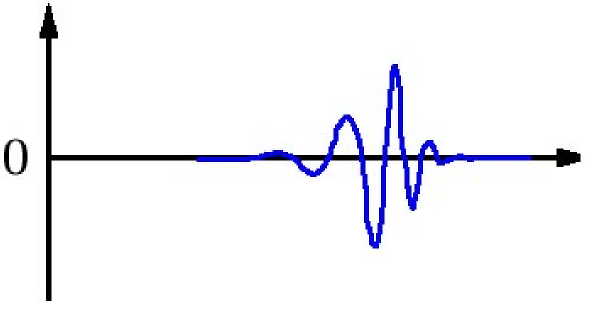
$$f(t) = \psi(4t) \quad ; \quad a = \frac{1}{4}$$

3. Wavelets: shifting

k: shift factor

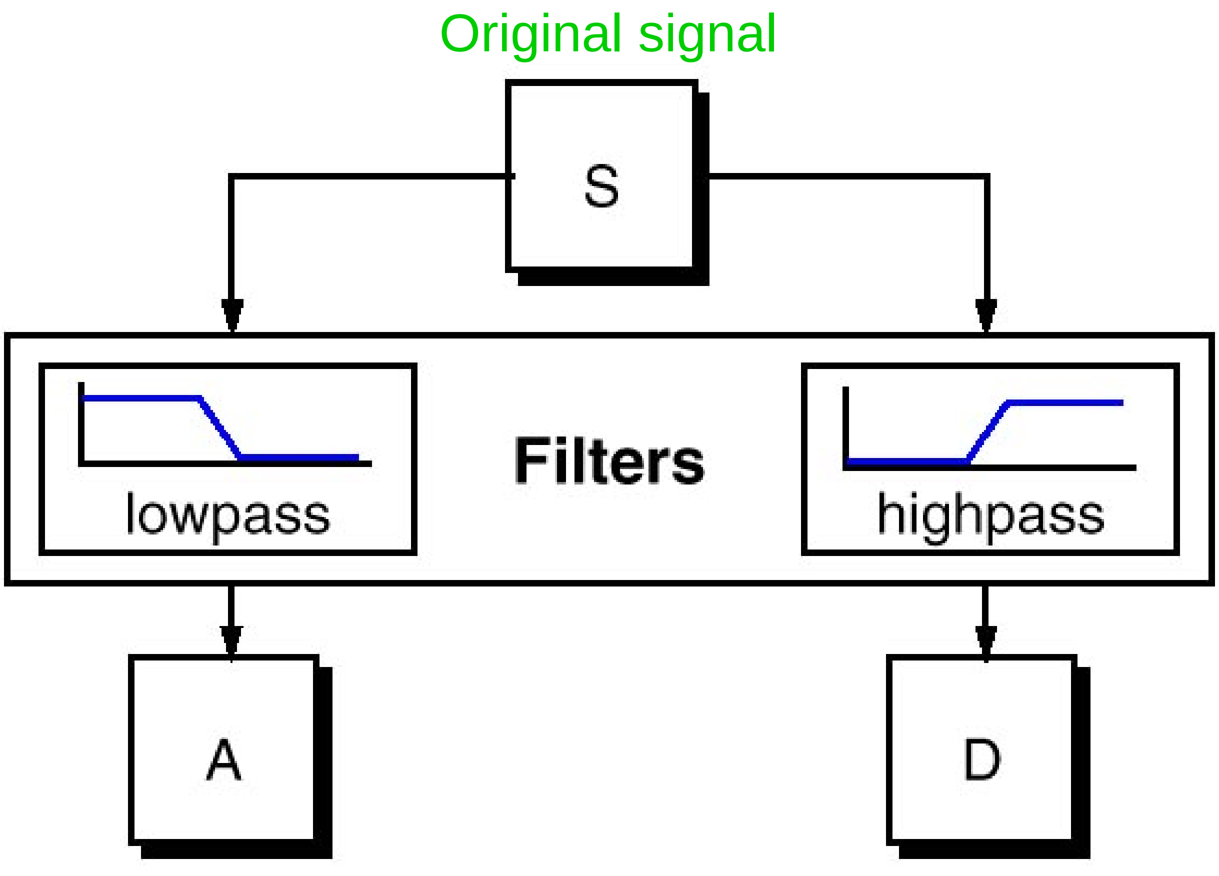


Wavelet function
$\psi(t)$



Shifted wavelet function
$\psi(t-k)$

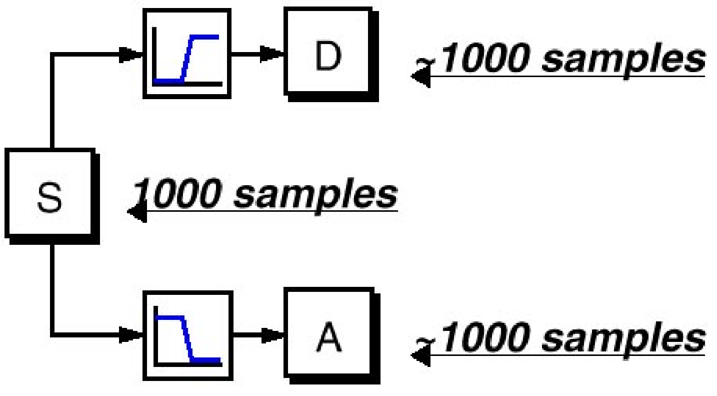
3. Discrete Wavelet Transform

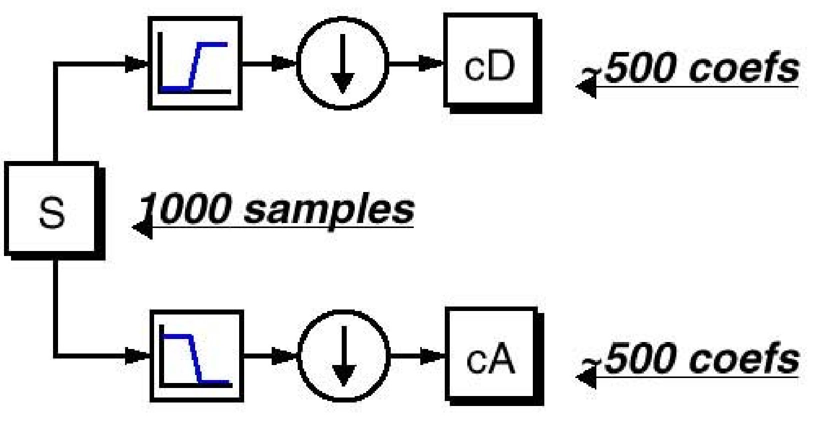


Problem: overcomplete presentation (like in pyraminds)

