Transform Domain Image Processing

Giacomo Boracchi

giacomo.boracchi@polimi.it

12th December 2011

ARTIFICIAL INTELLIGENCE and ROBOTICS LABORATORY

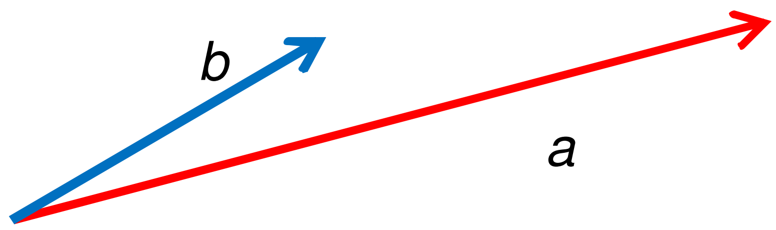
Preliminary Notation

- A basis $\{b_i\}_{i=1,..,N}$ of a vector space *V* is a **set of linearly independent vectors** that *spans* (i.e. that can represent with linear combinations) the whole vector space *V*.
  - *N,* the number of elements of the basis is said the dimension of *V* (potentially it could not be finite)

Preliminary Notation

- A basis $\{b_i\}_{i=1,..,N}$ of a vector space *V* is a **set of linearly independent vectors** that *spans* (i.e. that can represent with linear combinations) the whole vector space *V*.
  - *N,* the number of elements of the basis is said the dimension of *V* (potentially it could not be finite)

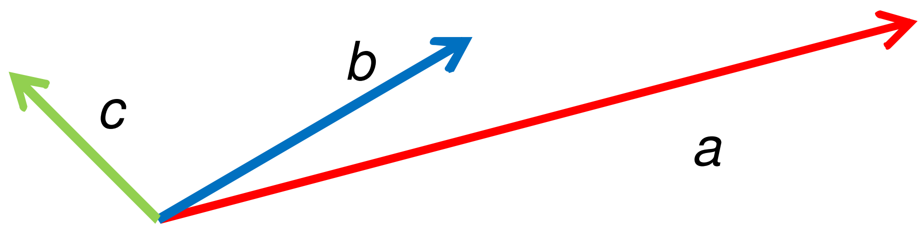
- Thus, these two vectors represent a basis



Preliminary Notation

- A basis $\{b_i\}_{i=1,..,N}$ of a vector space *V* is a **set of linearly independent vectors** that *spans* (i.e. that can represent with linear combinations) the whole vector space *V*.
  - *N*, the number of elements of the basis is said the dimension of *V* (potentially it could not be finite)

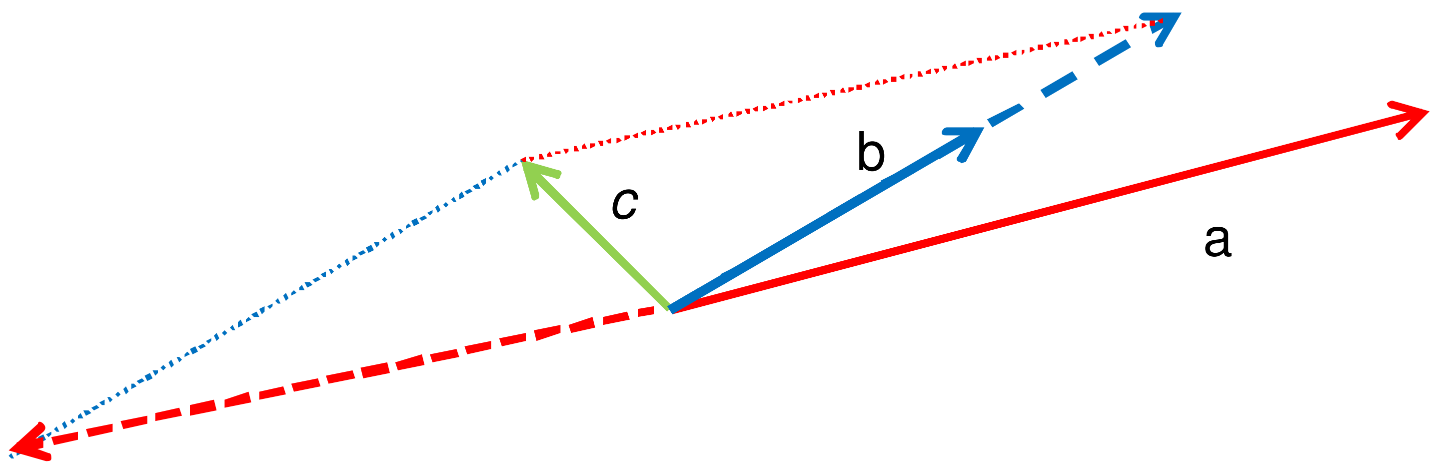
- any other element can be represented as a linear combination of them



Preliminary Notation

- A basis $\{b_i\}_{i=1,..,N}$ of a vector space *V* is a **set of linearly independent vectors** that *spans* (i.e. that can represent with linear combinations) the whole vector space *V*.
  - *N*, the number of elements of the basis is said the dimension of *V* (potentially it could not be finite)

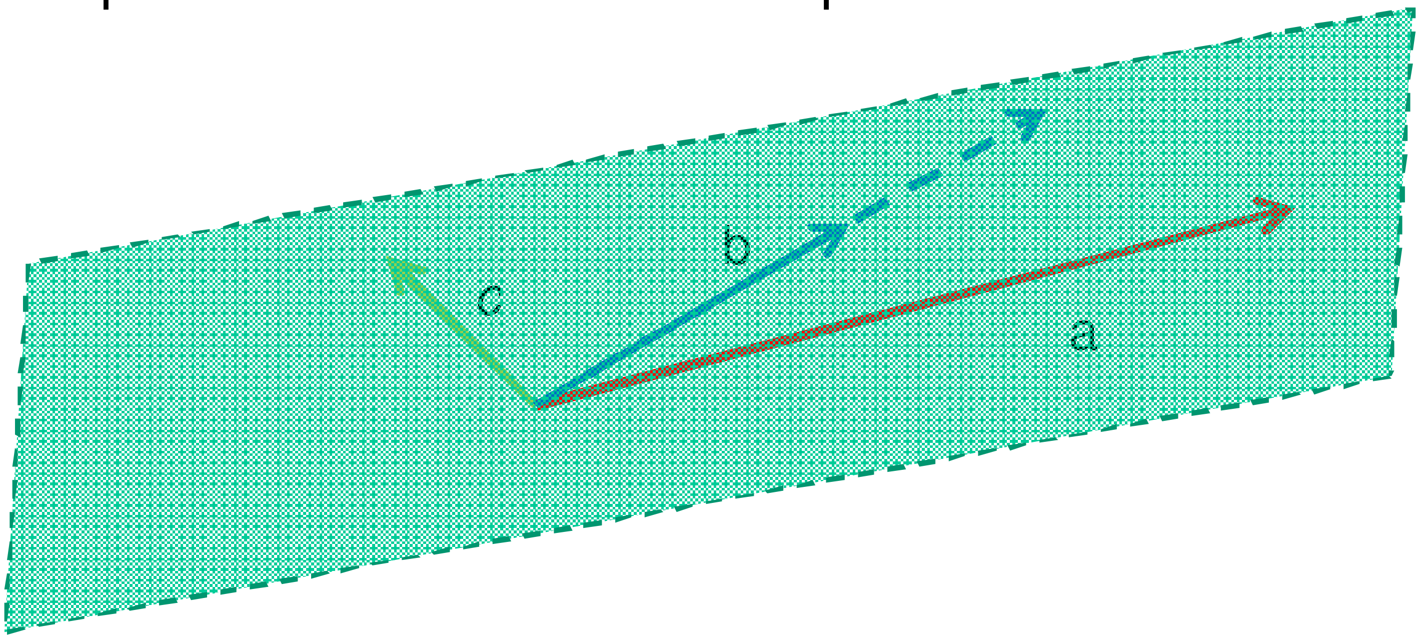
- c = (-1) * a + 1.5 * b



Preliminary Notation

- A basis $\{b_i\}_{i=1,..,N}$ of a vector space *V* is a **set of linearly independent vectors** that *spans* (i.e. that can represent with linear combinations) the whole vector space *V*.
  - *N,* the number of elements of the basis is said the dimension of *V* (potentially it could not be finite)

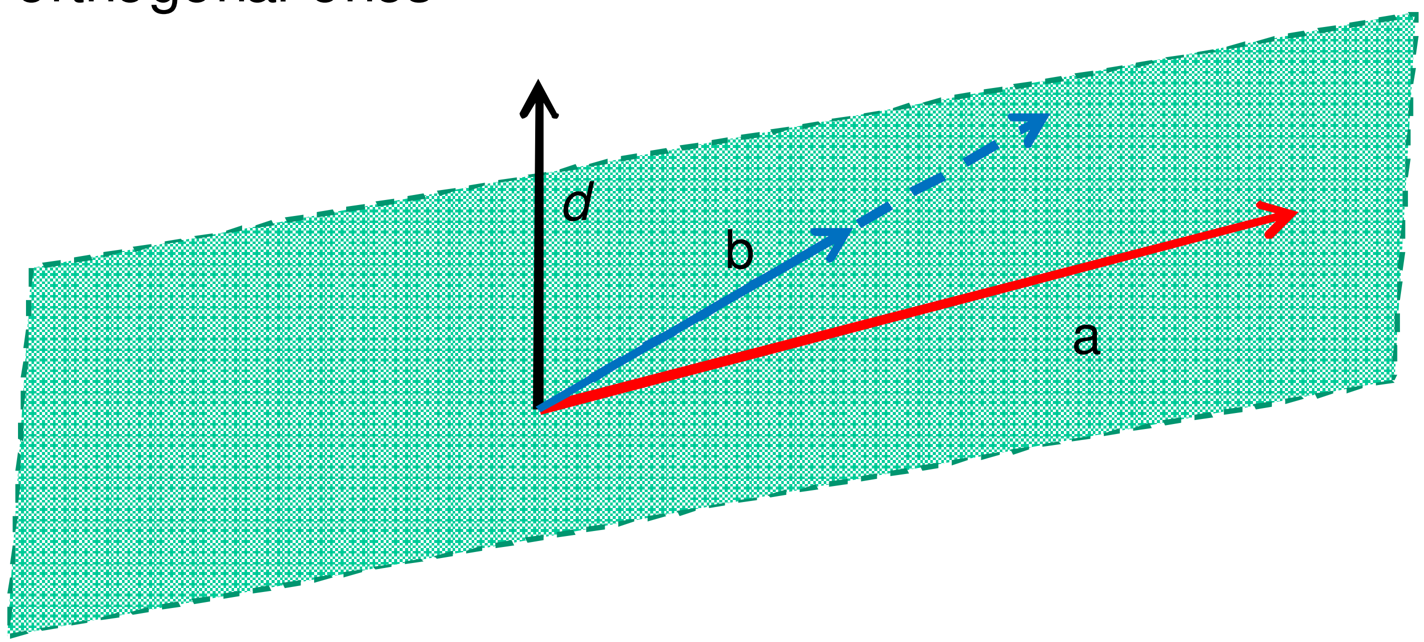
- thus we can represent all the vector in the plane



Preliminary Notation

- A basis $\{b_i\}_{i=1,..,N}$ of a vector space *V* is a **set of linearly independent vectors** that *spans* (i.e. that can represent with linear combinations) the whole vector space *V*.
  - *N*, the number of elements of the basis is said the dimension of *V* (potentially it could not be finite)

- but not the orthogonal ones



Preliminary Notation

- A basis $\{b_i\}_{i=1,..,N}$ of a vector space V is a **set of linearly independent vectors** that *spans* (i.e. that can represent with linear combinations) the whole vector space V.
  - *N,* the number of elements of the basis is said the dimension of V (potentially it could not be finite)

Preliminary Notation

- A basis $\{b_i\}_{i=1,..,N}$ of a vector space *V* is a **set of linearly independent vectors** that *spans* (i.e. that can represent with linear combinations) the whole vector space *V*.
  - *N,* the number of elements of the basis is said the dimension of *V* (potentially it could not be finite)

- Any vector space
  - admit infinite basis,
  - each element of the vector space admit a unique representation w.r.t. each basis.

$$v = \sum_{i=1}^{N} a_i b_i \quad a_i \in \Re,\ \forall v \in V$$

Preliminary Notation

- A basis $\{b_i\}_{i=1,..,N}$ of a vector space *V* is a **set of linearly independent vectors** that *spans* (i.e. that can represent with linear combinations) the whole vector space *V*.
  - *N,* the number of elements of the basis is said the dimension of *V* (potentially it could not be finite)

- Any vector space
  - admit infinite basis,
  - each element of the vector space admit a unique representation w.r.t. each basis.

$$v = \sum_{i=1}^{N} a_i b_i \quad a_i \in \Re, \; \forall v \in V$$

- In signal/image processing, basis allows us to represent signal or images by means of their coefficients

$$v \rightarrow \{a_i\}_{i=1,..,N}$$

but....

- ... if one needs a basis to represent a signal, how did we manage them so far?

- Which is the *canonical* basis for digital signals/images?

116	23	33
16	3	73
5	4	30

but....

- ... if one needs a basis to represent a signal, how did we manage them so far?

- Which is the *canonical* basis for digital signals/images?

116	23	33
16	3	73
5	4	30

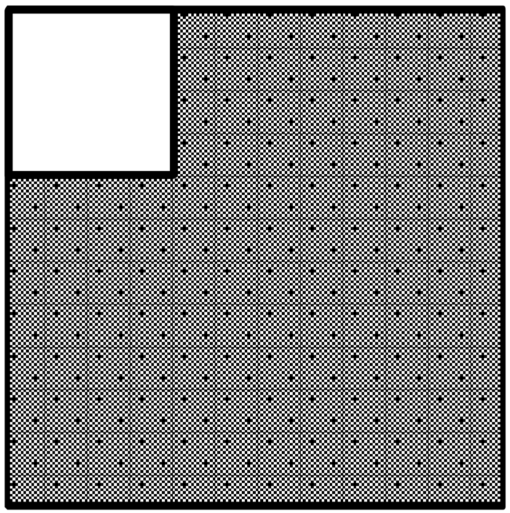
1	0	0
0	0	0
0	0	0

but....

- ... if one needs a basis to represent a signal, how did we manage them so far?

- Which is the *canonical* basis for digital signals/images?

116	23	33
16	3	73
5	4	30

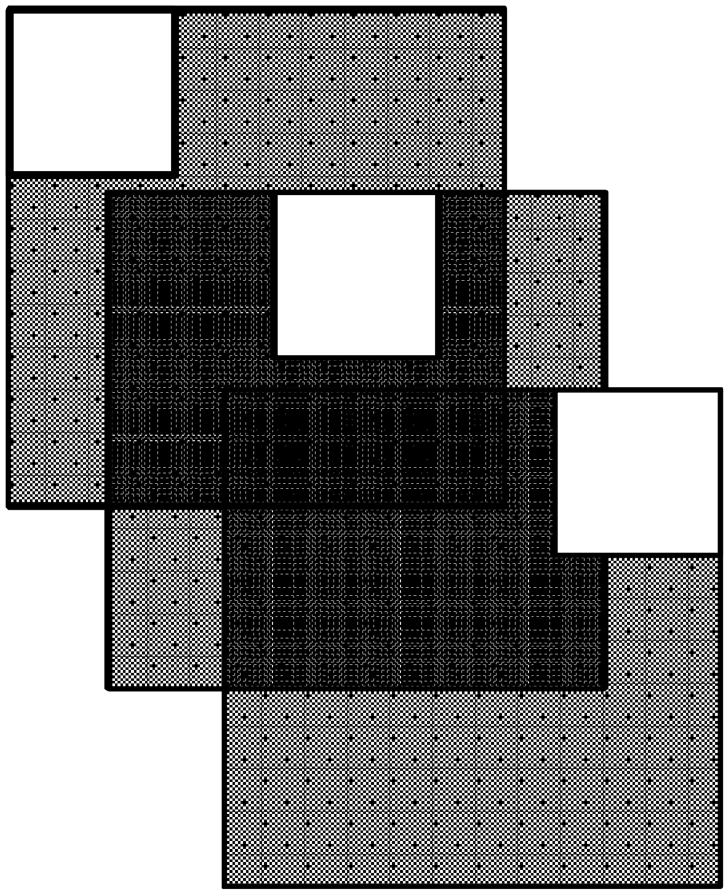


but....

- ... if one needs a basis to represent a signal, how did we manage them so far?

- Which is the *canonical* basis for digital signals/images?

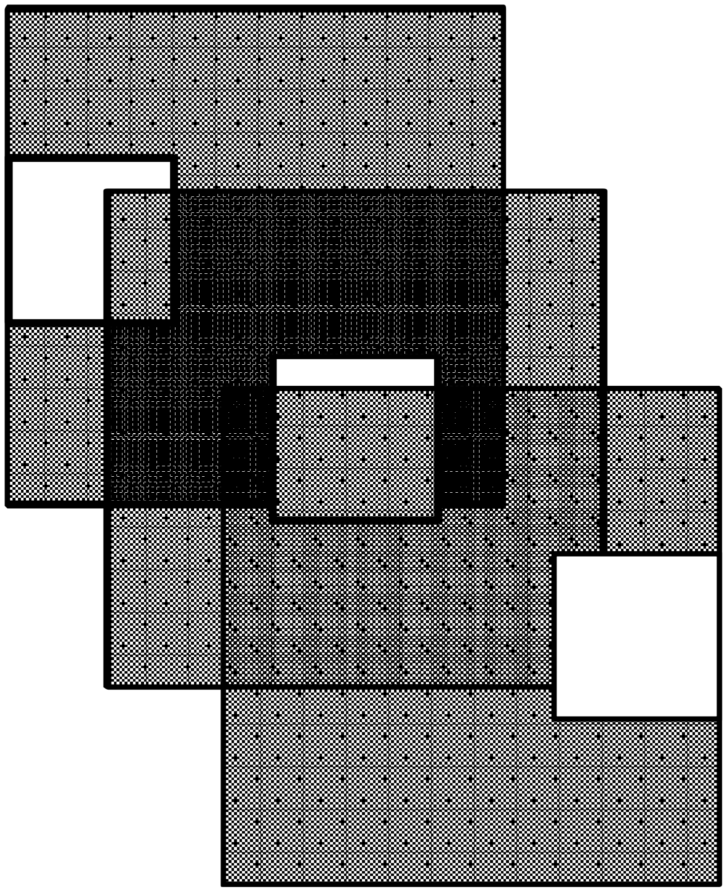
116	23	33
16	3	73
5	4	30



but....

- ... if one needs a basis to represent a signal, how did we manage them so far?
- Which is the *canonical* basis for digital signals/images?

116	23	33
16	3	73
5	4	30

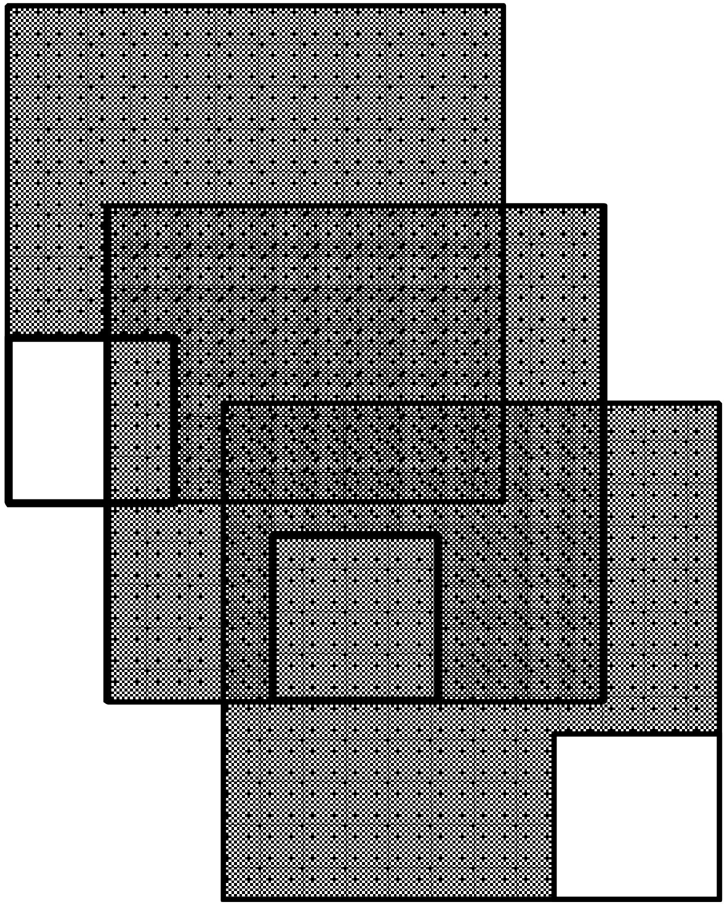


but....

- ... if one needs a basis to represent a signal, how did we manage them so far?

- Which is the *canonical* basis for digital signals/images?

116	23	33
16	3	73
5	4	30



Thus the canonica basis

- Uses each coefficient to represent a pixel:
  - all coefficients are equally meaningful
  - thus, it is not usefull at all for compression
  - it corresponds to the canonical basis of $\Re^N$

- Are there basis that ease image processing tasks?

Orthonormal Basis

- When V is an Hilbert space, the inner product allows us to define the **orthonormal basis** as a basis having orthonormal elements, i.e.

$$< b_i, b_j >= \delta_{i,j} \quad \forall i, j$$

- For orthonormal basis we have that if $\{b_i\}_{i=1,\ldots N}$ is a orthonormal basis,

$$v = \sum_{i=1}^{N} a_i b_i \quad a_i \in \Re, \ \forall v \in V$$

means $a_i =< v, b_i >$.

- Thus we know how to change basis, i.e. how to transform the signals/images w.r.t. a different basis.

- Furthermore, **such corresponding transforms are linear (thus invertible)**

2D Fourier Transform

- The *(u,v)*-element of the 2D Fourier basis is defined as

$$e^{-i2\pi(ux+vy)} = \cos(2\pi(ux+vy)) + i\sin(2\pi(ux+vy))$$

- Each Fourier coefficient is computed with an inner product with the correspondinf function.
- The Fourier basis functions are constant where $y = \frac{ux}{v} + c$
- The Fourier basis functions have unlimited support.
- The Fourier Transform is invertible (it is an orthonormal transform)
- Fourier domain is also called frequency domain.

2D Discrete Fourier Transform

- The Discrete Fourier Transform (DFT) is defined as

$$F[m,n]=\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} f[k,l]e^{-\pi i\left(\frac{km}{M}+\frac{\ln}{N}\right)}$$

- The Fourier Transform admit a fast implementation (FFT) when the signal/image sizes are powers of 2.