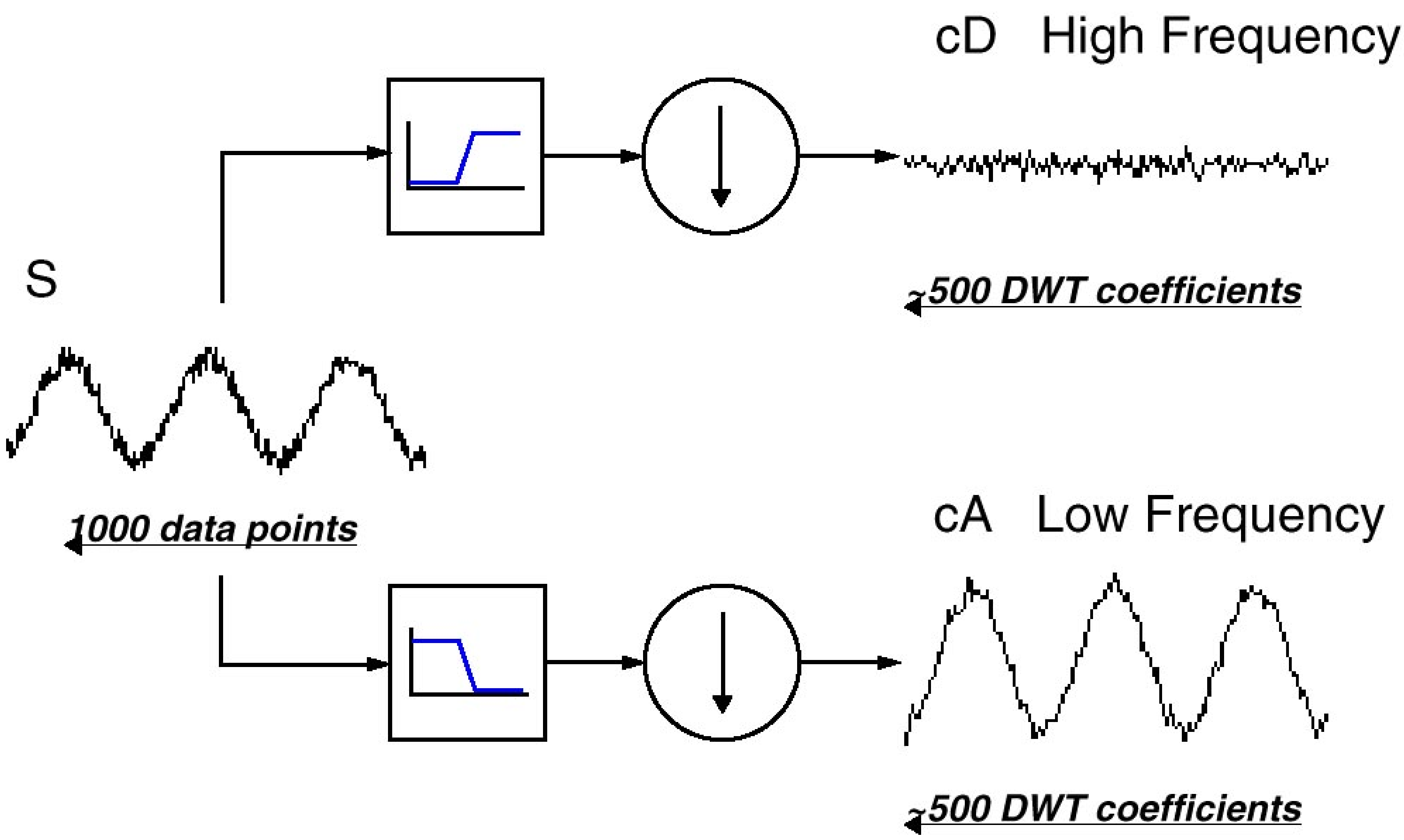
3. Discrete Wavelet Transform

Downsampling

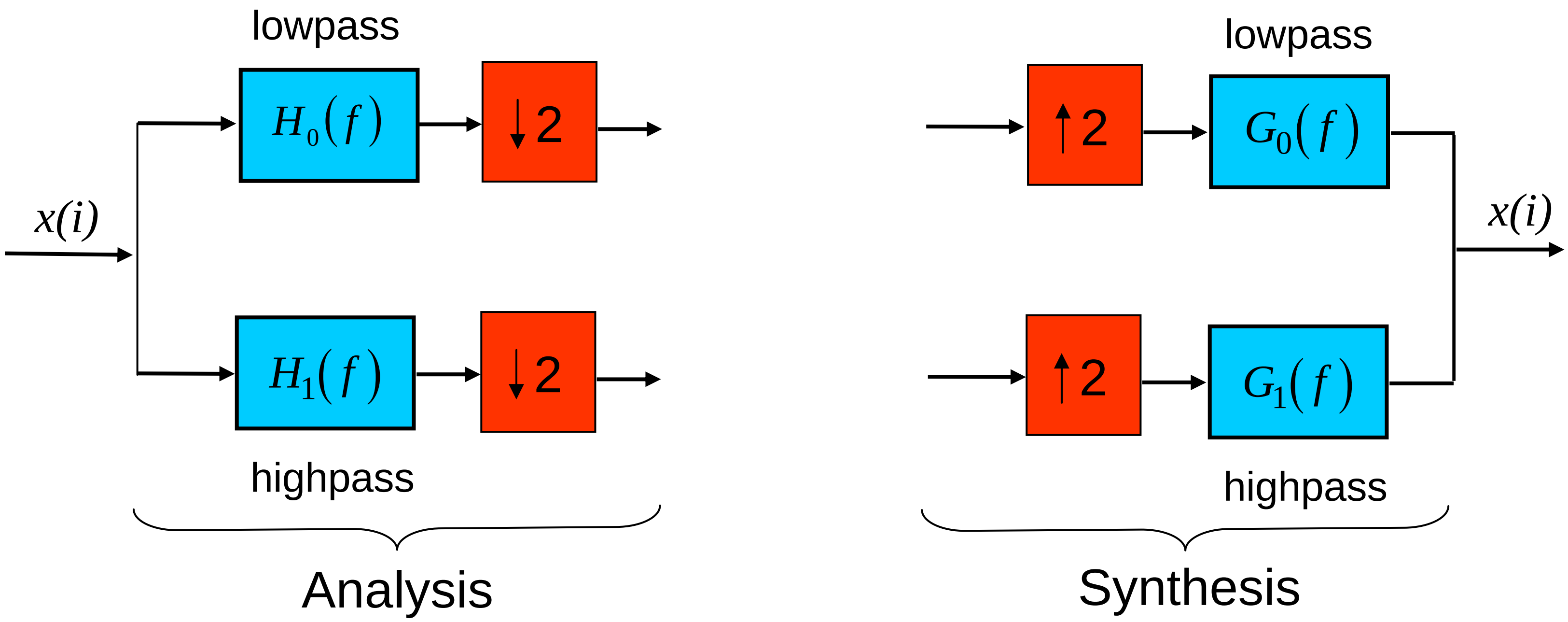




3. Discrete Wavelet Transform: Example



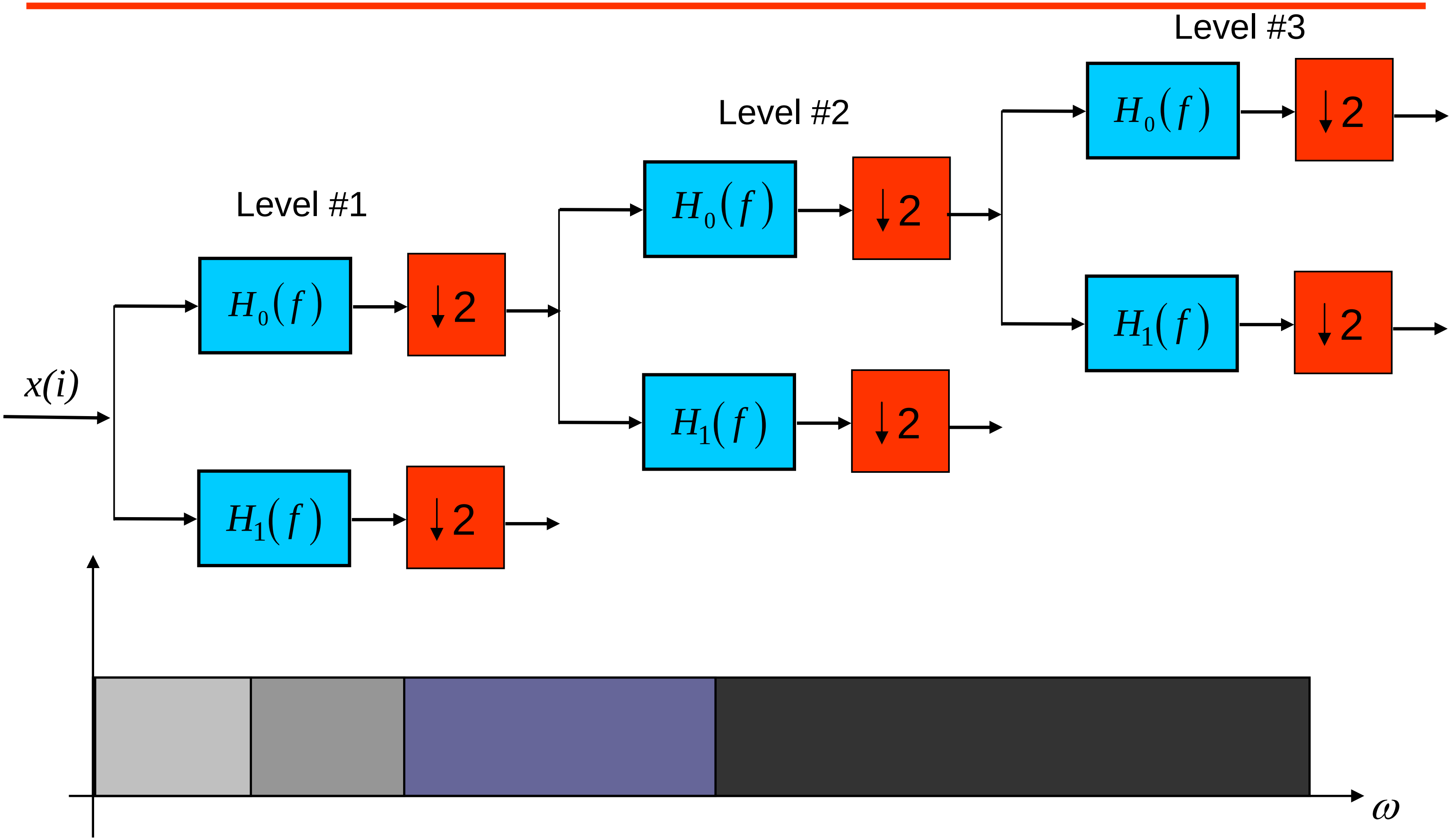
3.1. Filter banks



Low-pass filters: $H_0(f)$ and $G_0(f)$

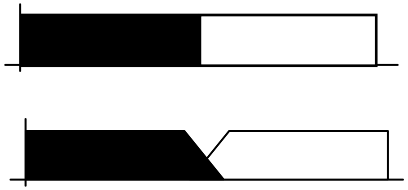
High-pass filters: $H_1(f)$ and $G_1(f)$

3.1. Filter banks: pyramidal decomposition



3.1. Filter banks: pyramidal decomposition

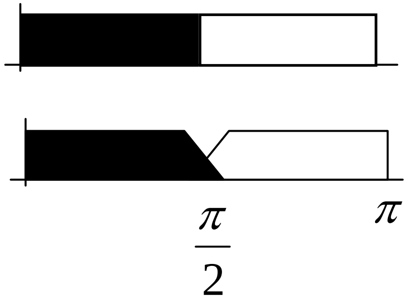
- In order to decompose a full-band, 1-D signal into two subbands, the analysis filter bank would ideally consist of a lowpass and higpass filter set with frequency responses that:
  - are nonoverlapping, but contiguous (connected);
  - have unity gain over their bandwidths.



- However, ideal filters are unrealizable in practice.
- Therefore, it is necessary to use filters with overlapping responses in order to prevent frequency gaps in signals represented by the subbands.
- The problem with overlapping filters is that aliasing is introduced when the subbands are downsampled.

3.1. Filter banks: pyramidal decomposition

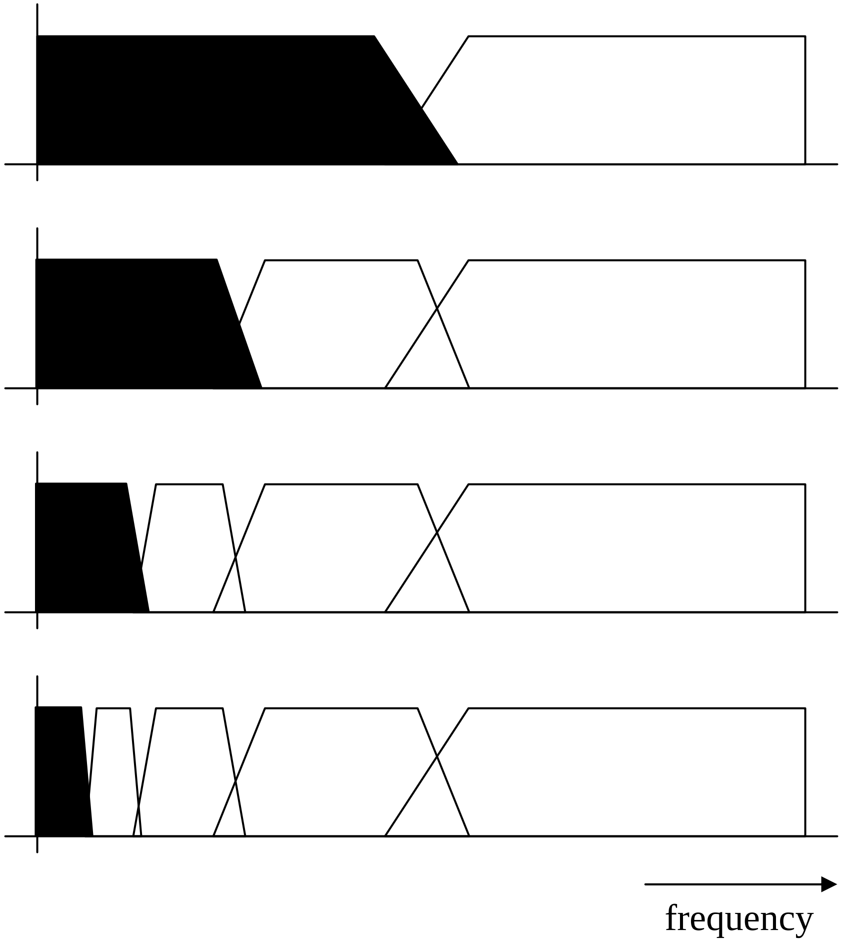
- To overcome this problem, analysis and synthesis filters are designed using quadrature mirror filters (QMFs).



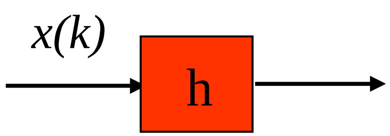
P.P. Vaidyanathan, Quadrature mirror filter banks, M-band extensions and perfect-reconstruction technique, IEEE ASSP Magazine, 4(3), 1987.

- The name QMFs arises from the fact that the filters exhibit mirror symmetry about $\frac{\pi}{2}$.
- The main idea of QMFs: to allow aliasing to be introduced by using overlapping filters for analysis bank and then design the synthesis filters in such a way that any aliasing is exactly cancelled out in the reconstruction (synthesis) process.

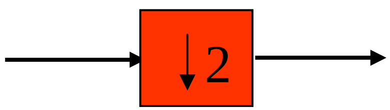
3.1. Filter banks: pyramidal decomposition



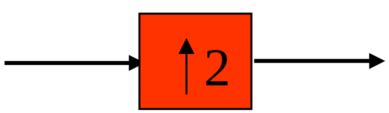
3.1. Filter banks: fundamentals



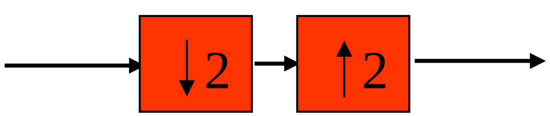
$$(h*x)(k)=\sum_{l\in Z}h(l)x(k-l) \overset{Z}{\longleftrightarrow} Y(z)=H(z)X(z)$$



$$[x]_{\downarrow 2}(k)=x(2k) \overset{Z}{\longleftrightarrow} \frac{1}{2}\left[X\left(z^{1/2}\right)+X\left(-z^{1/2}\right)\right]$$



$$[x]_{\uparrow 2}(k)=\begin{cases}0, k\ odd\\ x(l), k=2l\ even\end{cases} \overset{Z}{\longleftrightarrow} X\left(z^{2}\right)$$



$$[x]_{\downarrow 2\uparrow 2}(k)=x(k) \overset{Z}{\longleftrightarrow} \frac{1}{2}\left[X(z)+X(-z)\right]$$

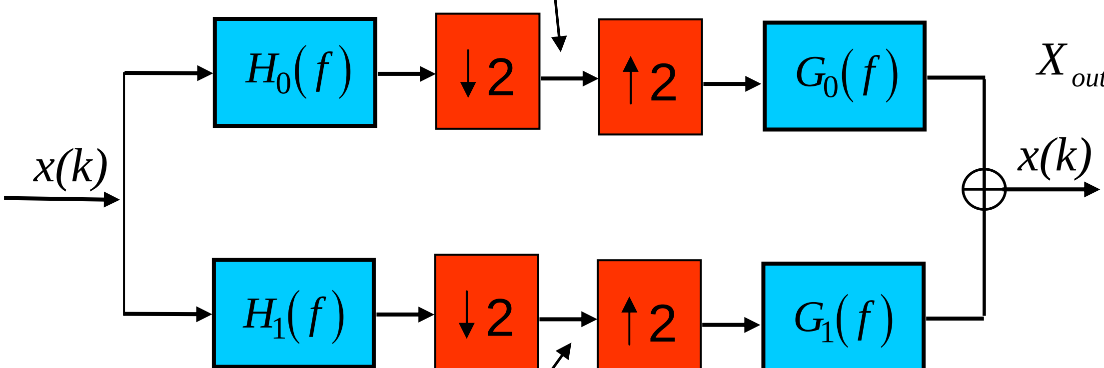
3.1. Filter banks: Perfect Reconstruction

One can design orthonormal analysis and synthesis filter banks.

Orthonormality implies that the energy of samples is preserved under the transform.

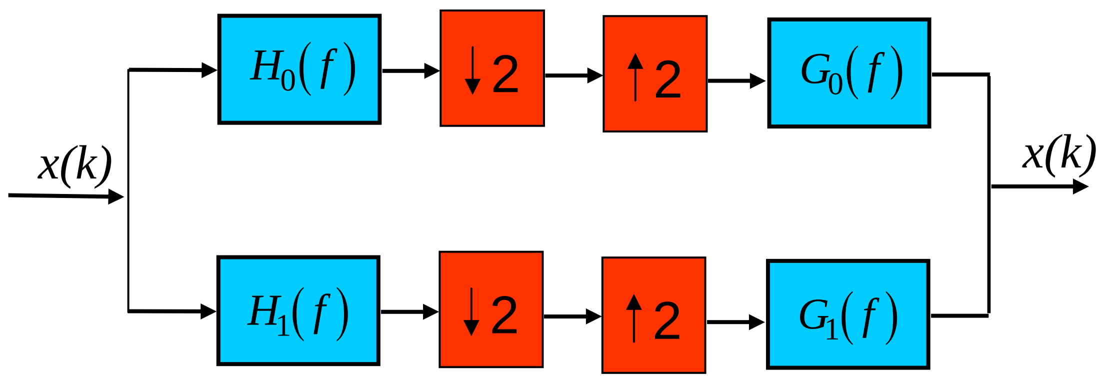
$$Y_0(z)=\frac{1}{2}\left[X\left(z^{\frac{1}{2}}\right)H_0\left(z^{\frac{1}{2}}\right)+X\left(-z^{\frac{1}{2}}\right)H_0\left(-z^{\frac{1}{2}}\right)\right]$$

$$X_{out}(z)=G_0(z)Y_0\left(z^2\right)+G_1(z)Y_1\left(z^2\right)$$



$$Y_1(z)=\frac{1}{2}\left[X\left(z^{\frac{1}{2}}\right)H_1\left(z^{\frac{1}{2}}\right)+X\left(-z^{\frac{1}{2}}\right)H_1\left(-z^{\frac{1}{2}}\right)\right]$$

3.1. Filter banks: Perfect Reconstruction



$$X_{out}(z)=\frac{1}{2}G_0(z)\left[X(z)H_0(z)+X(-z)H_0(-z)\right]+\frac{1}{2}G_1(z)\left[X(z)H_1(z)+X(-z)H_1(-z)\right]=X(z)$$

3.1. Perfect Reconstruction Condition

$$\boxed{X_{out}(z) = X(z)}$$

$$X_{out}(z) = \frac{1}{2} X(z) \underbrace{\left[H_0(z)G_0(z) + H_1(z)G_1(z)\right]}_{2} + \frac{1}{2} X(-z) \underbrace{\left[H_0(-z)G_0(z) + H_1(-z)G_1(z)\right]}_{0} = X(z)$$

(I) $H_0(z)G_0(z) + H_1(z)G_1(z) = 2$ (distortion free)

(II) $H_0(-z)G_0(z) + H_1(-z)G_1(z) = 0$ (aliasing free)

Wavelet transform design:

Construct 4 filters such that conditions I and II are satisfied.

3.1. Filter banks:wavelets

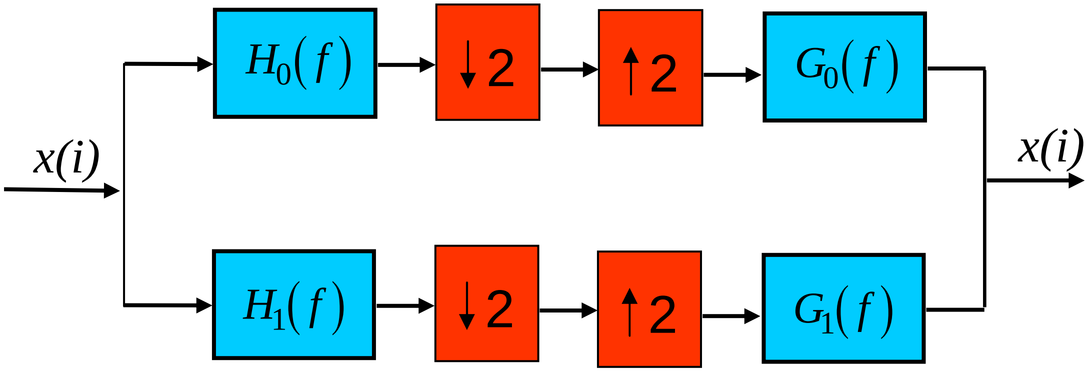
- If the prototype filter $H_0(\omega)$ has a zero at frequency $\omega = \pi$, the filters are said to be <u>regular filters</u>, or <u>wavelets</u>.

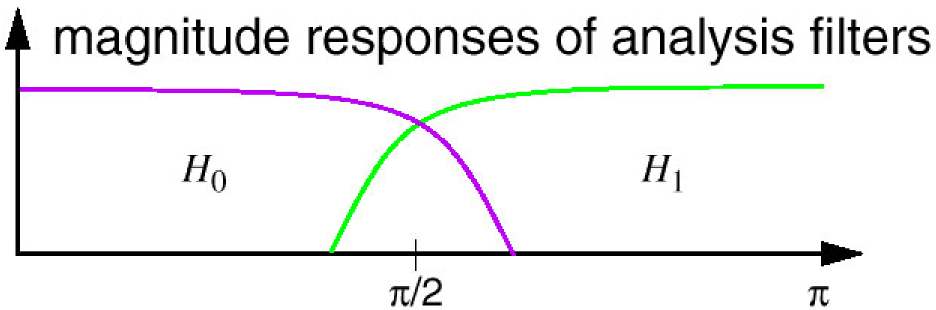
- Aliasing-free filter design:
  - If we use symmetric (linear phase) FIR filters and let:

$$
\boxed{\begin{aligned}
H_1(z) &= H_0(-z) \Rightarrow h_1(n) = (-1)^n h_0(n) \\
G_0(z) &= 2H_0(z) \Rightarrow g_0(n) = 2h_0(n) \\
G_1(z) &= -2H_1(z) \Rightarrow g_1(n) = -2(-1)^n h_0(n)
\end{aligned}}
\qquad X_{out}(z) = X(z)\left[H_0^2(z) - H_1^2(z)\right]
$$

  - Since both $H_0(z)$ and $H_1(z)$ have linear phase, the system introduces no phase distortions.

3.1. QMF Filter banks:perfect reconstruction





One problem with linear phase filters: only two-taps (coefficients) filters.

3.1. Filter banks: perfect reconstruction

There are other possible approaches to subband filter design that yield substantially different filter characteristics:

- Conjugate quadrature filters (CQFs)

  M. Smith and T. Barnwell, IEEE ASSP, 3, pp. 434-441, 1985.

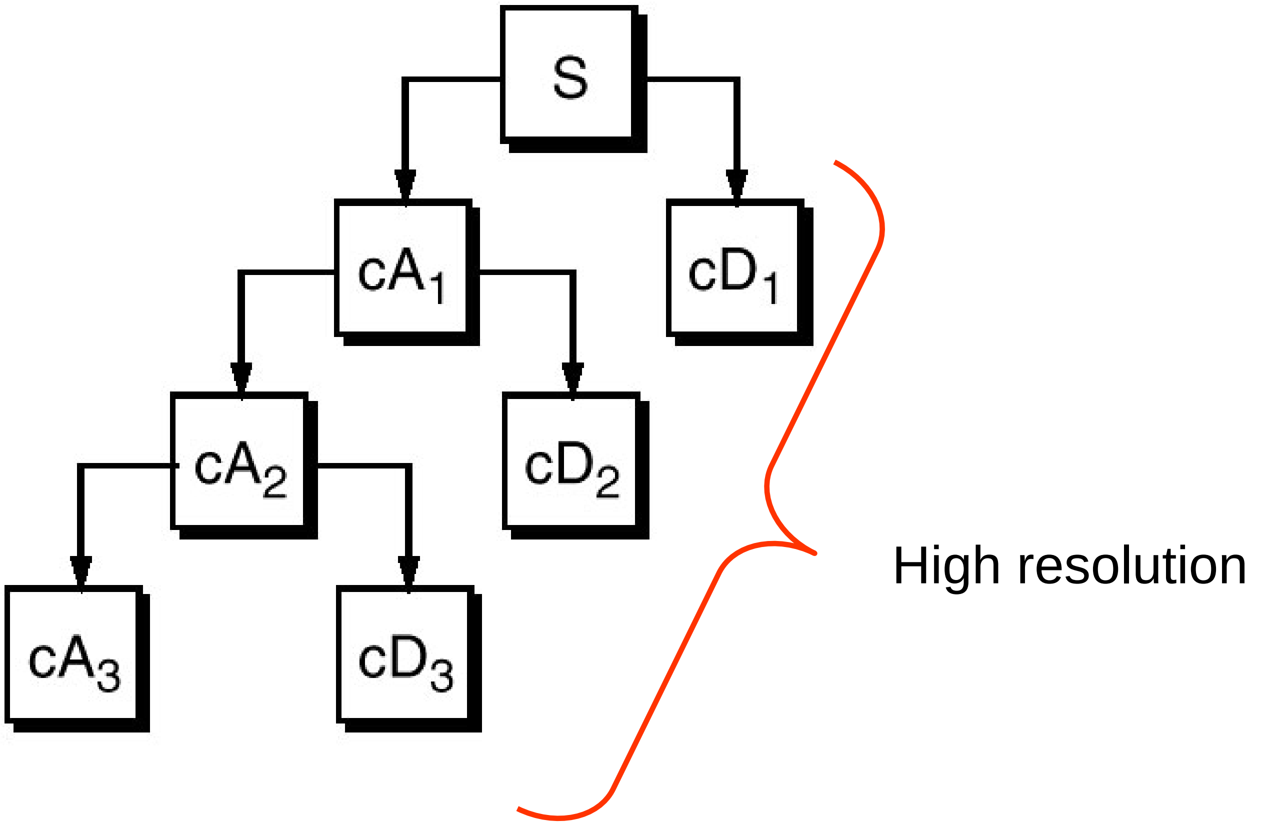
  The analysis/synthesis filters have non-symmetrical coefficients and nonlinear phase, but finally they possess a perfect reconstruction.

- Biorthogonal wavelets.

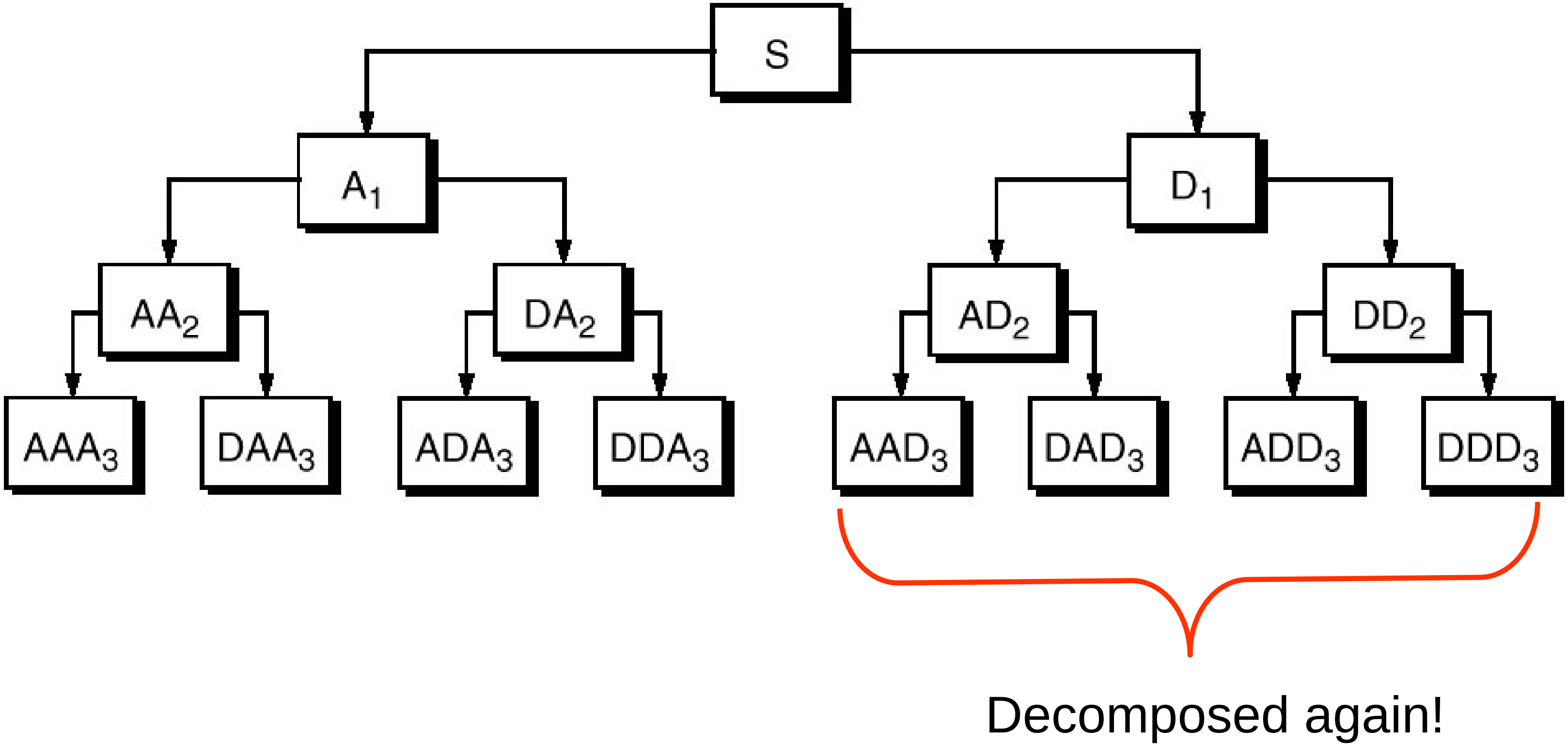
3.1. Requirements to filters

- There are many filter banks available for general use.
- However, when choosing filters for subband image decomposition, there are additional requirements that are specific to image coding (denoising and restoration):
  - Analysis filters should have a short impulse response to preserve the localization of image features.
  - Synthesis filters should be also short to prevent spreading of artifacts resulting from quantization errors at edges and other local features.
  - Long synthesis filters often have very good mean squared error performance but lead to annoying ringing artifacts around edges.
  - Linear phase filters are desirable for subband image coding (nonlinear phase filters introduce subjectively unpleasant distortions in lowpass channels).