2D Fourier Basis Elements

- Frequency and Orientation of the 2D Fourier Basis Elements

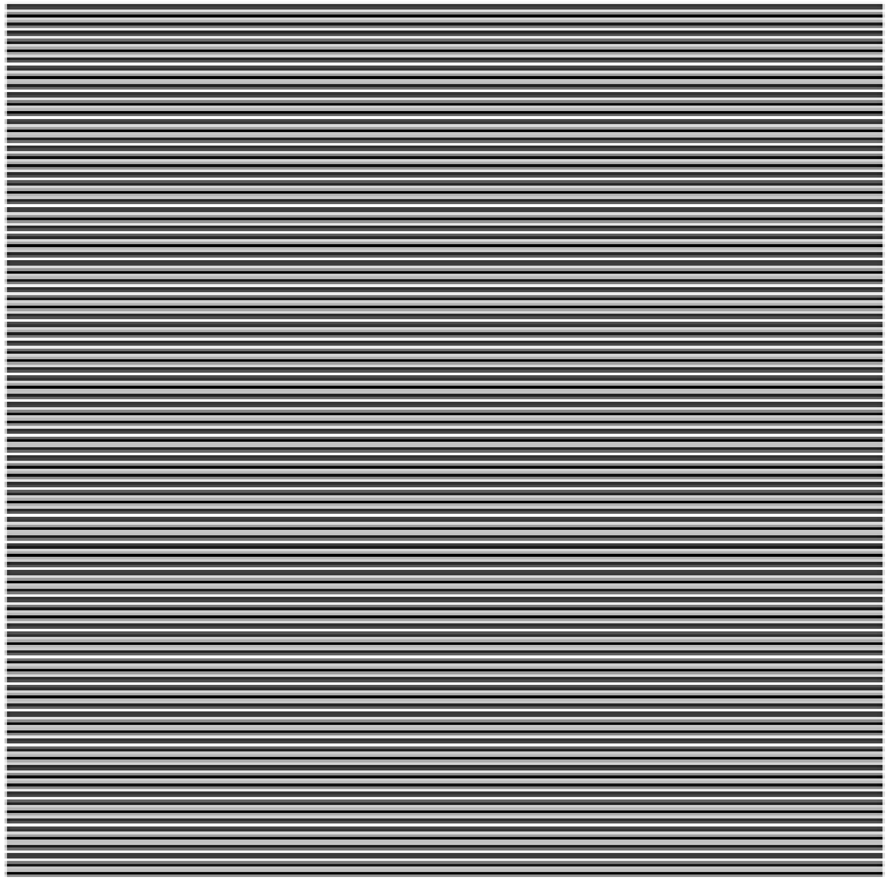
$$b_{u,v} \quad u = 1, v = 10$$

2D Fourier Basis Elements

- Frequency and Orientation of the 2D Fourier Basis Elements

$$b_{u,v} \quad u = 100, v = 1$$

the u=100v=1Fourier domain basis element



2D Fourier Basis Elements

- Frequency and Orientation of the 2D Fourier Basis Elements

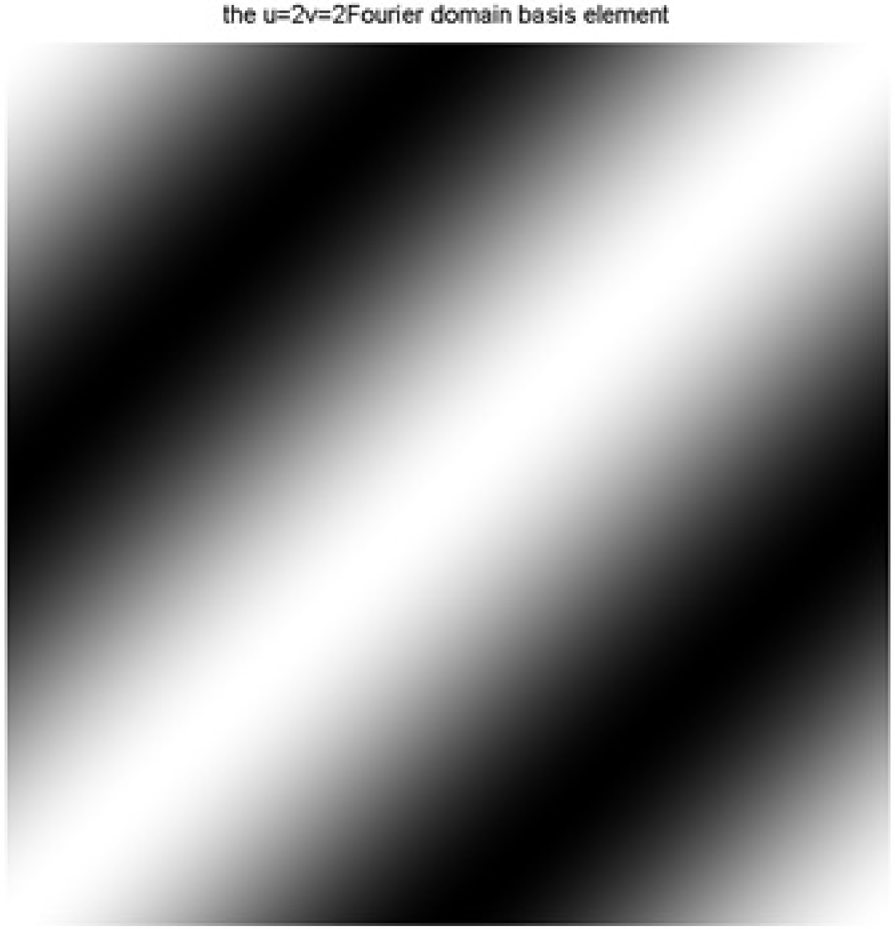
$$b_{u,v} \quad u = 10, v = 10$$

2D Fourier Basis Elements

- Frequency and Orientation of the 2D Fourier Basis Elements

$$b_{u,v} \quad u = 2, v = 2$$

(note that in matlab Fourier coefficients are indexed starting from 1)



Try it in Matlab

```
%Examples of Basis Elements
Y=zeros(512);
u=10
v=10
Y(u,v)=1;
% Y is the image in Fourier Domain having only the (u,v)-coefficient set
    to 1 and the other set to 0
figure(1),imshow(Y,[]),title(['the  u=',num2str(u),  'v=',num2str(v),  'space
    domain basis element']  )
 % by inverting the Fourier transform one get the (u,v)-element of the 2D
    discrete Fourier basis
y=ifft2(Y);
figure(2),imshow(real(y),[]),title(['the  u=',num2str(u),  'v=',num2str(v),
    'Fourier domain basis element']  )
```

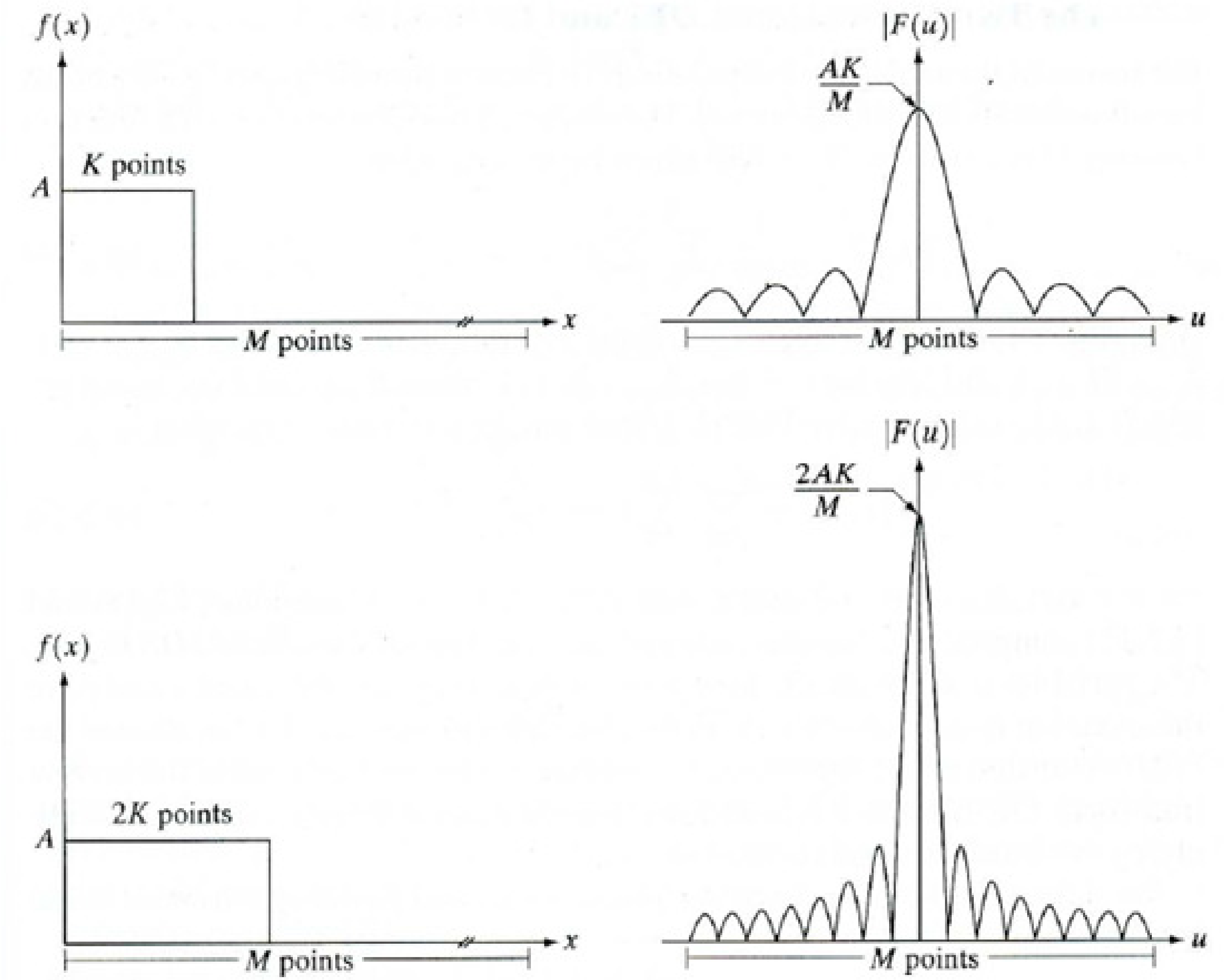
Properties of Fourier Coefficients

- The Fourier Transform is a Global Transform: every image pixel influences the value of each Fourier coefficient.
- The (0,0) coefficient corresponds to the image average.
- Thus not all coefficients are "equally meaningful" in the image representation.

Fourier Transform Properties

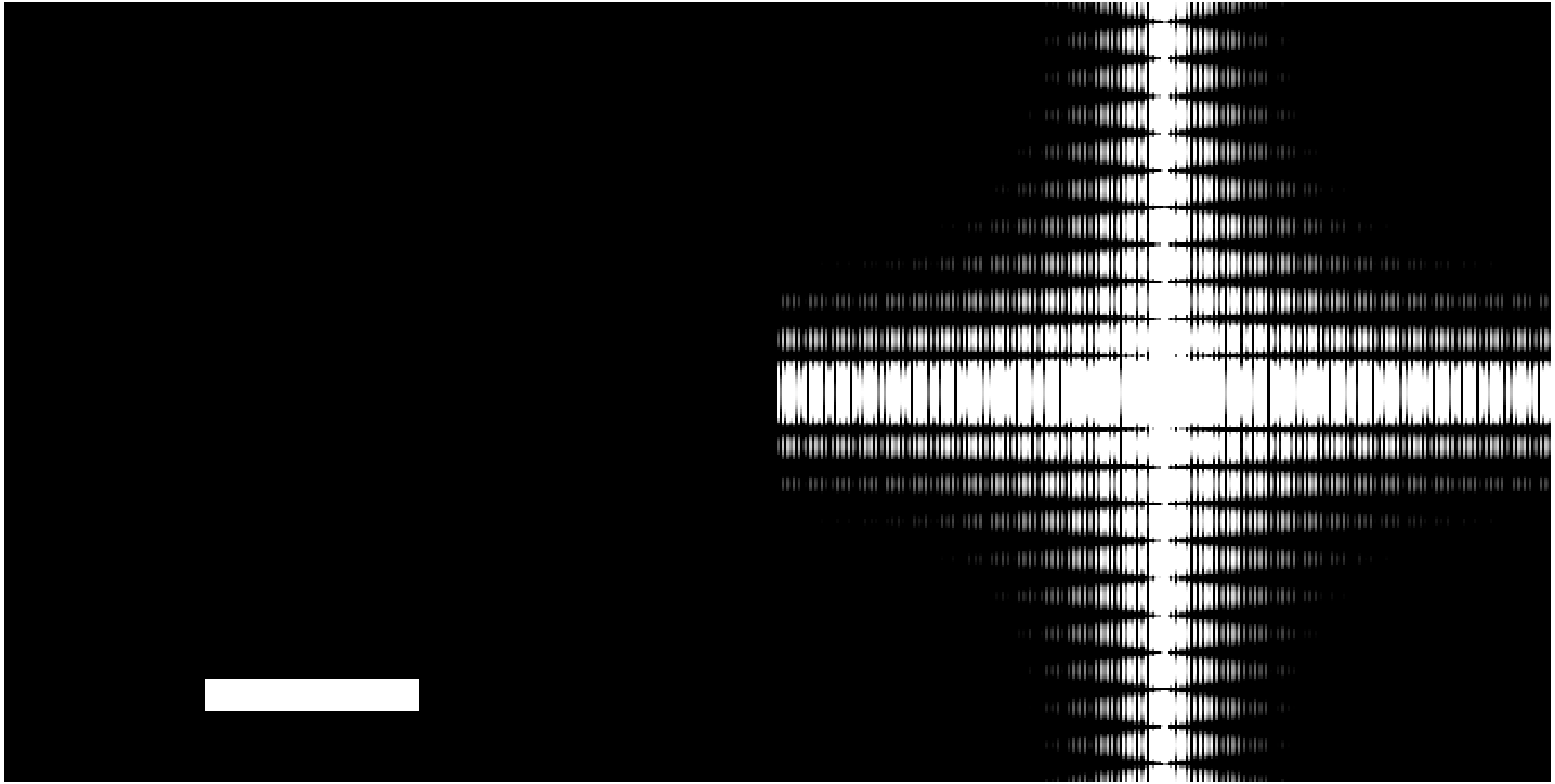
Property	Function	Fourier Transform	
	$f(t)$	$\hat{f}(\omega)$	
Inverse	$\hat{f}(t)$	$2\pi f(-\omega)$	(2.15)
Convolution	$f_1 \star f_2(t)$	$\hat{f}_1(\omega)\hat{f}_2(\omega)$	(2.16)
Multiplication	$f_1(t) f_2(t)$	$\frac{1}{2\pi}\hat{f}_1 \star \hat{f}_2(\omega)$	(2.17)
Translation	$f(t-u)$	$e^{-iu\omega}\hat{f}(\omega)$	(2.18)
Modulation	$e^{i\xi t} f(t)$	$\hat{f}(\omega-\xi)$	(2.19)
Scaling	$f(t/s)$	$\|s\|\,\hat{f}(s\omega)$	(2.20)
Time derivatives	$f^{(p)}(t)$	$(i\omega)^p \hat{f}(\omega)$	(2.21)
Frequency derivatives	$(-it)^p f(t)$	$\hat{f}^{(p)}(\omega)$	(2.22)
Complex conjugate	$f^*(t)$	$\hat{f}^*(-\omega)$	(2.23)
Hermitian symmetry	$f(t) \in \mathbb{R}$	$\hat{f}(-\omega) = \hat{f}^*(\omega)$	(2.24)

Function supports in Fourier and Space Domain



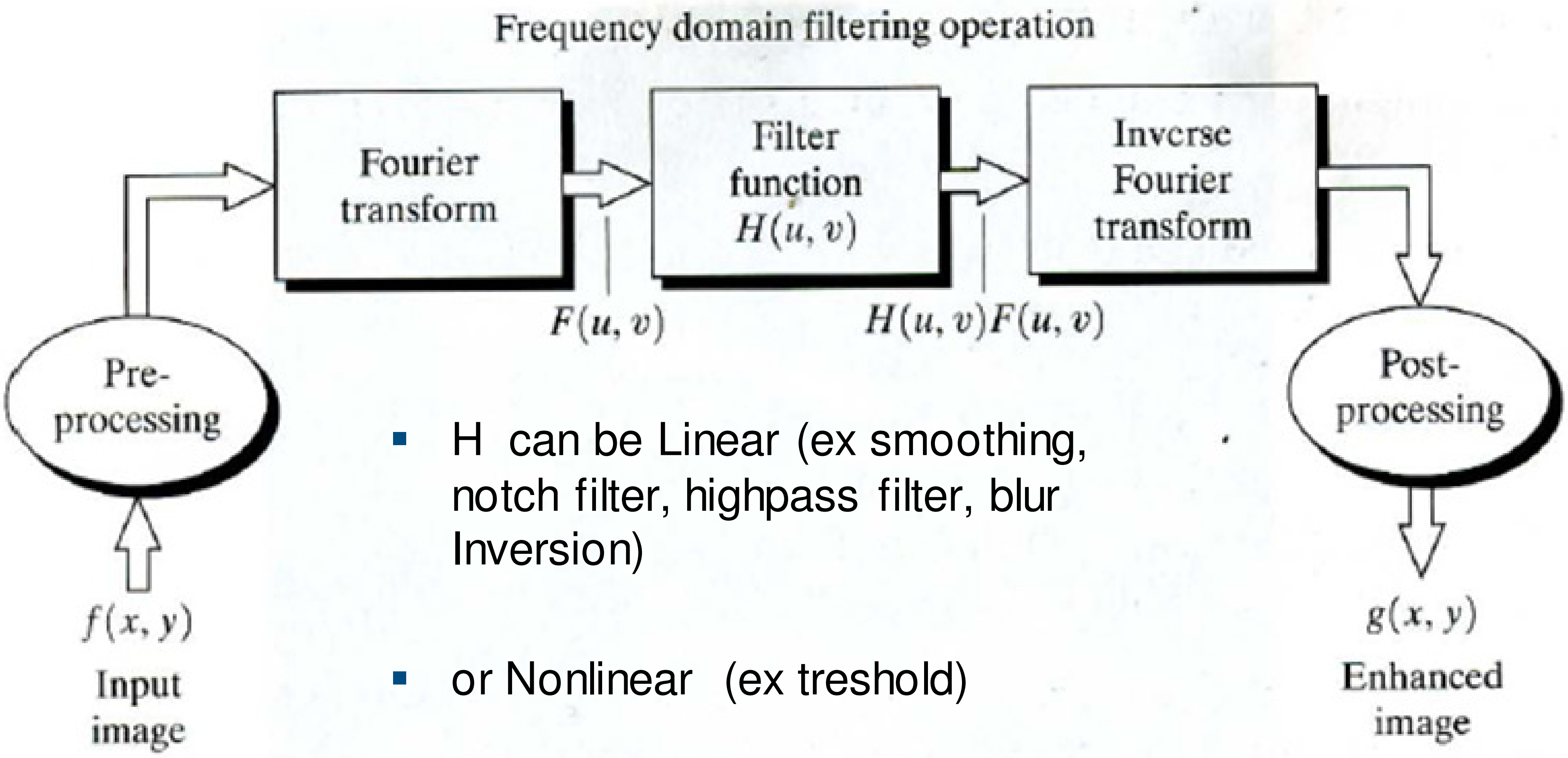
Shifted zero-frequency fourier transform

I , fftshift(log(fft2(I)))



Fourier Domain Image Processing

- The typical apprach is as follows



- H can be Linear (ex smoothing, notch filter, highpass filter, blur Inversion)
- or Nonlinear (ex treshold)

Denoising via Convolution

- In Fourier domain a 2D convolution a gainst a kernel *h*

  $$g = (f \circledast h)$$

  corresponds to element-wise (i.e. pixel-wise) multiplication, i.e.

  $$G = F\,H$$

  being ***G, F***, and ***H*** the Fourier transform of f, g, h, respectively

- Thus filtering in space domain with a kernel, corresponds to masking the Fourier transform of the original signal/image.

- In particular, when the kernel *h* is a Gaussian, ***H*** is also a Gaussian

  $$f(x) = e^{-ax^2} \qquad F(\omega) = e^{-\omega^2/a}$$

  then Gaussians having small σ in space domain corresponds to large Gaussians having large σ in Fourier Domain

Denoising via Convolution

- In particular, when the kernel h is a Gaussian, $\boldsymbol{H}$ is also a Gaussian

$$f(x) = e^{-ax^2} \qquad F(\omega) = e^{-\omega^2/a}$$

then Gaussians having small σ in space domain corresponds to large Gaussians having large σ in Fourier Domain.

  - Small σ means smoothing few pixels in space domain and shrink few coefficients in Fourier Domain
  - Large σ means smoothing a lot of pixels in space domani, and shrink a lot coefficients in time domain

The Deblurring Problem

- The convolution theorem states that

$$Fourier((f \circledast g)) = FG$$

  It follows that:
  - Convolution can be computed by performing
    - Fourier transofrm,
    - Element-wise product
    - Inverse Fourier transform

    (and if FFT is possible, this is less expensive than computing convolutions in space domain)

  - Blur can be easily inverted in Fourier domain

- Note that the convolution theorem holds when the signal is periodic, and thus the **circular convolution** has to be computed

The Deblurring Problem

- Given the image *y* and the kernel *h* a blurred observation is given by

$$z = (y \circledast h)$$

- then, whenever the filter *h* is exactly known, the convolutional blur can be inverted in Fourier Domain as we have

$$Z = Y H$$

- and thus

$$\widehat{y} = Fourier^{-1}\left(\frac{Z}{H}\right)$$

- in order to "avoid problems" in frequencies where ***H****=0*

$$\widehat{y} = Fourier^{-1}\left(\frac{Z\bar{H}}{\|H\|^2 + \epsilon}\right)$$

being $\epsilon > 0$ a regularization parameter (difficult to chose in practice).

The Deblurring Problem

- but what's up when noise appears:

$$z = (y \circledast h) + \eta$$

- then, when computing the Fourier transform of the observation we have

$$Z = YH + N$$

- and thus

$$\widehat{y} = Fourier^{-1}\left(\frac{Y\bar{H}}{\|H\|^2 + \epsilon} + \frac{N\bar{H}}{\|H\|^2 + \epsilon}\right)$$

.. thus even unperceptible amount of noise, may become problematic in the second term of the sum.

A Brief Introduction on Wavelets

Giacomo Boracchi
giacomo.boracchi@polimi.it

ARTIFICIAL INTELLIGENCE and ROBOTICS LABORATORY

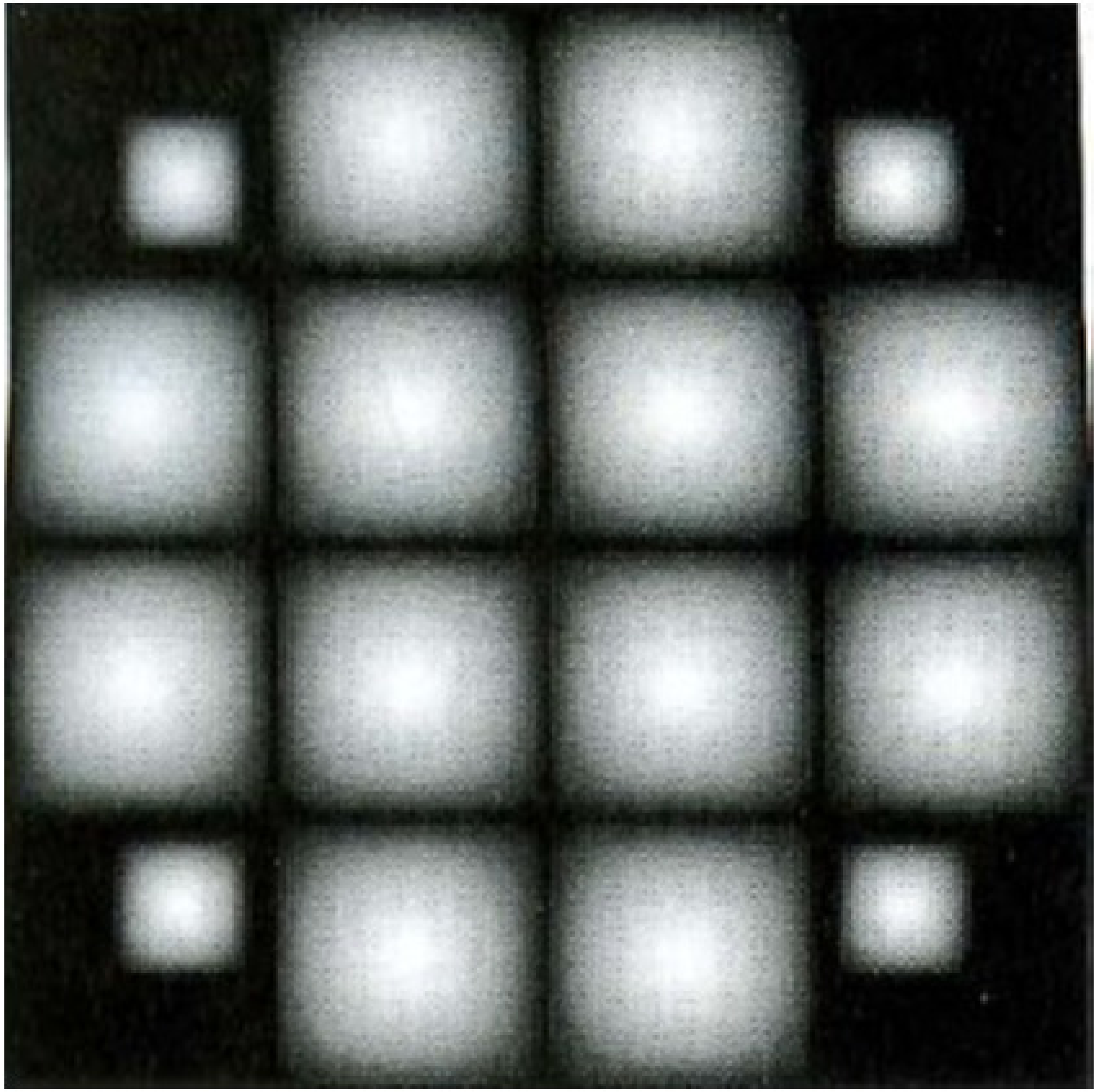
Why wavelets

- Fourier coefficients are global, while we would like transform coefficients to be **local**
  - Thus they provide a signal description i a neighbor of each pixel
  - Thus a change of a few pixels in the original image would determine a change in few coefficients