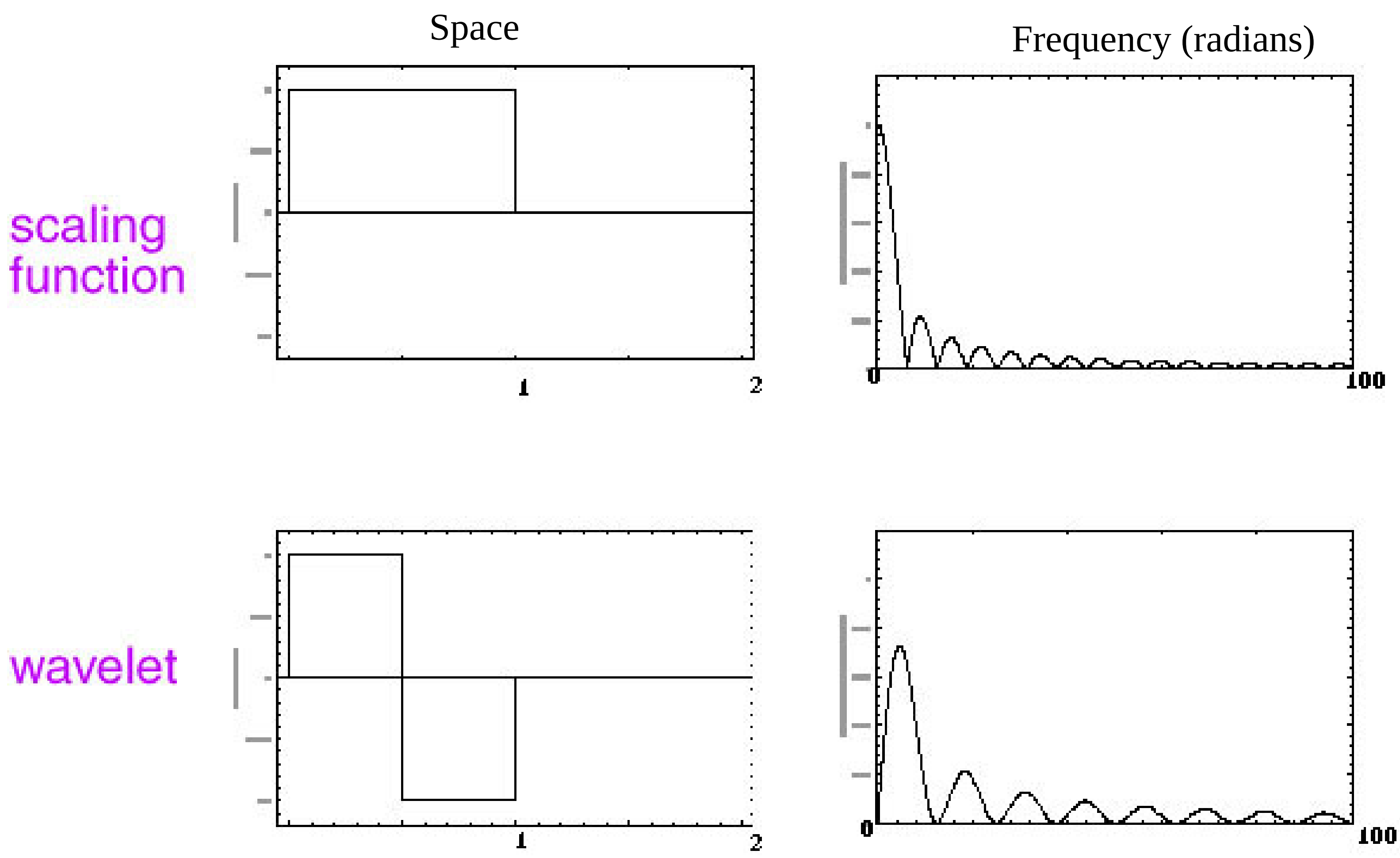
3.2. Wavelets



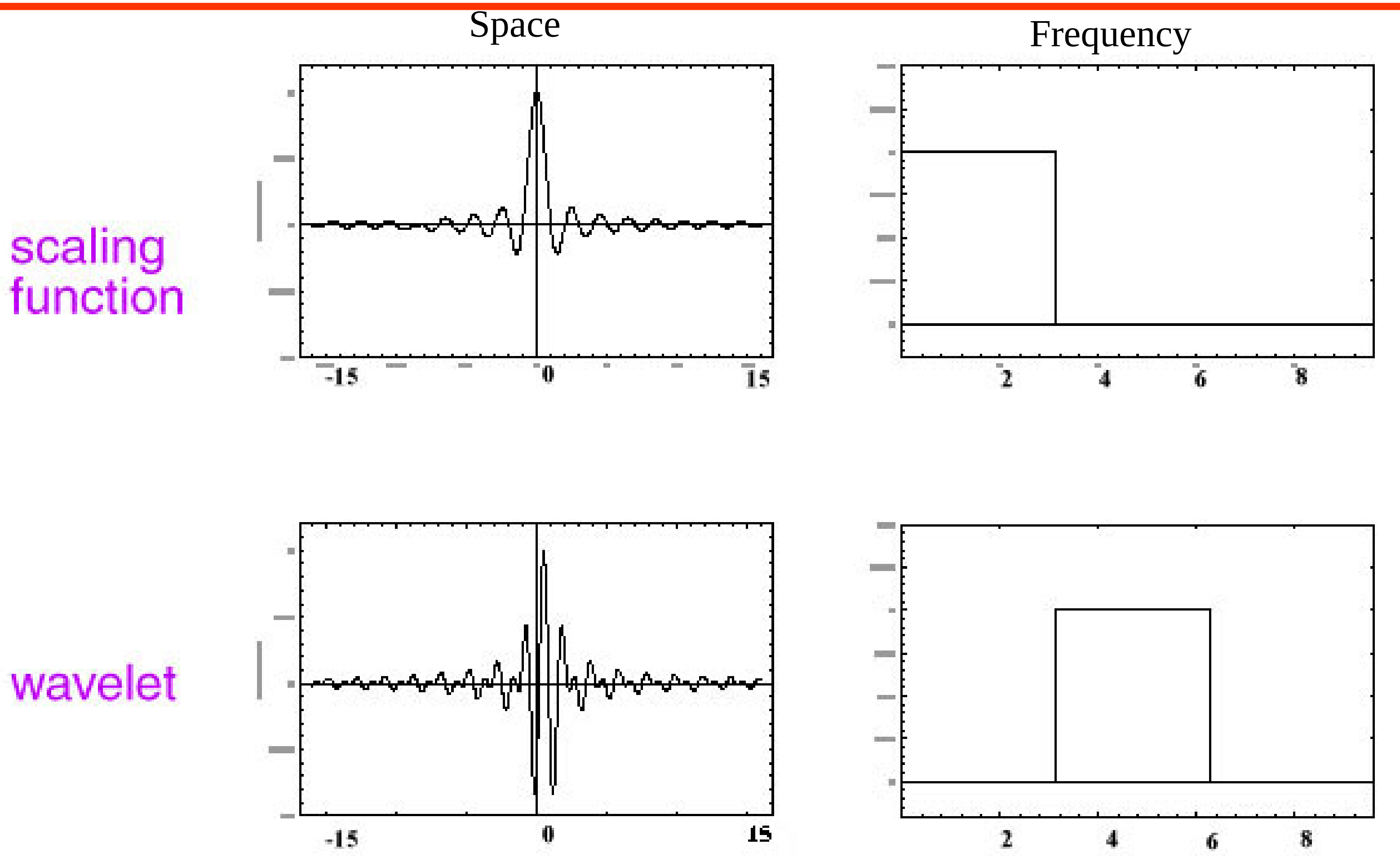
3.2. Wavelet packets



3.2. Basis functions: Haar system



3.2. Basis functions: Sinc system



3.2. Basis functions: common features

Haar and sinc systems: either good time OR frequency localization

↓

Compromise: Daubechies system

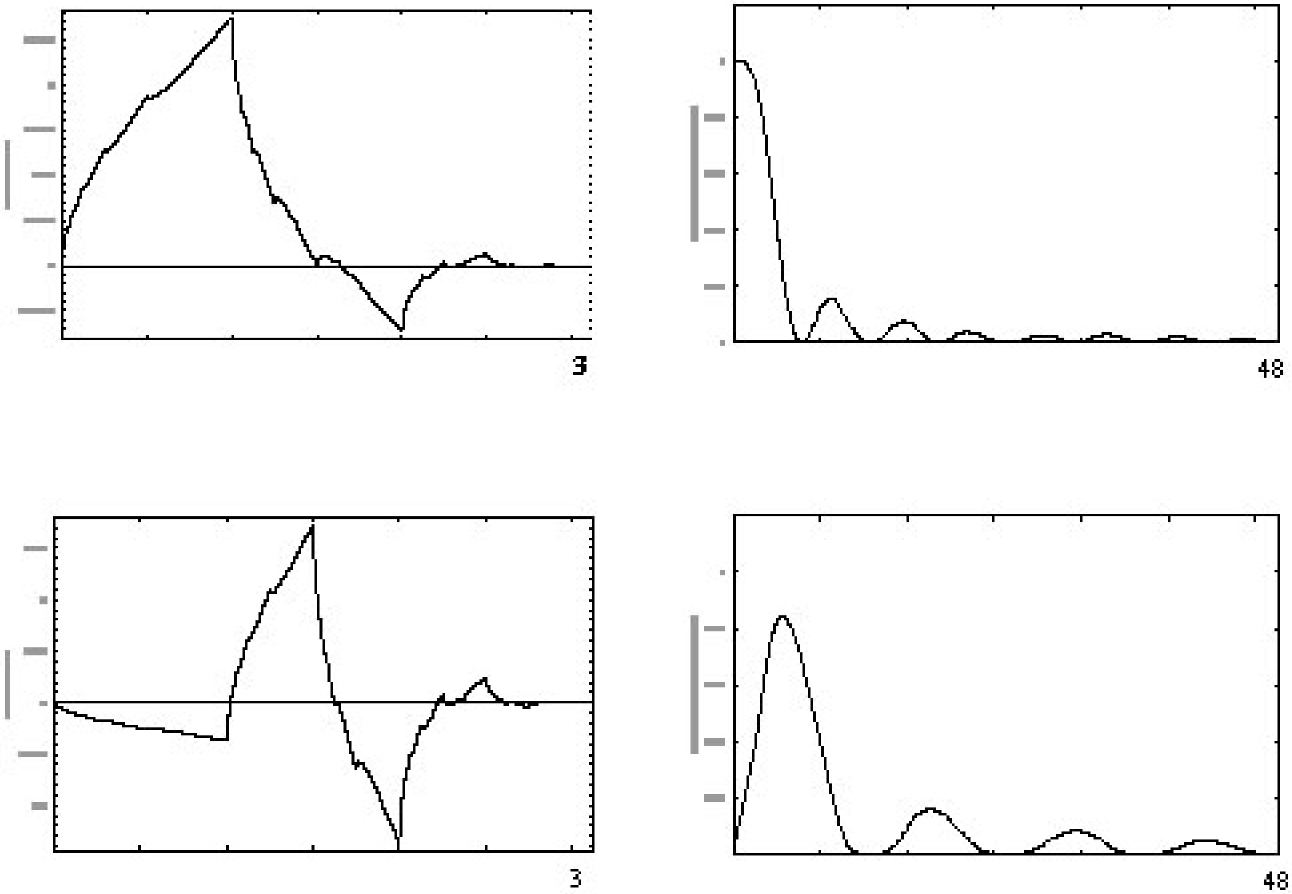
Good time AND frequency localization

3.2. Basis functions: Daubechies' system

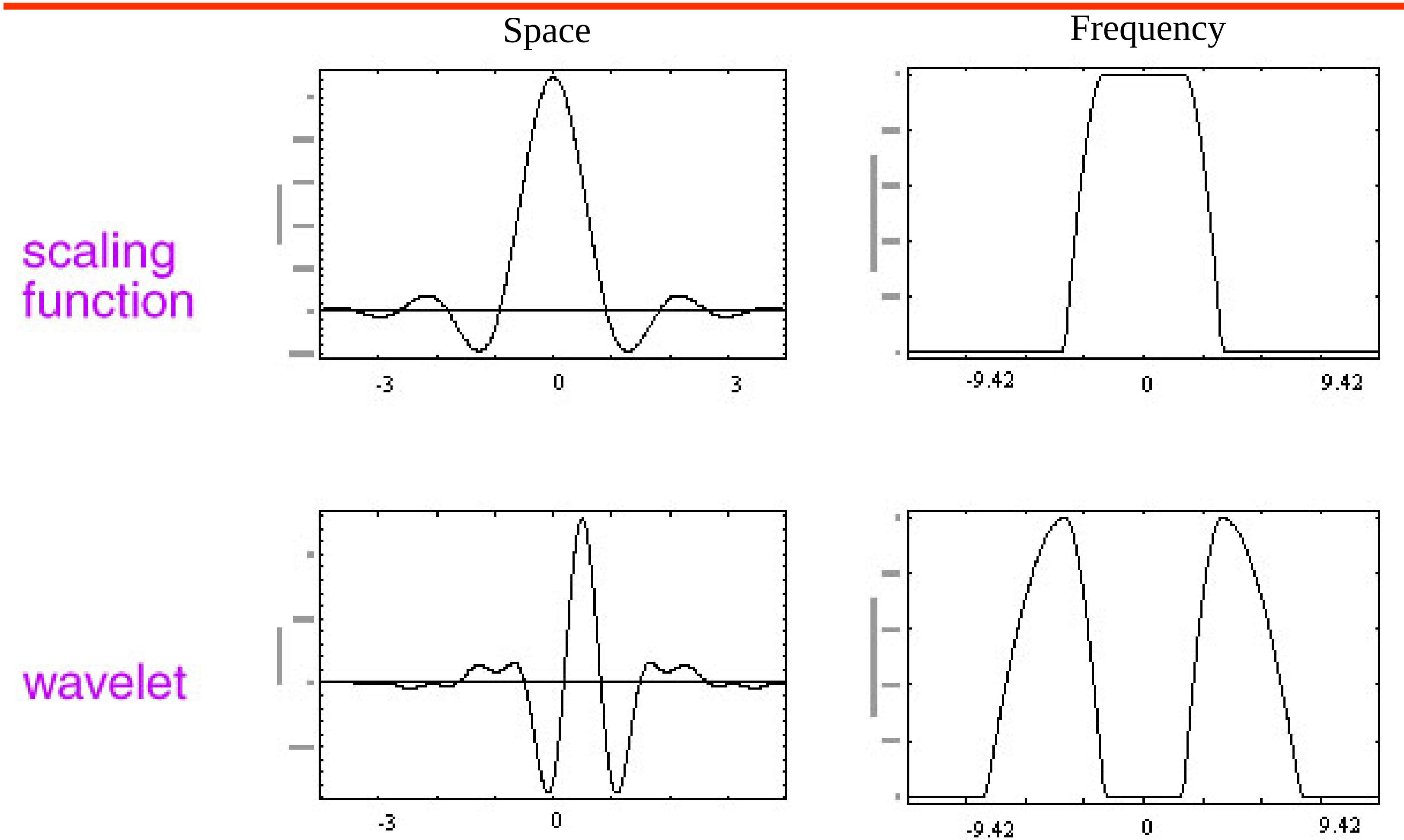
Space　　　　Frequency

scaling function

wavelet



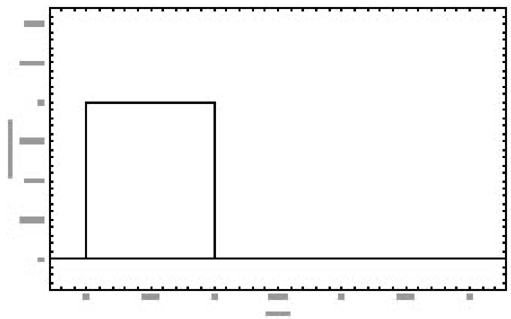
3.2. Basis functions: Meyer's system

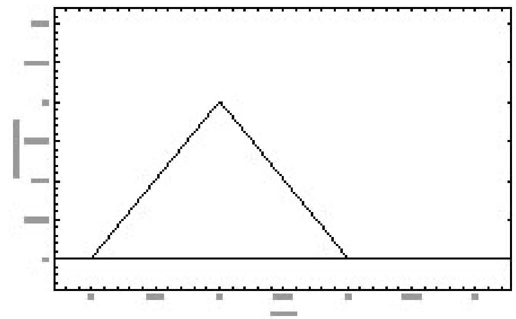


3.2. Splines-wavelets

constant

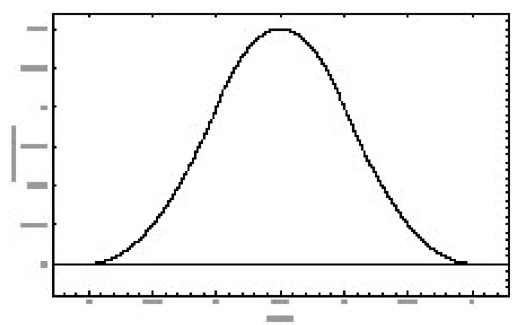
$n=0$





linear

$n=1$



quadratic

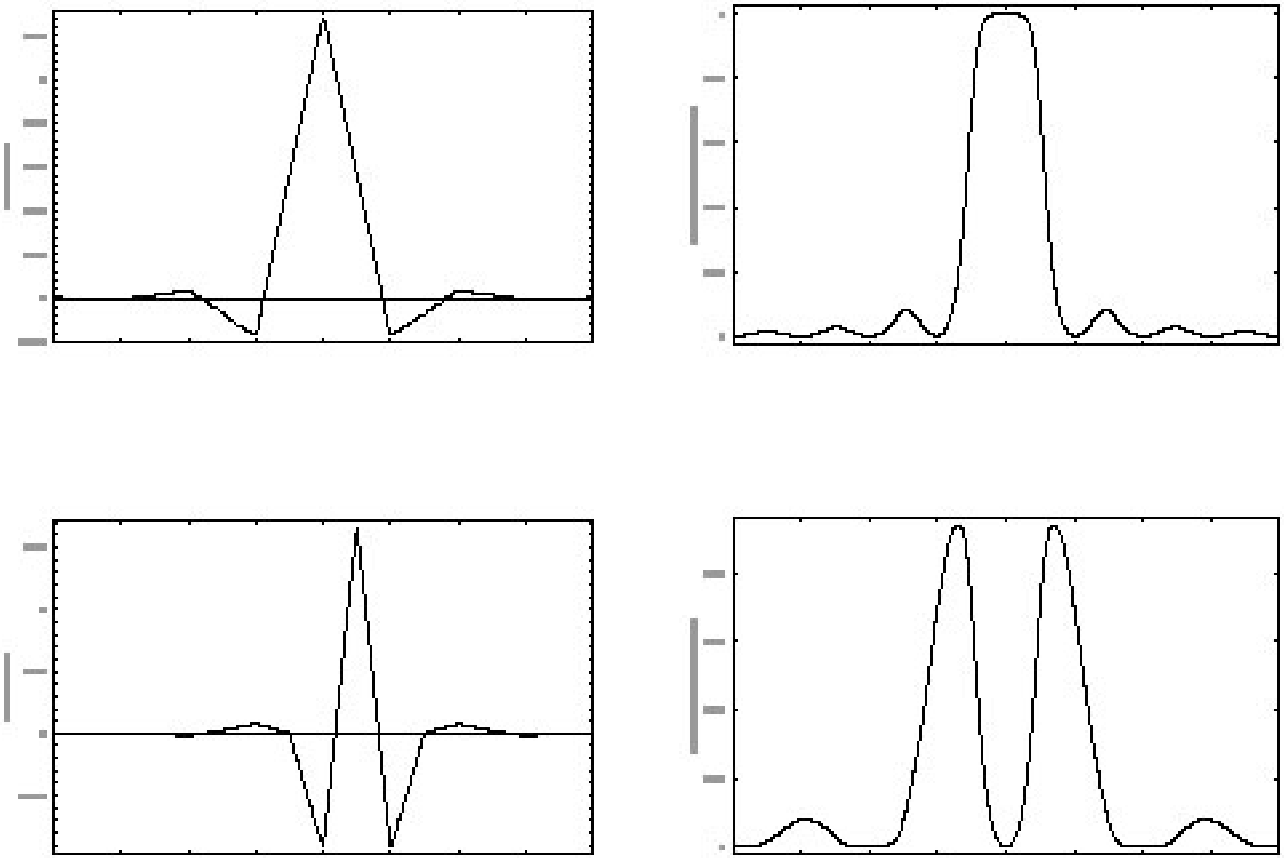
$n=2$

Only splines with $n=0,1$ are orthogonal, the rest need to orthogonilize.

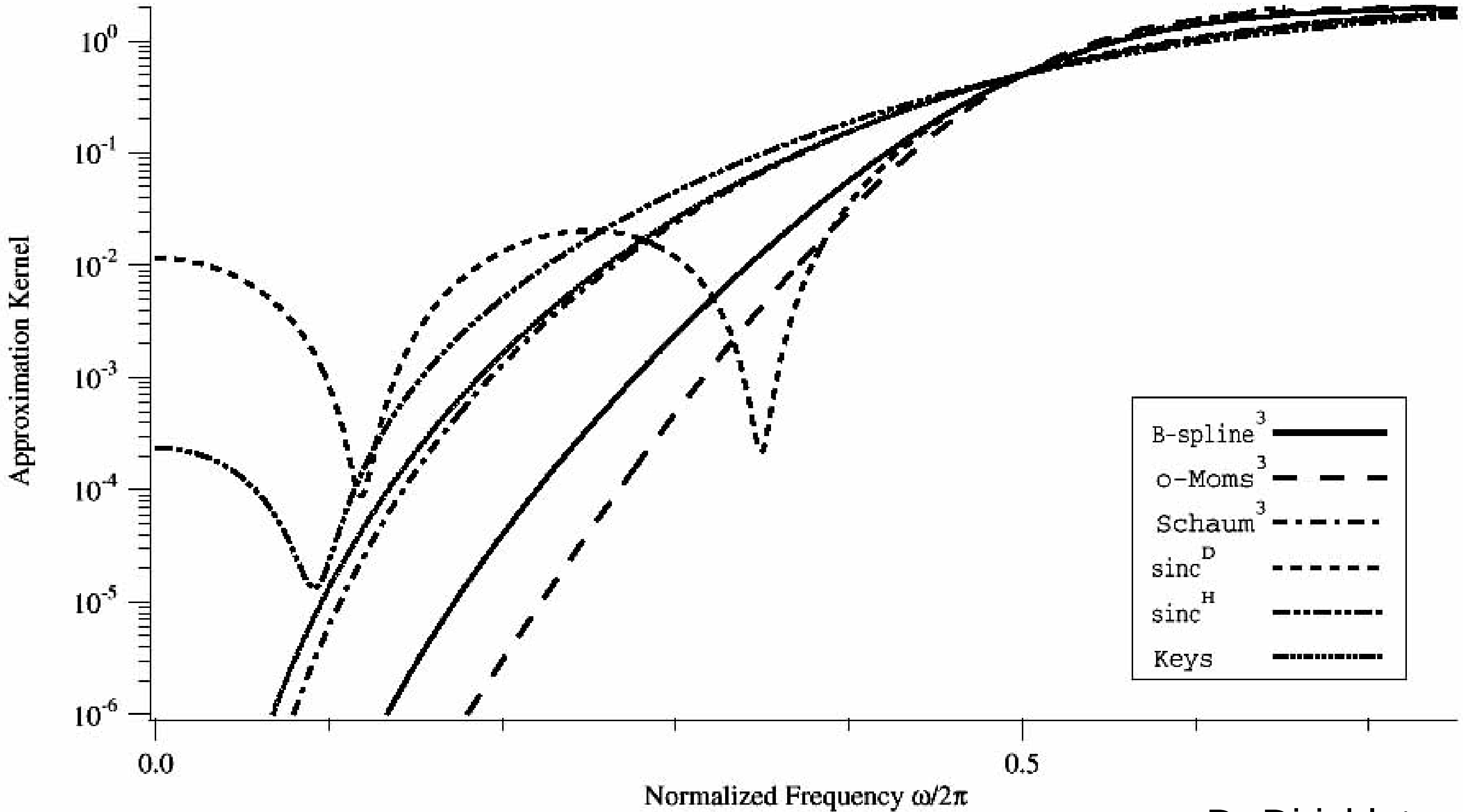
3.2. Basis functions: splines

scaling function

wavelet



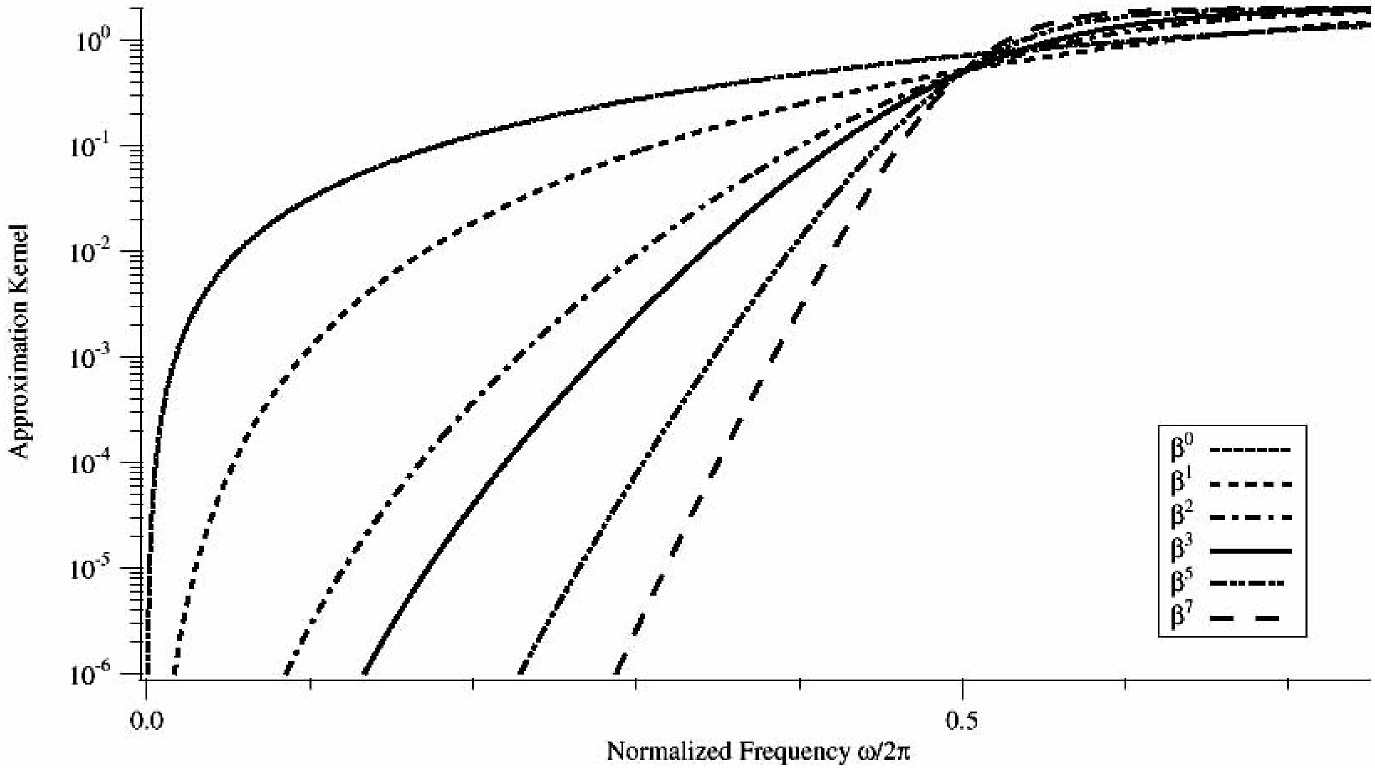
3.2. Spline synthesis functions (Unser)