- We would like different coefficients to yield information about the behavior of the image
  - w.r.t. some particular frequency bands
  - along different directions
  - at different resolution levels

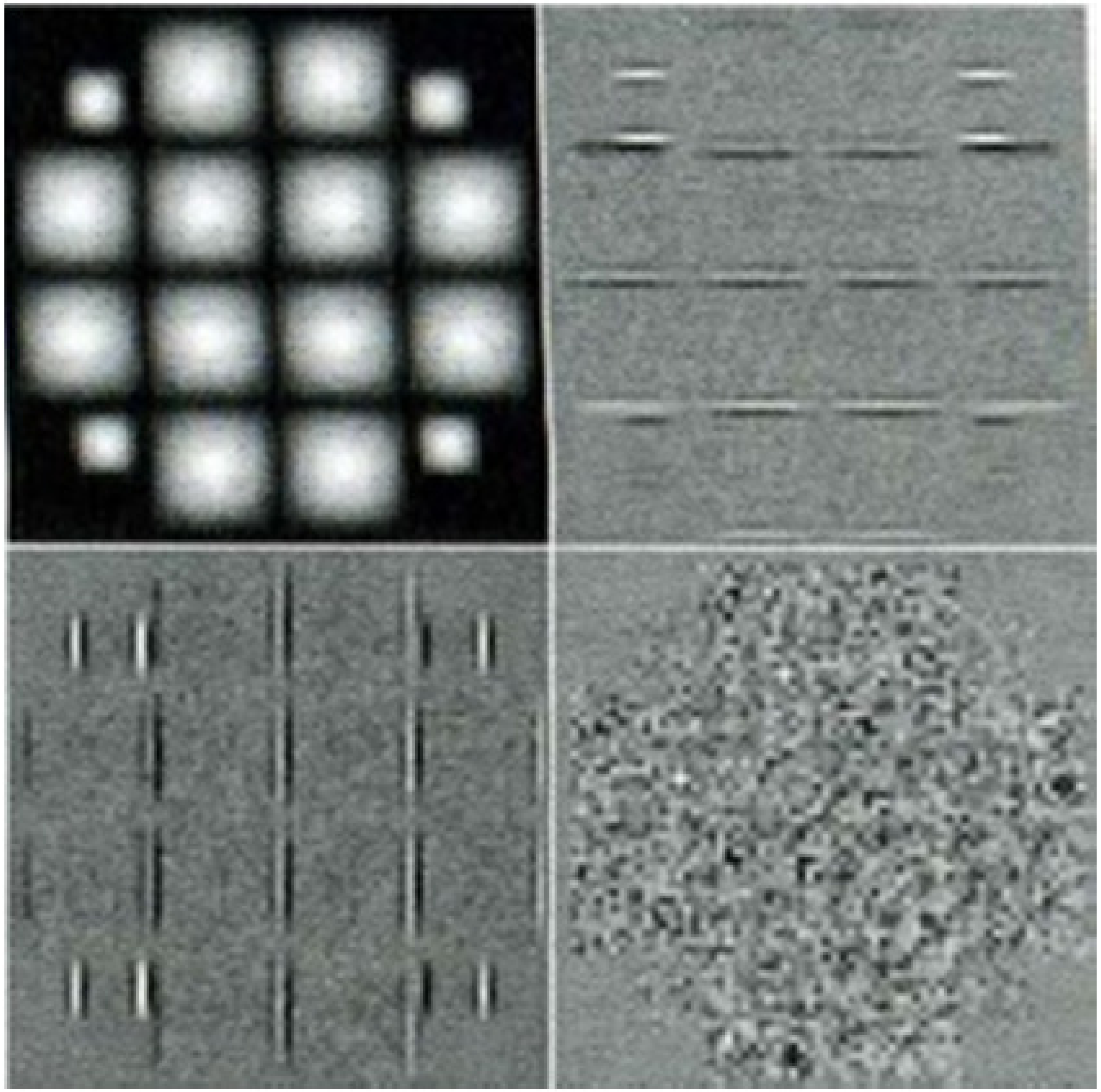
An example,

- given such an imput image



An example,

- one would obtain a similar representation, where each coefficient is
  - determined by the values of the image in a neighbor of each pixel
  - have different meaning such as
    - horizontal detail
    - vertical detail
    - low resolution representation of the image



What are Wavelets?

- Essentially wavelets are basis functions
- Usually one refers to wavelets meaning the Wavelet Transform (both continuous or discrete)
- It is possible to express each function $f \in L^2(\Re)$ with respect to a given wavelet basis

Basis For What?

- We introduce first the Multi Resolution Analysis (MRA)
- A MRA is a collection of nested subspaces of functions (usually belonging to $L^2(\Re)$) which is described by a *scaling function* ϕ and a *wavelet function* ψ.
- A subspace in the MRA is described by the scaling function at a resolution level i.e.

MRA

- A function f belonging to V_{j_0} can be written as

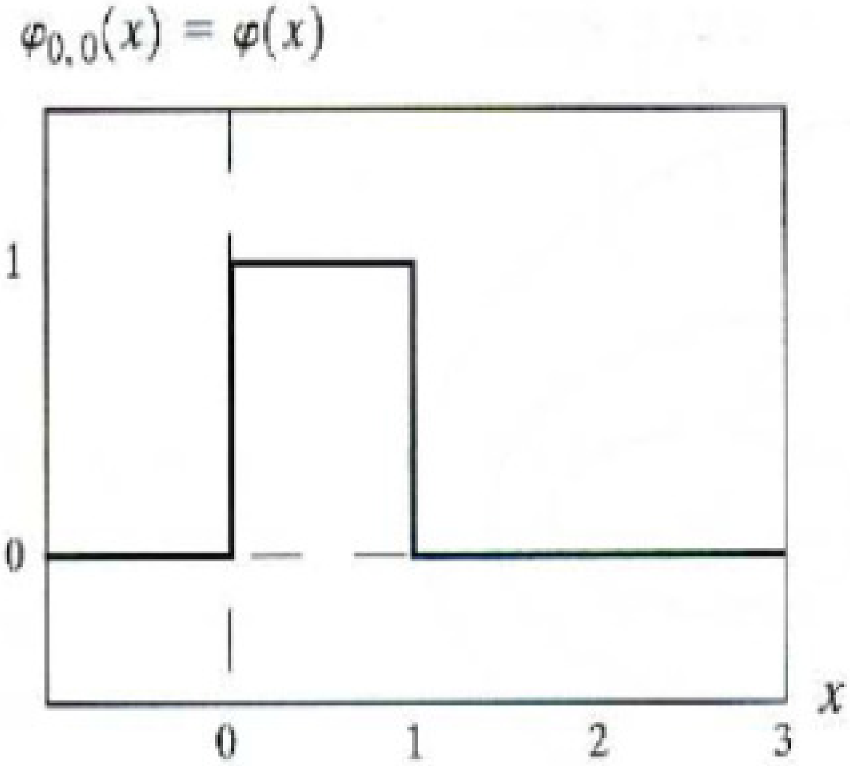
$$f(x) = \sum_k \alpha_k \varphi_{j_0,k}(x).$$

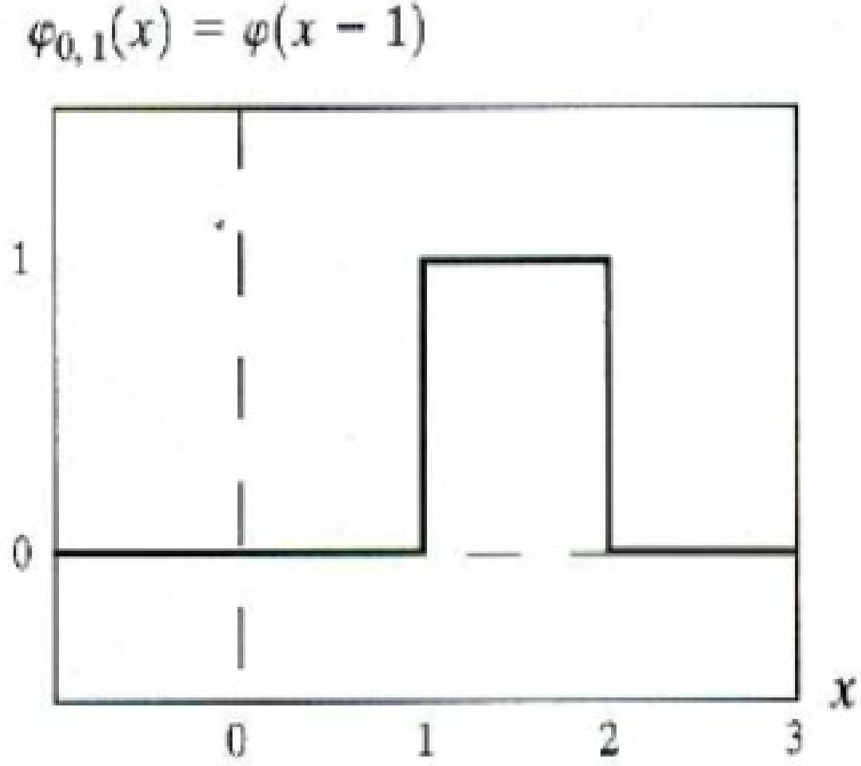
- and this is the only representation because the collection of

$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$ is an (orthonormal) basis for V_{j_0}