D- Dirichlet

H- Hanning apodization

3.2. B-spline synthesis functions (Unser)

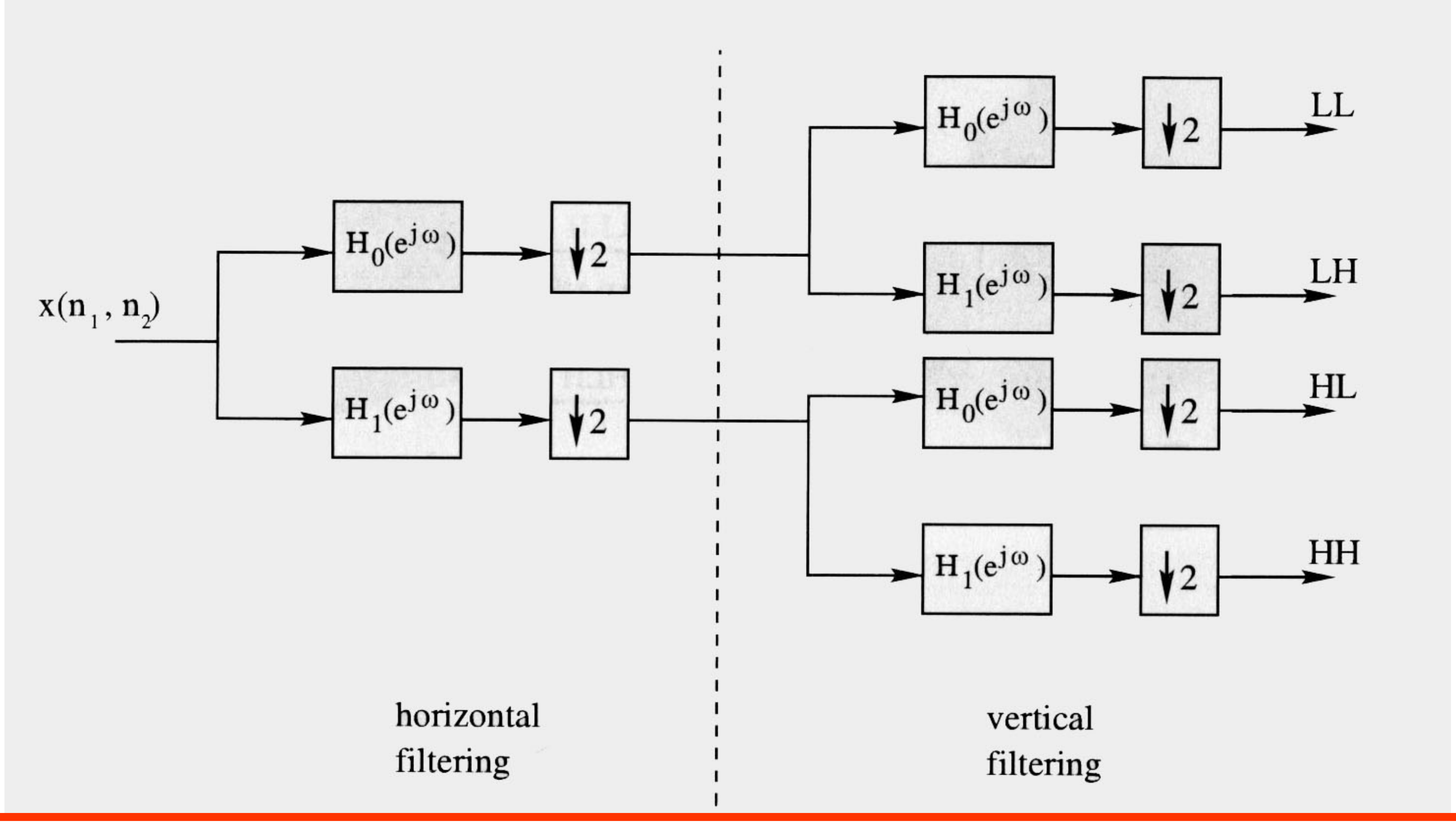


3.2. Reasons to use splines (Unser)

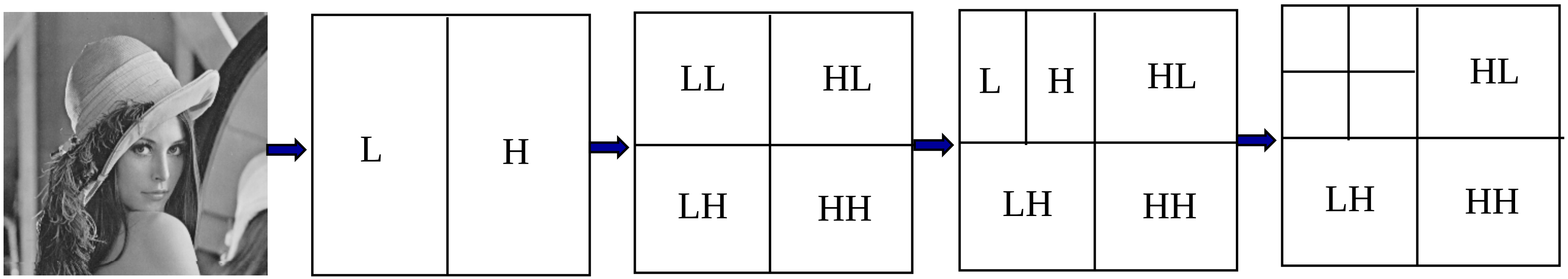
- closed-form representation
- simple manipulation
- symmetry
- shortest scaling function of order N
- maximum regularity for a given order N
- best approximation
- optimal time-frequency localization
- convergence to the ideal filter

3.2. Wavelet Decomposition

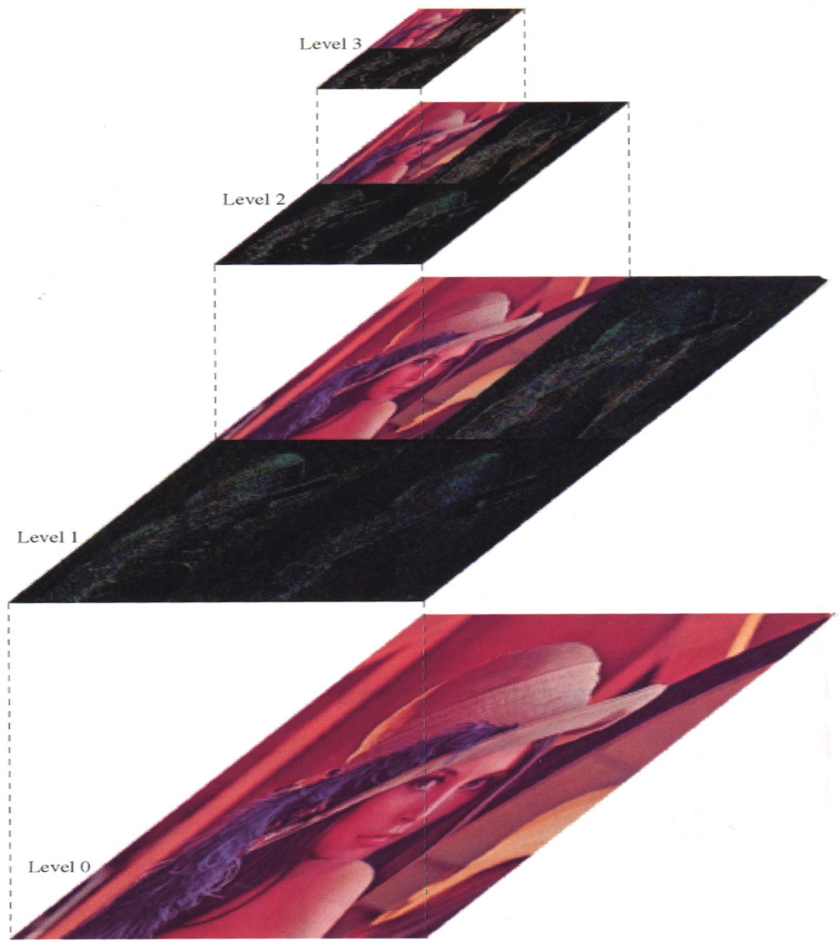
Separable 2-D decomposition can be applied to images using first 1-D filtering along rows of the image and then along columns, or vice versa.



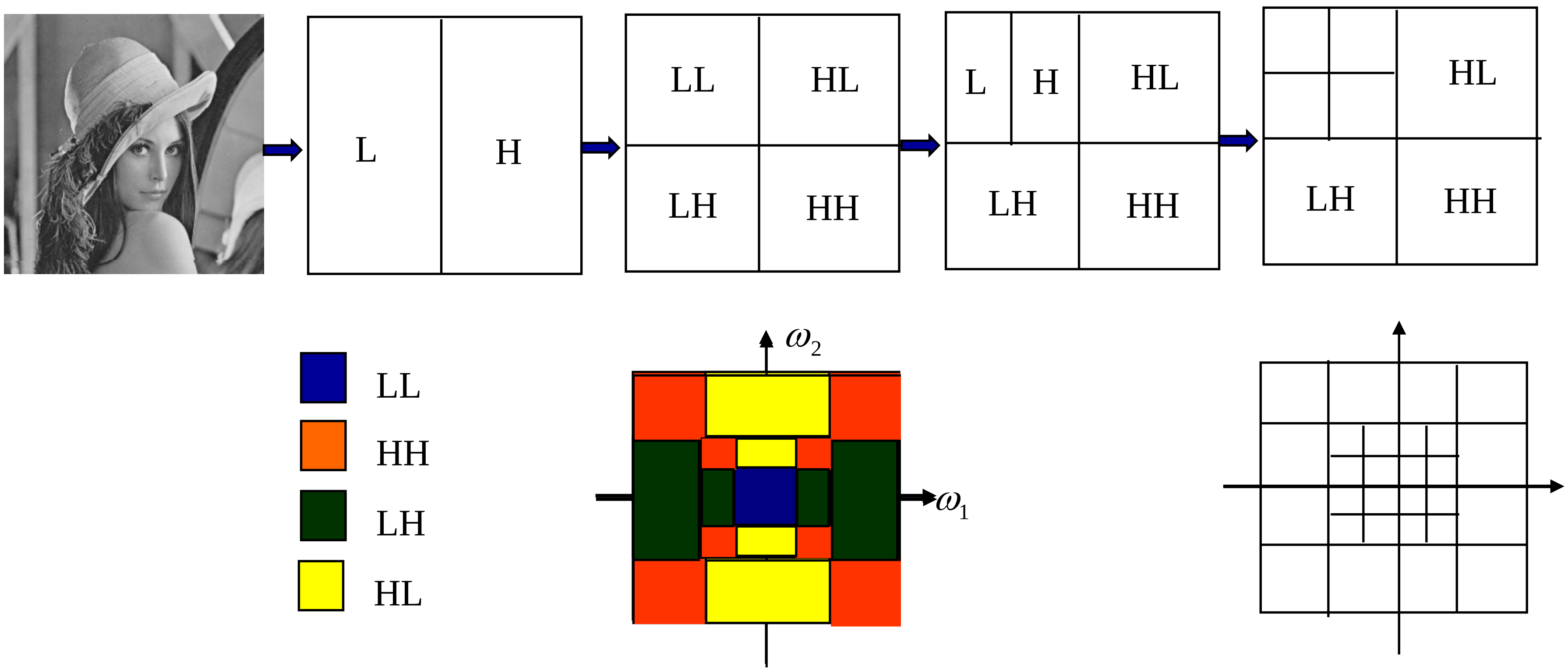
3.2. Wavelet Decomposition



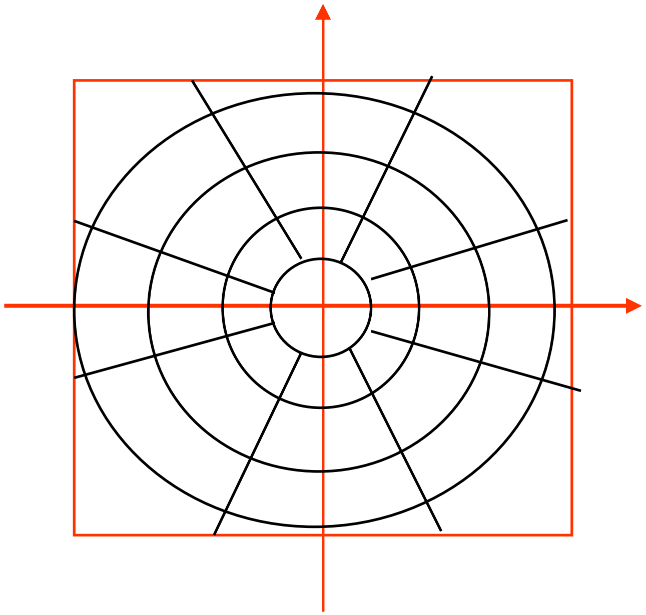
3.2. Wavelet Decomposition (Bovik, C-24)



3.2. Wavelet Decomposition: frequency domain



3.2. Steerable Decomposition: frequency domain

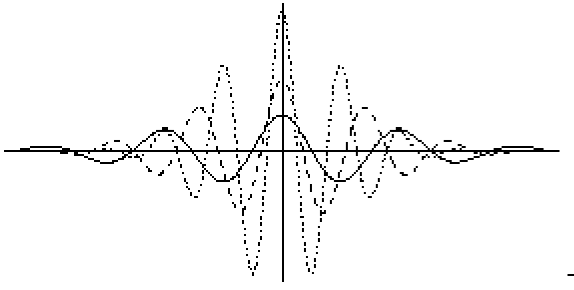


Eero Simoncelli

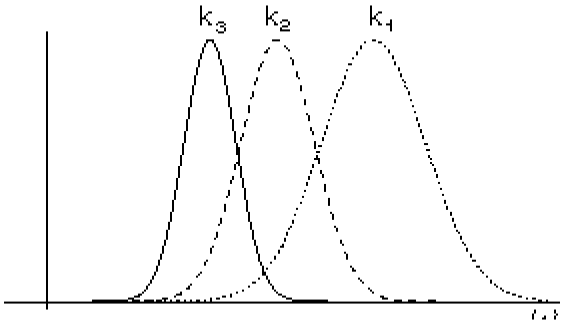
http://www.cis.upenn.edu/~eero/steerpyr.html

3.2. Gabor filter

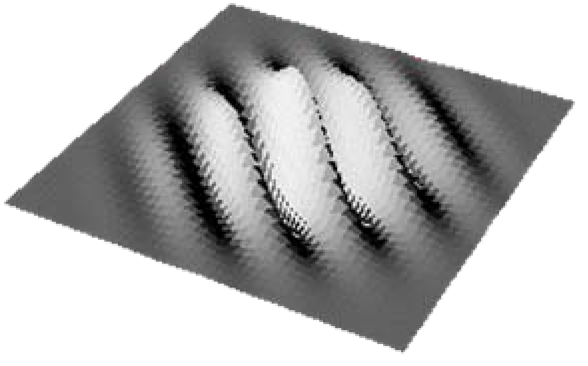
Basis function: Exp modulated by sine wave



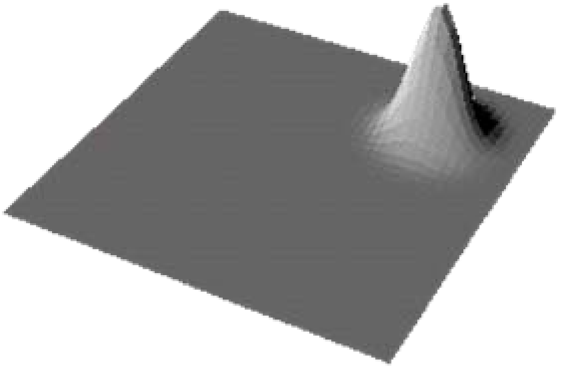
Frequency domain



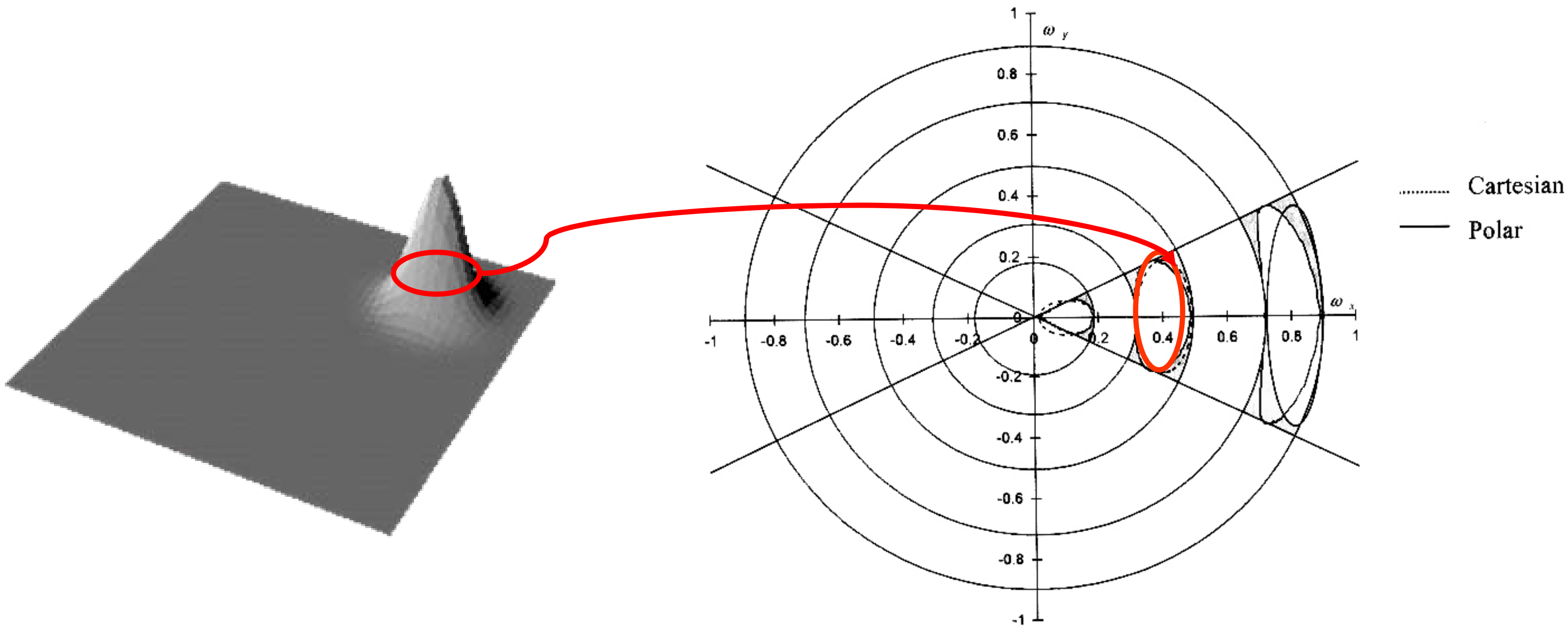
2-D Coordinate domain



2-D Frequency domain



3.2. Gabor filter



3.2. Fundamentals of wavelets

Representation of signal $x(t)$ defined over some (discrete or continuous) domain T:

$$x(t)=\sum_k a_k \varphi_k(t),\ t\in T$$

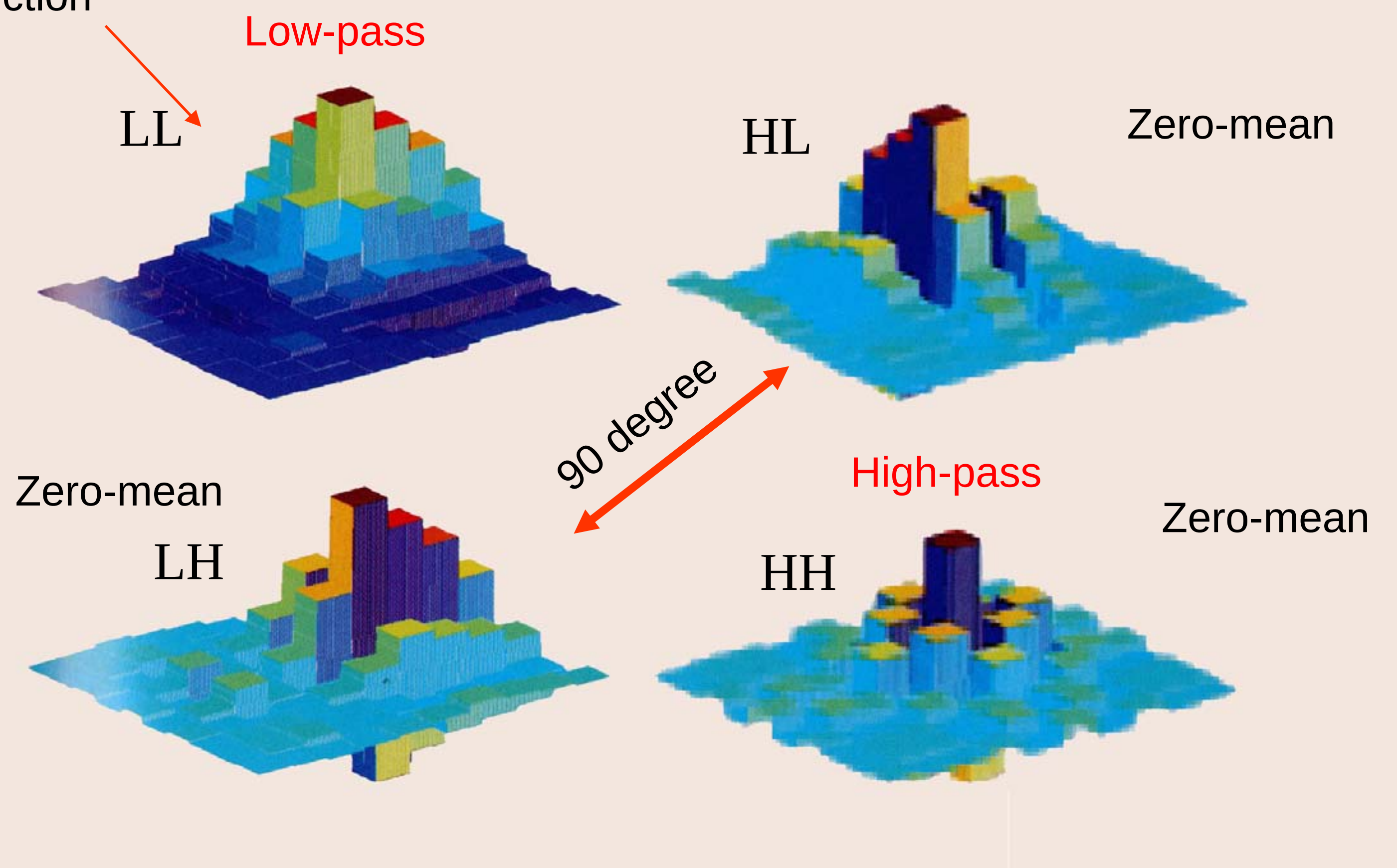
Basis functions

Coefficients of $x(t)$ *in the basis* $\{\varphi_k(t)\}$

The shape of basis functions depends on the scale and orientation of subbands.

3.2. Discrete basis functions in the four subbands

Scaling function

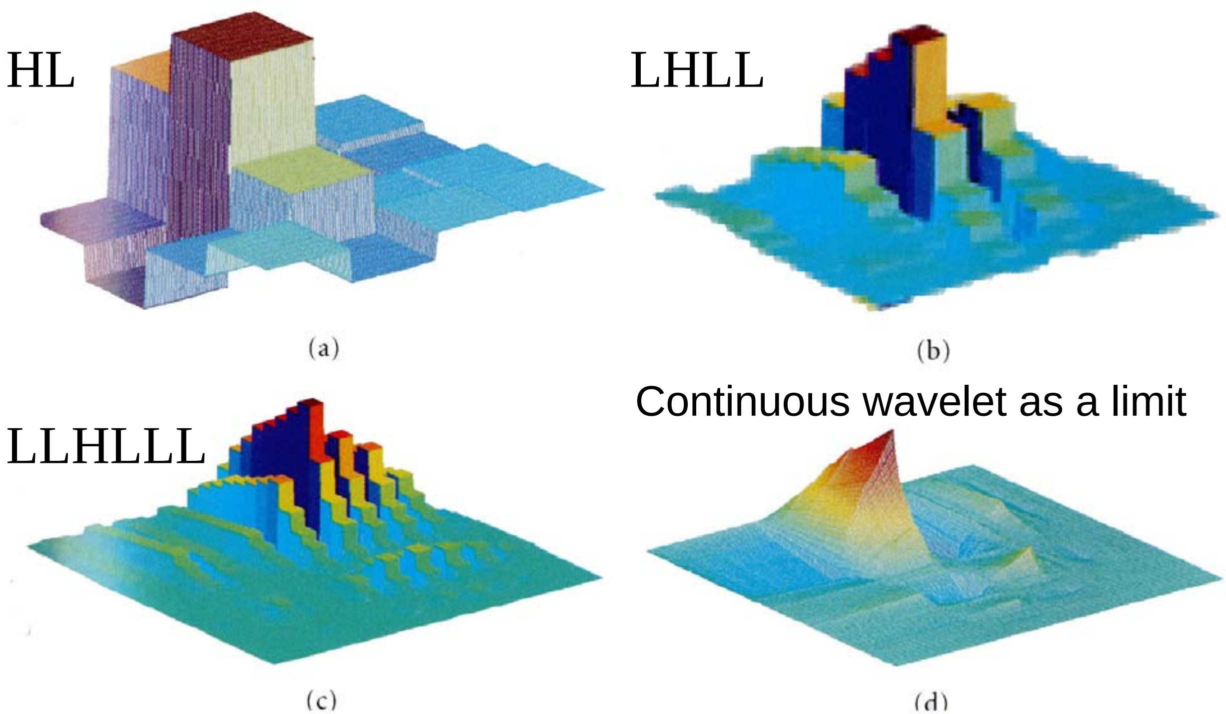


The rest are called *wavelets*.

Bovik, Ch.4.2

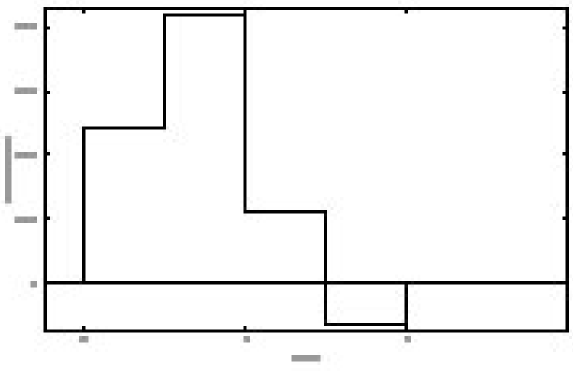
3.2. Basis functions with vertical orientation

Basis functions with the same orientation in different subbands have a *similar shape*.