A classical Example the Haar Scaling Function

V_0

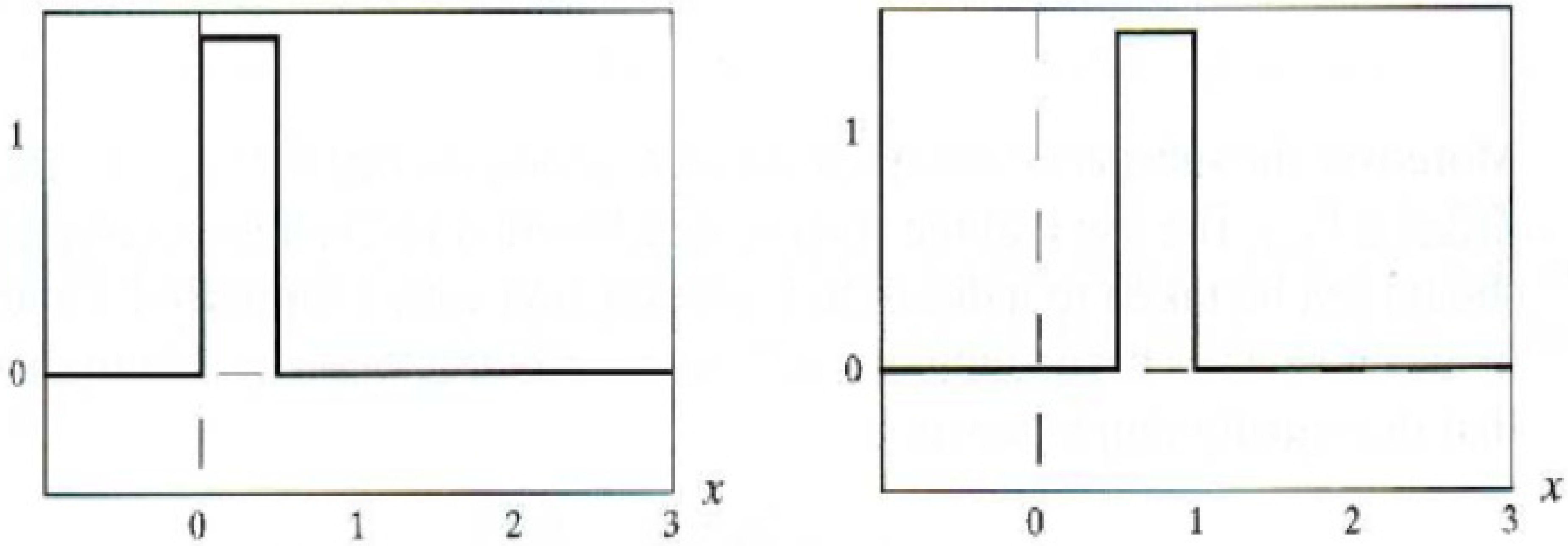




A classical Example the Haar Scaling Function

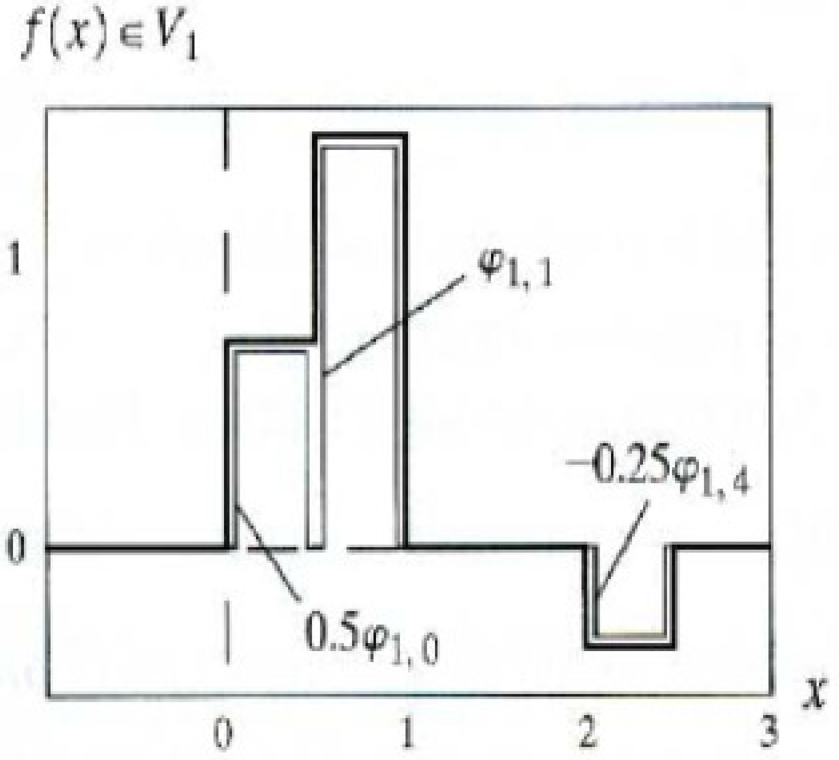
V_1

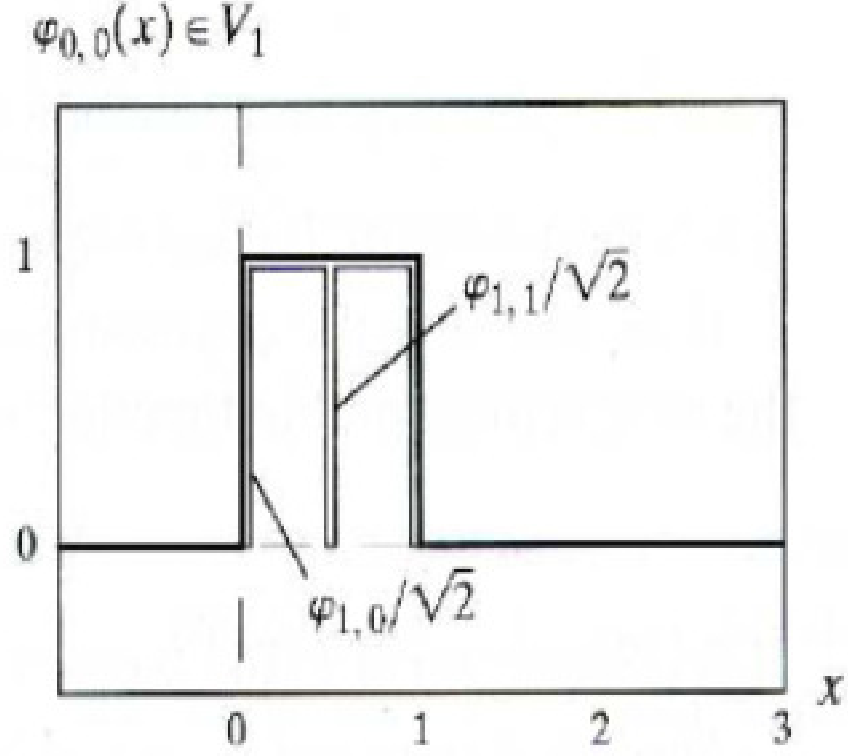
$\varphi_{1,0}(x) = \sqrt{2}\,\varphi(2x)$ $\varphi_{1,1}(x) = \sqrt{2}\,\varphi(2x-1)$



A classical Example the Haar Scaling Function

V_1 $V_0 \subset V_1$





Basis For What?

- A Multi Resolution Analysis (MRA) is a collection of nested subspaces of functions (usually belonging to $L^2(\Re)$) which is described by
  - the *scaling function* ϕ
  - the *wavelet function* ψ.

- A subspace in the MRA is described by the scaling function at a given resolution level j_0 i.e.

$$V_{j_0} = \overline{\operatorname*{Span}_k \{\varphi_{j_0,k}(x)\}}. \qquad \varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$$

- V_{j_0} does not span the whole $L^2(\Re)$ but it is just a subspace.
- In the previous example the spatial resolution of V_{j_0} is 2^j

MRA requirements

- All the integer translates of the scaling function at fixed resolution form an (orthonormal) basis

- The subspaces spanned by the scaling function at low scales are nested within those spanned at highest scales

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

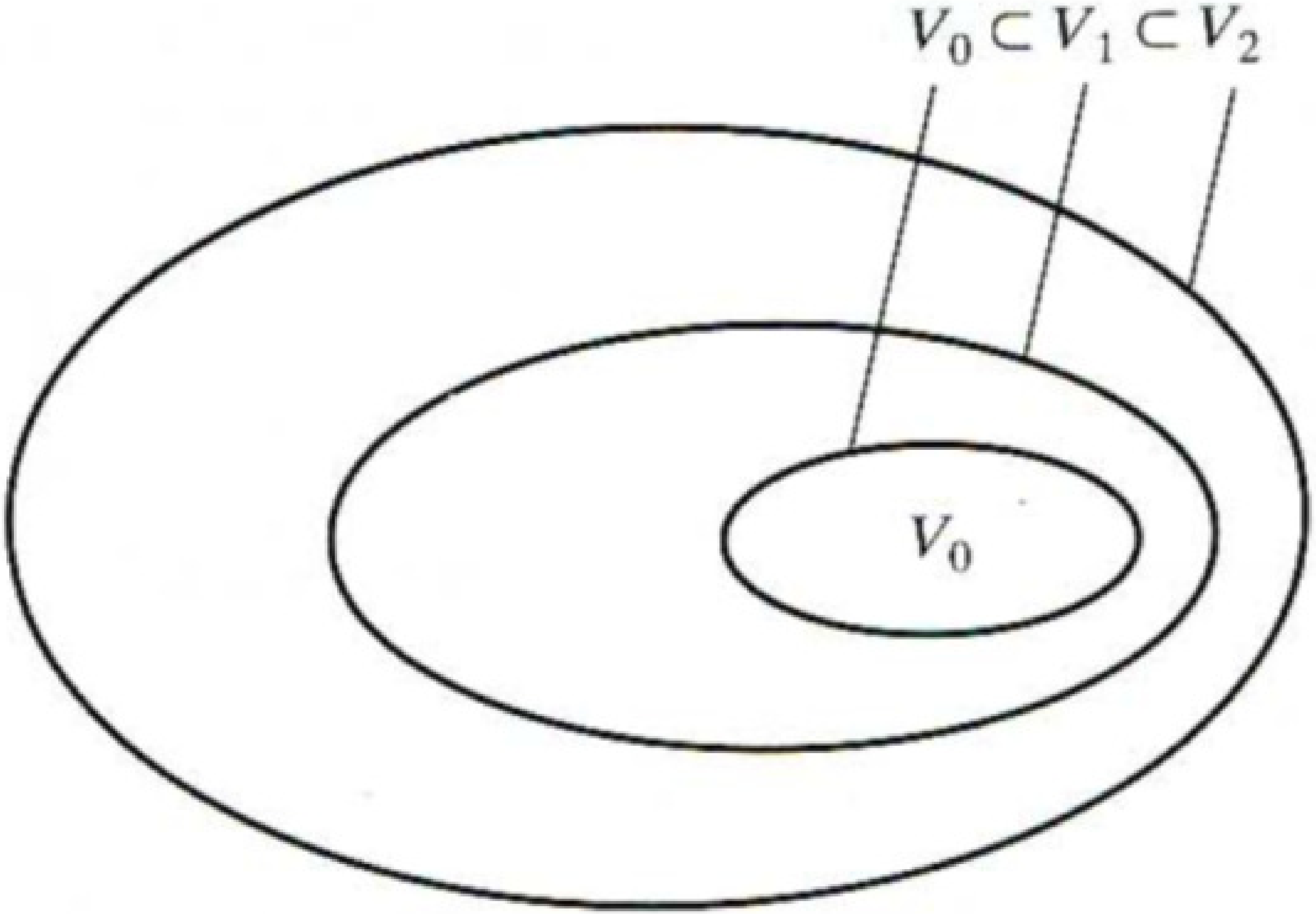
- The only function that is common to all V_1 is $f = 0$

- Any function $f \in L(\Re)$ can be represented at an arbitrary precision

$$\lim_{j \to \infty} V_j = L^2(\Re)$$

MRA

- Then some subspace of MRA can be represented as



(orthonormal) MRA refinement equation

- Express each basis function of a given MRA subspace V_j w.r.t the basis of the finer subspace, i.e. the basis of V_{j+1}

$$\varphi_{j,k}(x) = \sum_n \alpha_n \varphi_{j+1,n}(x)$$

- .. which are the scaling functions

$$\varphi_{j,k}(x) = \sum_n h_\varphi(n) 2^{(j+1)/2} \varphi(2^{j+1}x - n).$$

- h are the scaling filter coefficient

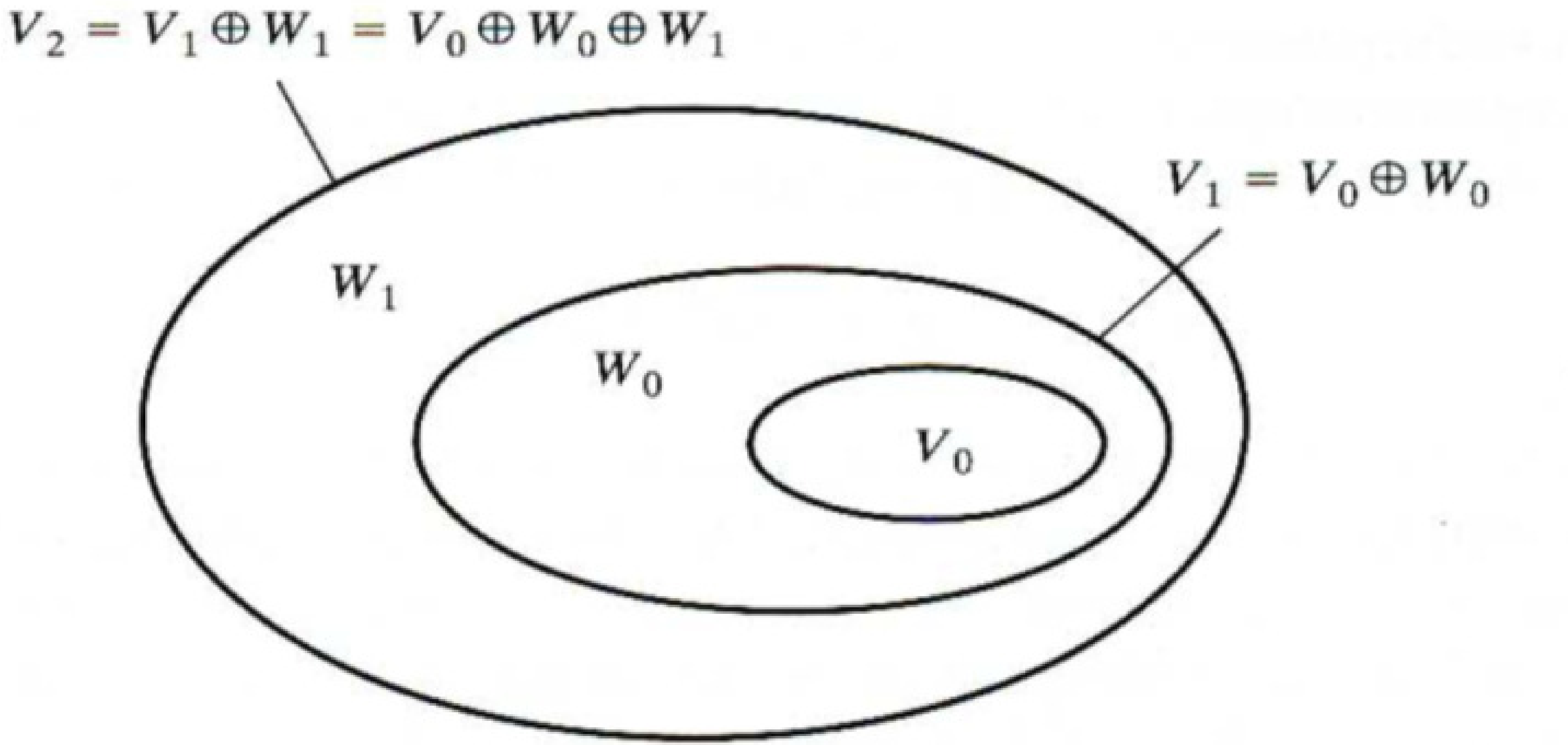
- in case of Haar basis we have

$$\varphi(x) = \frac{1}{\sqrt{2}}[\sqrt{2}\varphi(2x)] + \frac{1}{\sqrt{2}}[\sqrt{2}\varphi(2x-1)].$$

thus $h = [\sqrt{2}/2, \sqrt{2}/2]$

What about Wavelets

- Wavelets spans the differences between the any two adjacent scaling subspaces



Wavelets have a Mother Function too..

- Wavelets subspaces are also built by translating and scaling the mother function ψ

$$W_j = \overline{\operatorname{Span}_k \{\psi_{j,k}(x)\}} \qquad \psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$$

- And Subspaces in MRA are related by the following relation,

$$V_{j+1} = V_j \oplus W_j$$

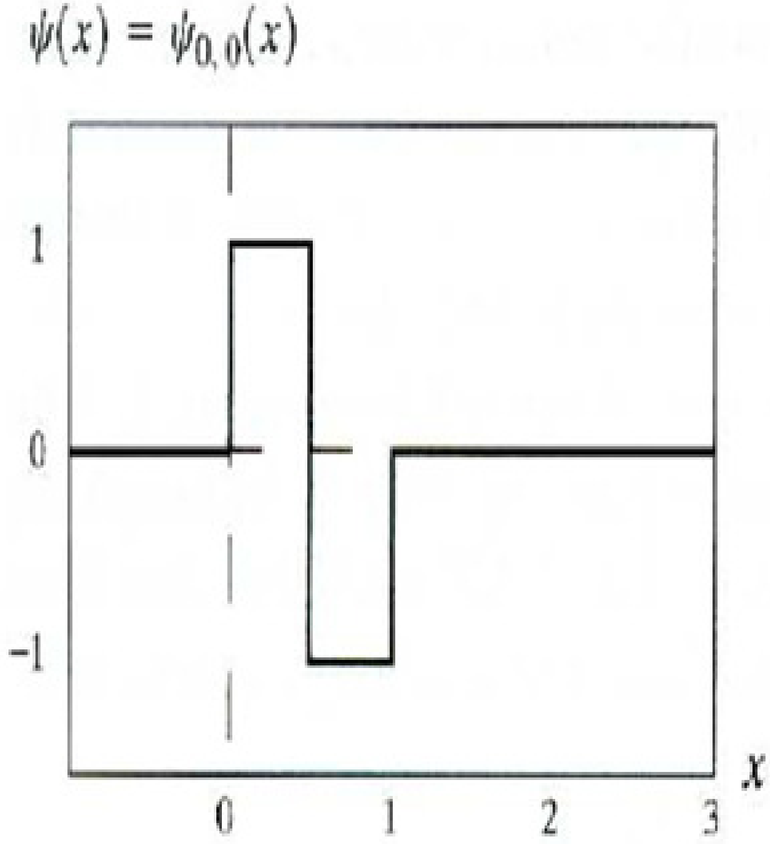
where $\oplus$ denotes the orthogonal complement (direct sum)

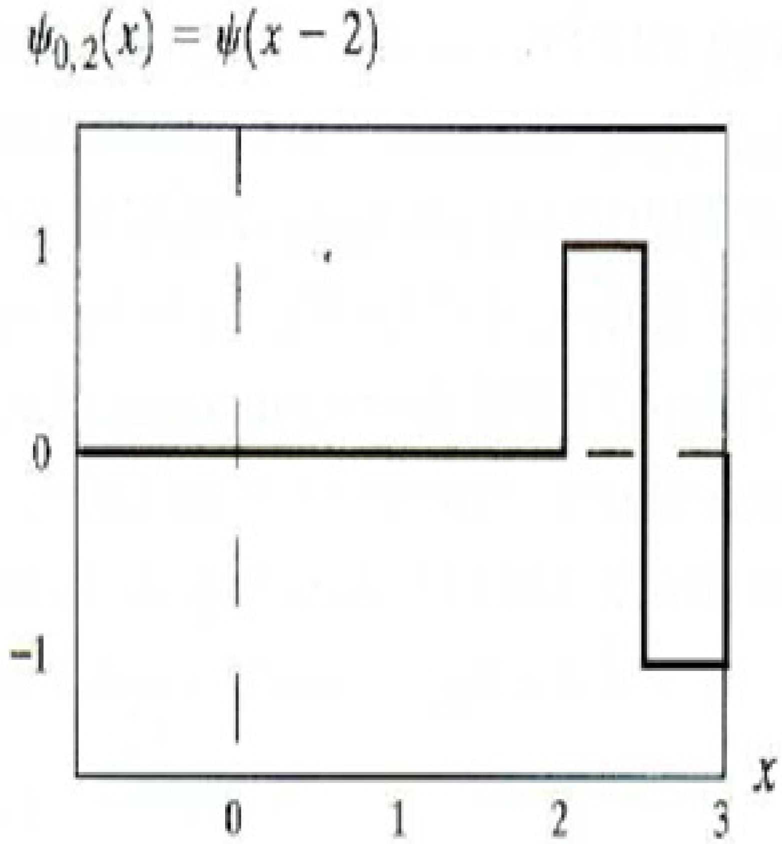
- Then

$$L^2(\mathbf{R}) = V_0 \oplus W_0 \oplus W_1 \oplus \ldots$$

A classical Example the Haar Wavelets

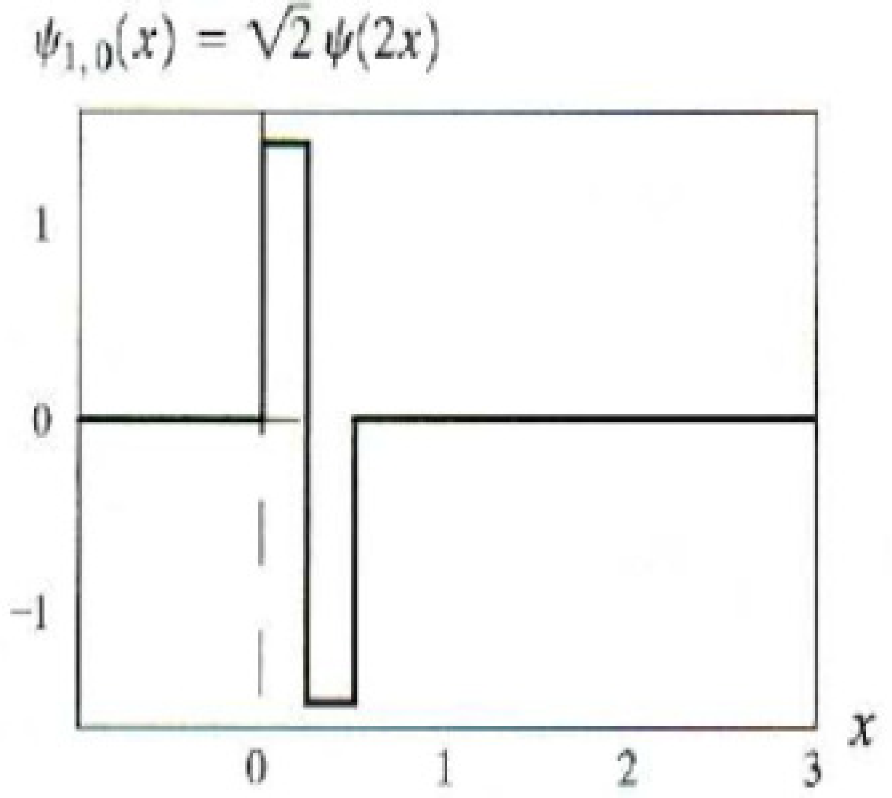
W_0



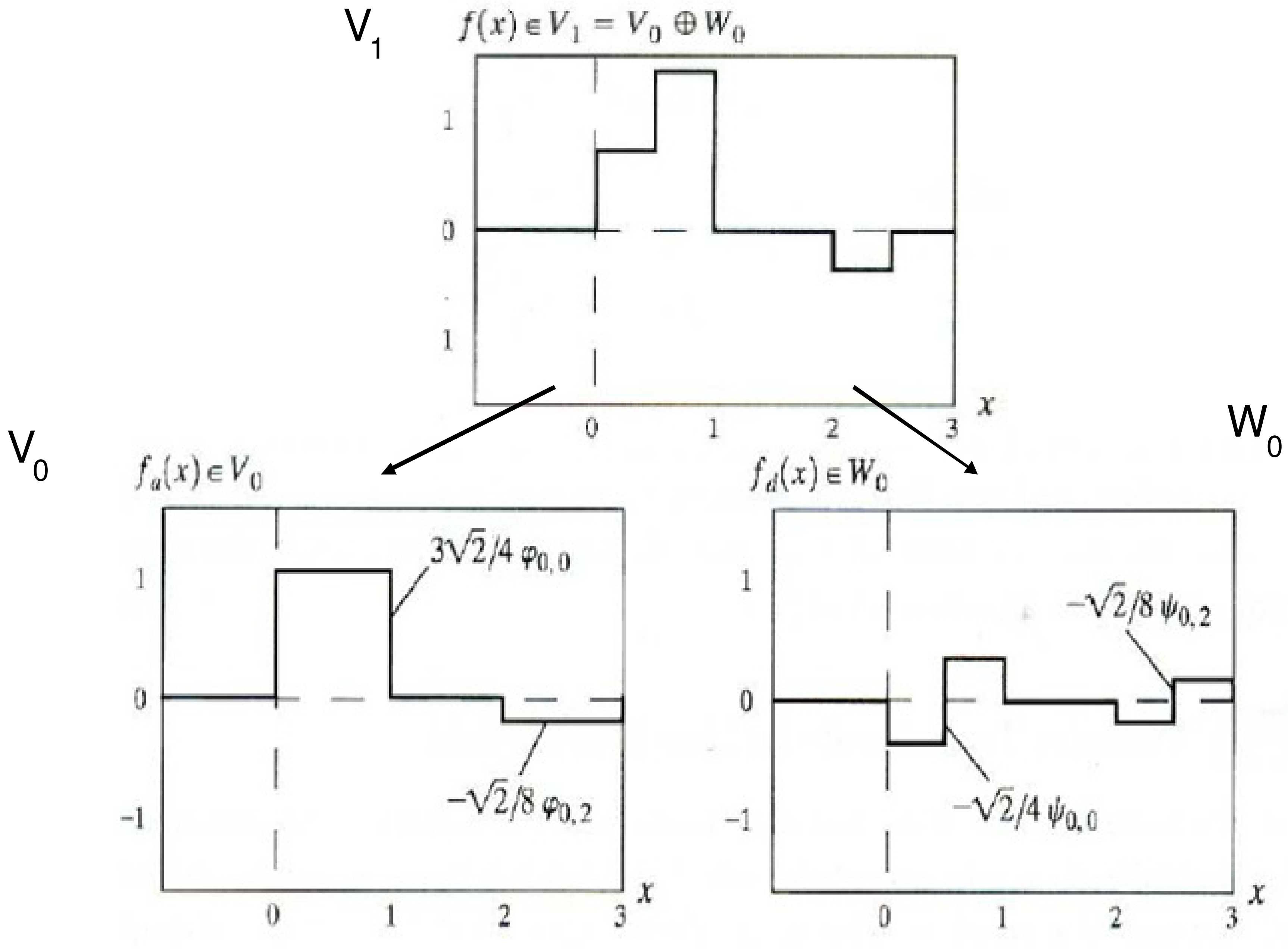


A classical Example the Haar Wavelets

W_1



A classical Example of Decomposition



Wavelet Filter

- The counterpart of the refinement equation holds for wavelets too

$$\psi(x) = \sum_n h_\psi(n)\sqrt{2}\varphi(2x - n)$$

- And it holds the following relation between scaling filter and wavelet filter coefficient