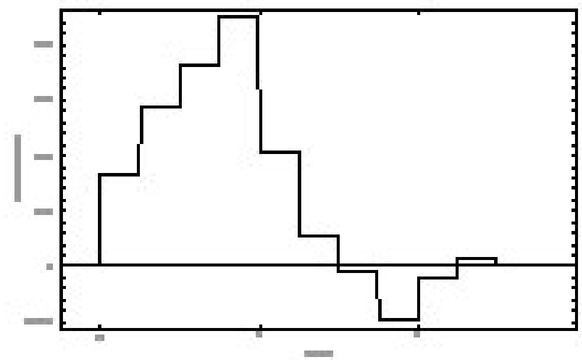


3.2. Limit of scale and wavelet functions

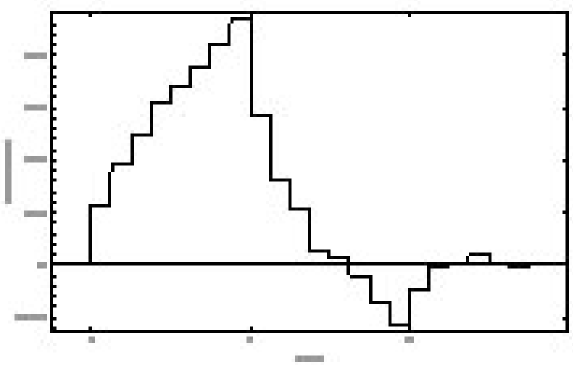
i = 1



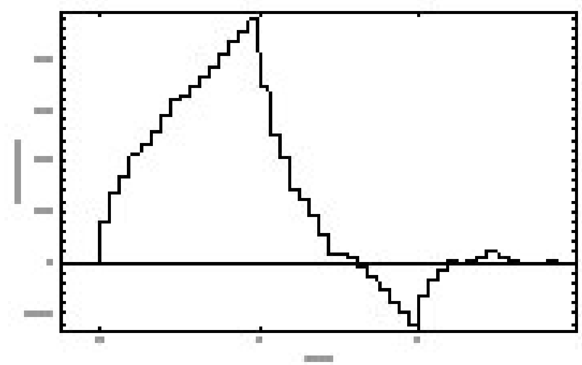
i = 2



i = 3

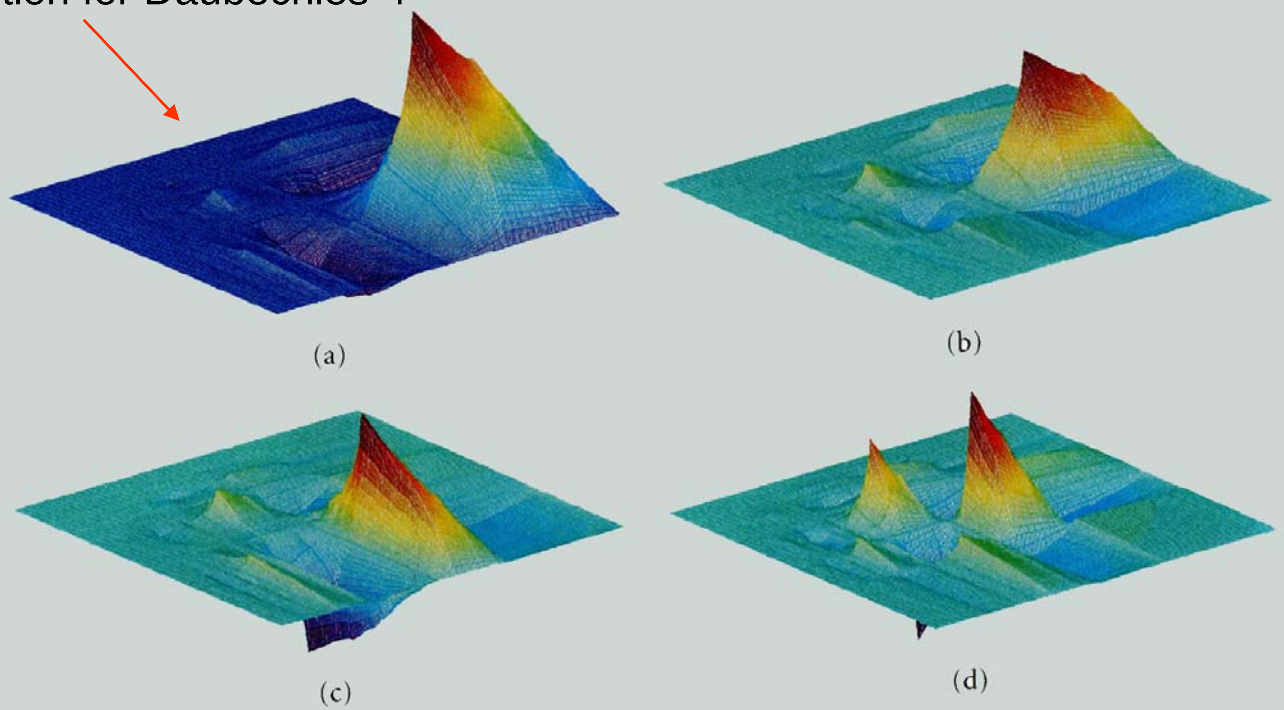


i = 4



3.2. Limit of scale and wavelet functions

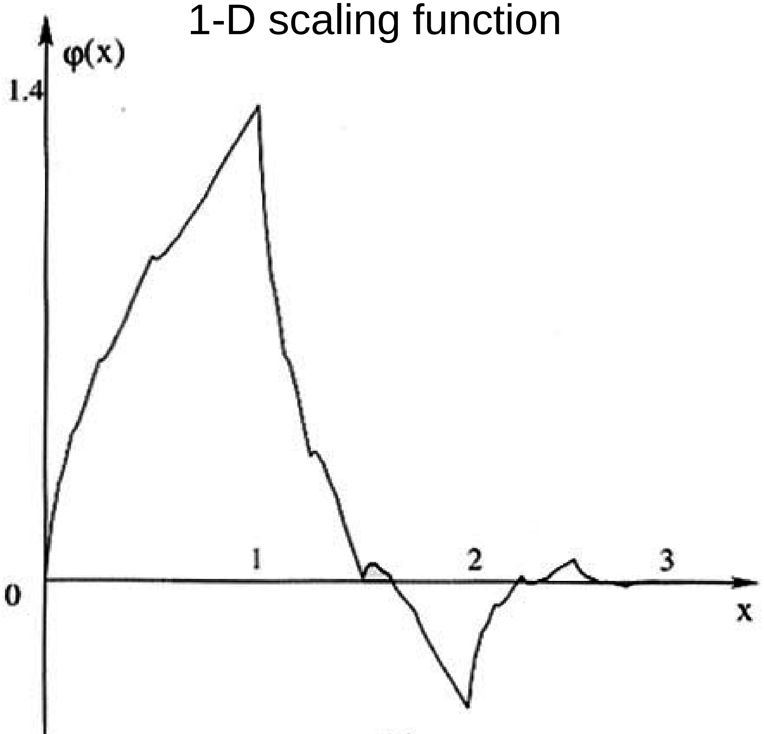
Scaling function for Daubechies-4



The remaining 3 are wavelets.

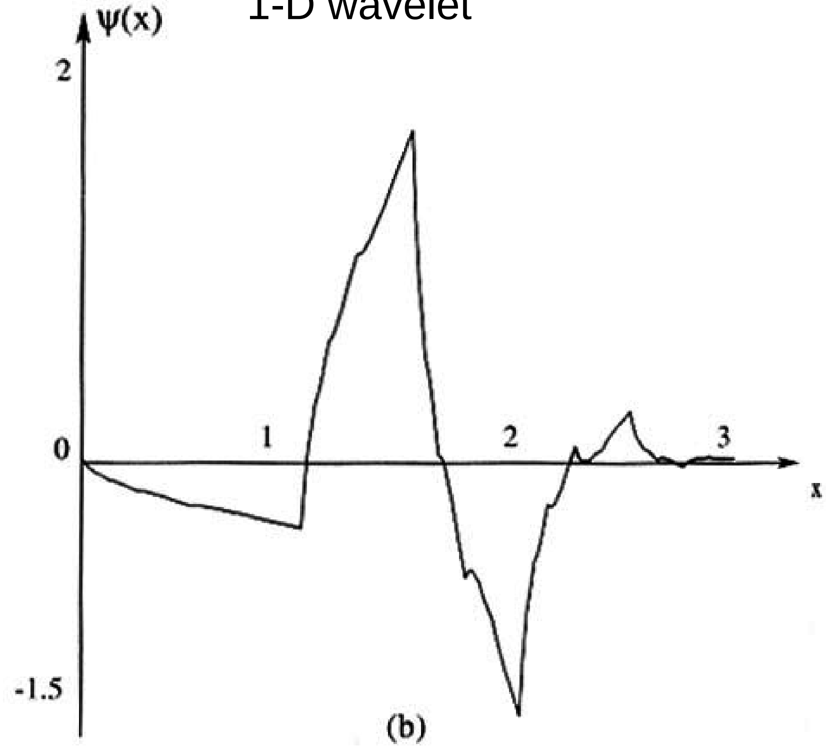
3.2. Generation of 2-D wavelets from 1-D functions

1-D scaling function



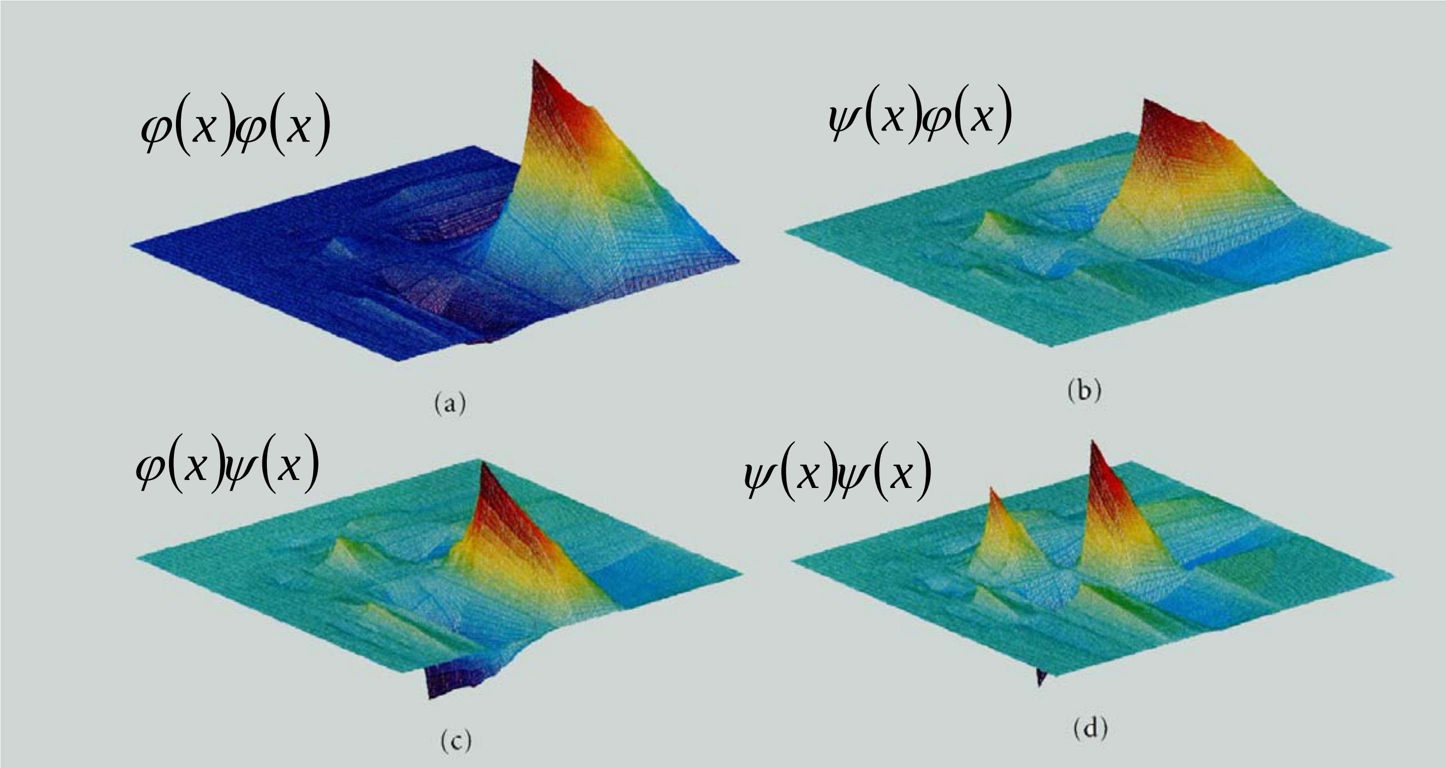
(a)

1-D wavelet



(b)

3.2. Generation of 2-D wavelets from 1-D functions



The remaining 3 are wavelets.

3.2. Properties of wavelets

The wavelet transform is linear and leads to the sparse image presentation.

There is strong connection between the type of wavelet filters and properties of the image.

The best decomposition (the sparsest representation):

- images with sharp edges would benefit from the use of short wavelets;
- images with mostly smooth areas would benefit from the use of longer wavelet filters with several vanishing, moments, since such filters generate smooth wavelets.

3.3. Relation to Human Visual System

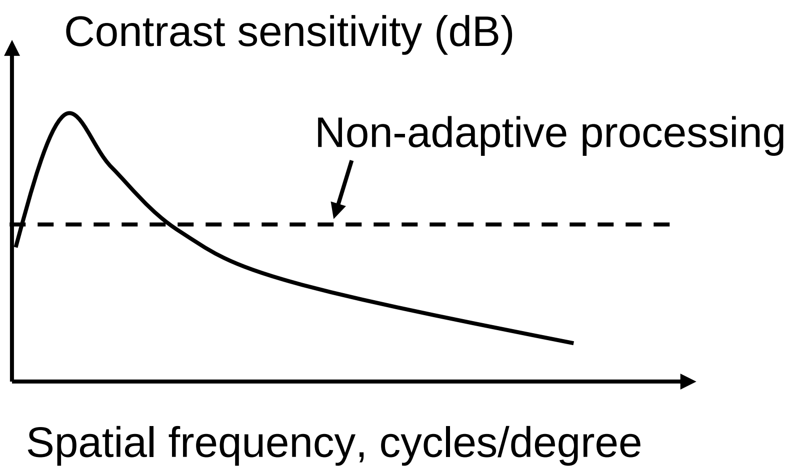
The sensitivity of the HVS depends on spatial frequency (MTF). Additionally, the image is decomposed into bandpass channels with different scales and spatial orientations in the visual cortex of the brain.

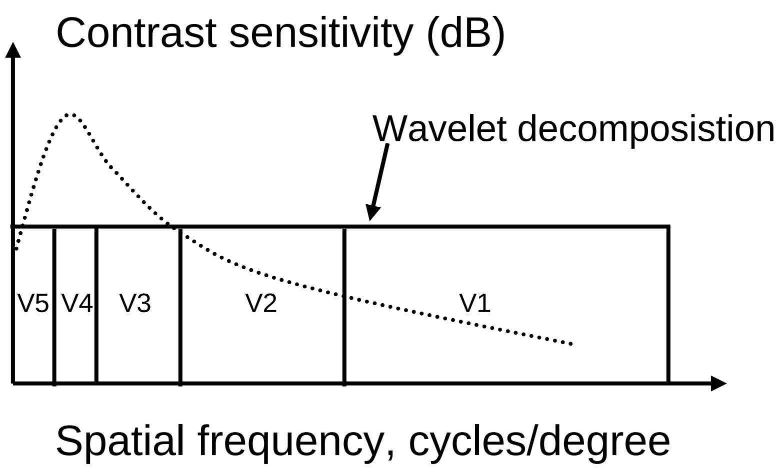
Wavelets reflects the multiresolution decomposition with 3 spatial orientations (H, V and D).

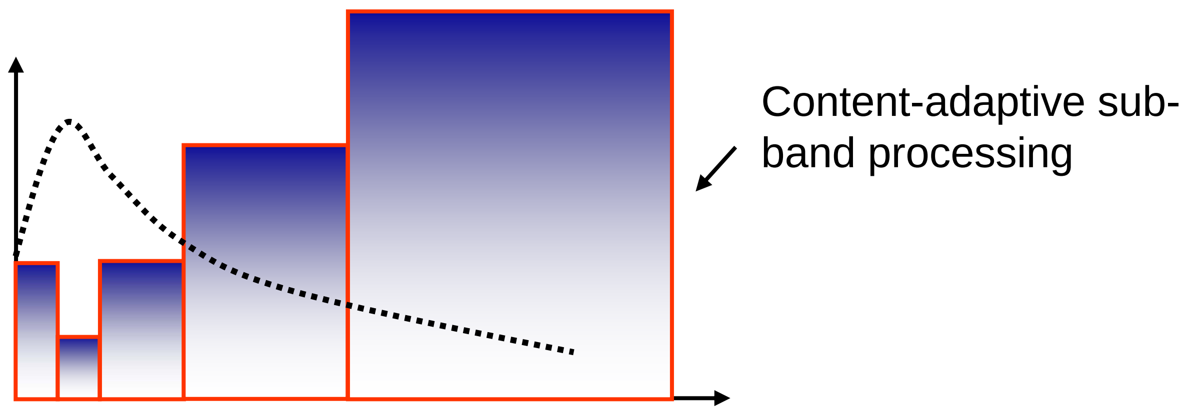
The other linear approximations:

- *Gabor transform* (basis functions are Gaussian functions modulated by sine waves).
- *Cortical transforms of Watson and Daly*.
- *Steerable pyramids.*

3.3. Relation to Human Visual System: wavelets and MTF (see Part I)







3.3. Relation to Human Visual System

Disadvantages of Gabor filter banks:

- Computationally quite expensive (especially for the evaluation of low-frequency components).
- The outputs of Gabor filter banks are not mutually orthogonal:
  - that results in a significant correlation between image coefficients;
  - this transform is not reversible, which limits the applicability for compression, denoising and texture synthesis.

Most of the problems are avoided if one uses the wavelet transform, which provides a precise and unifying framework for the analysis and characterization of an image at different scales.