$$h_\psi(n) = (-1)^n h_\varphi(1 - n).$$

Wavelet Transform

- Express a function in a given subspace V_J w.r.t to a basis mixed of scaling and wavelet coefficient, according to the decomposition

$$L^2(\mathbf{R}) = V_0 \oplus W_0 \oplus W_1 \oplus \ldots$$

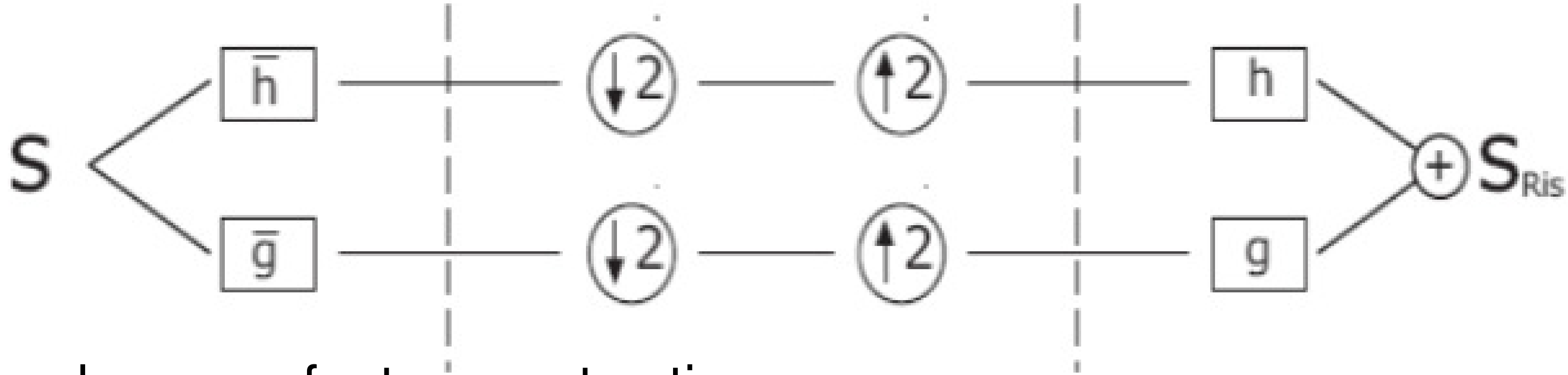
- And any function in MRA is usually written
  - as a coarse component (the approximation f_a)
  - and the details, the wavelets coefficients at **different scales** f_d

$$f(x) = f_a(x) + f_d(x)$$

$$\forall f \in L^2(\Re) \quad f_J = \sum_{k \in Z} a_0[k]\phi_{0,k} + \sum_{k \in Z;\, j=0}^{j=J-1} d_j[k]\psi_{j,k}\,.$$

Discrete Fourier Transform and Subband Coding

- Whenever filters *h* and *g* are given by an orthonormal MRA the this subband coding



produces perfect reconstruction

- And the if $S \in V_j$ then the coefficients

$$a_j[p] = \sum_{n=-\infty}^{+\infty} h[n-2p]a_{j+1}[n] = a_{j+1} \star \bar{h}[2p] \,,$$

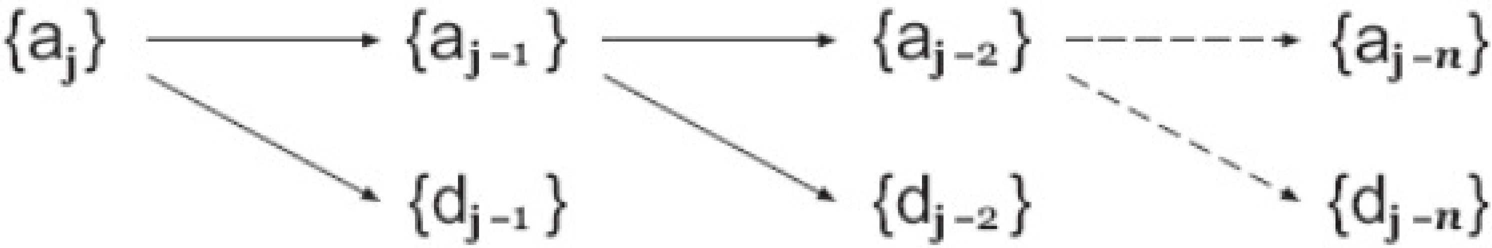
are scaling the coefficient in V_{j-1}

$$d_j[p] = \sum_{n=-\infty}^{+\infty} g[n-2p]a_{j+1}[n] = a_{j+1} \star \bar{g}[2p] \,.$$

and the wavelets coefficients in W_{j-1}

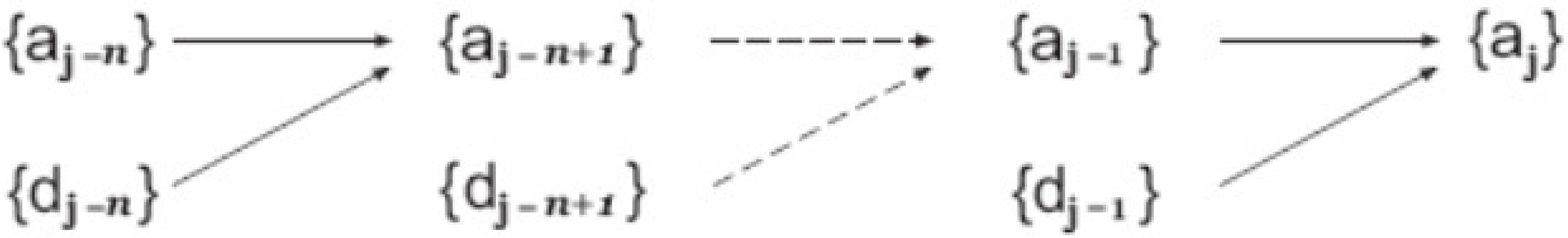
Iterative Subband Decomposition

- The Wavelet Expansion of a signal *S* is then given by



$$\left\{ \{d_j[n]\}_{j=1..J, n\in N}, \{a_k[n]\}_{n\in N} \right\}$$

- and the reconstruction is given by



I assure you… there are as many wavelets basis as you want

Wavelets and Signal Processing

- Wavelet Theory and Filter Design are strictly related
  - Mathematical Properties of Wavelets used to design digital filters
  - Transform Domain Methods are widely used in signal processing

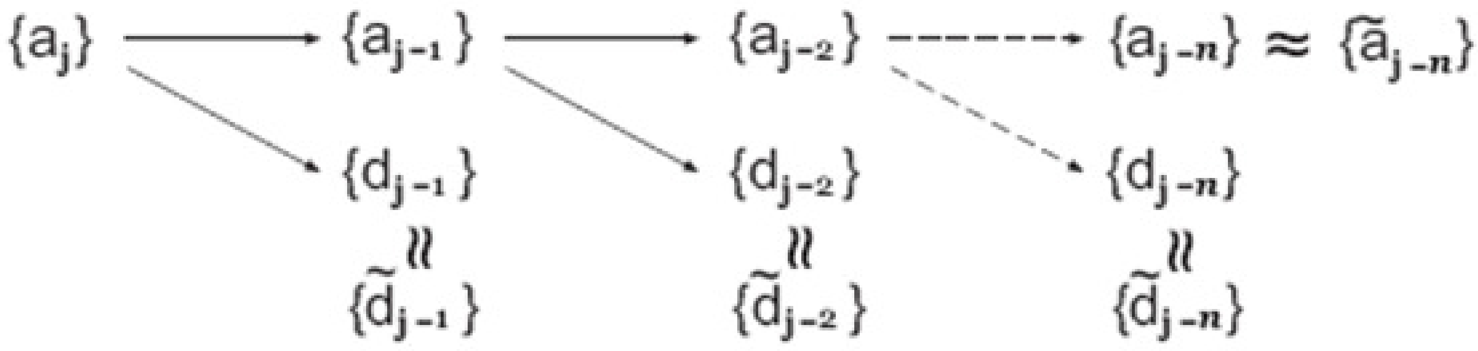
- Previous schemes hold for orthogonal wavelets, for biorthogonal wavelets and tight frames, as these have an explicit method both for decomposition and reconstruction

- The Lifting Scheme allows to generate (“by lifting”) an infinite family of biorthogonal wavelets filters (used in Jpeg2000) and offers a fast and easy interpretation as a filtering scheme

Why Wavelets in Signal Processing

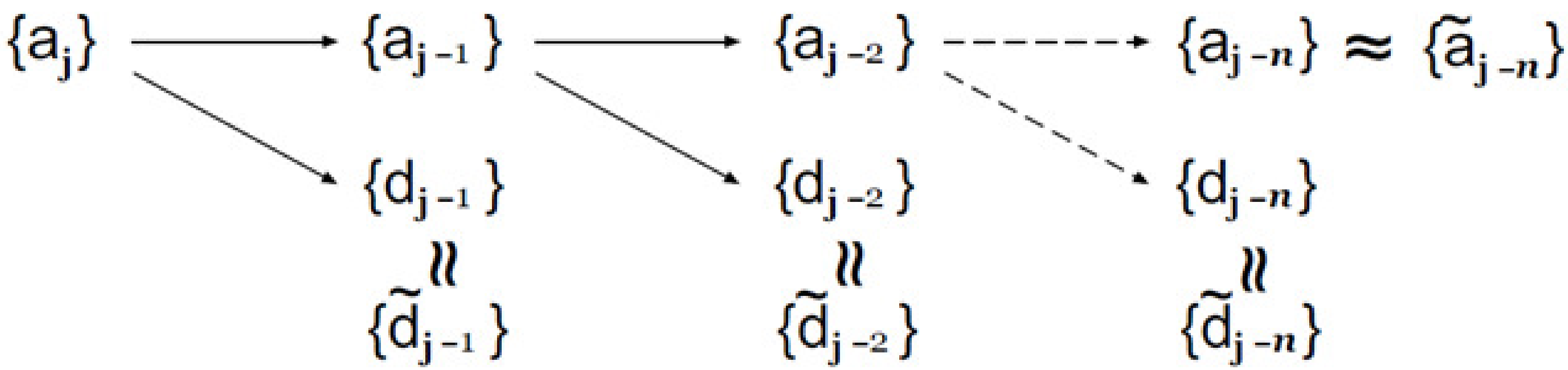
- Localized Support
  - Touching a Wavelets Coefficient modifies the signal only in a small (spatial) region. This does not happen for Fourier Transform

- Wavelets expansion is a **sparse** representation for some classes of signal, therefore can be used for compression.

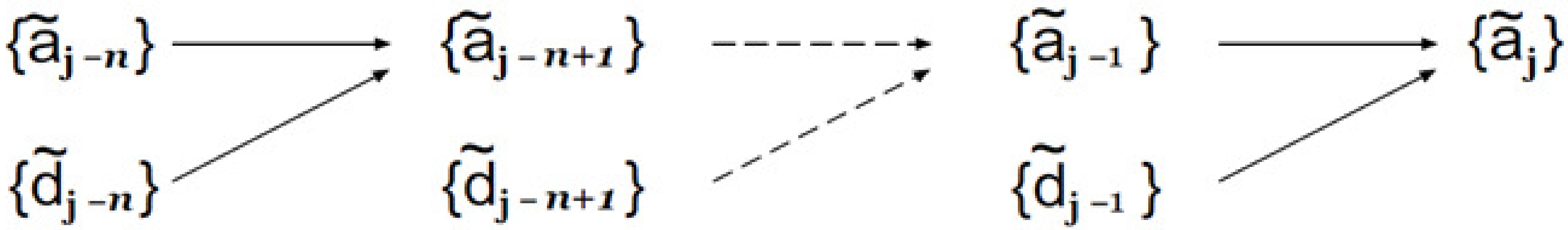


Basic Idea of Denoising / Compression

- Decomposition + Approximation



Restoration



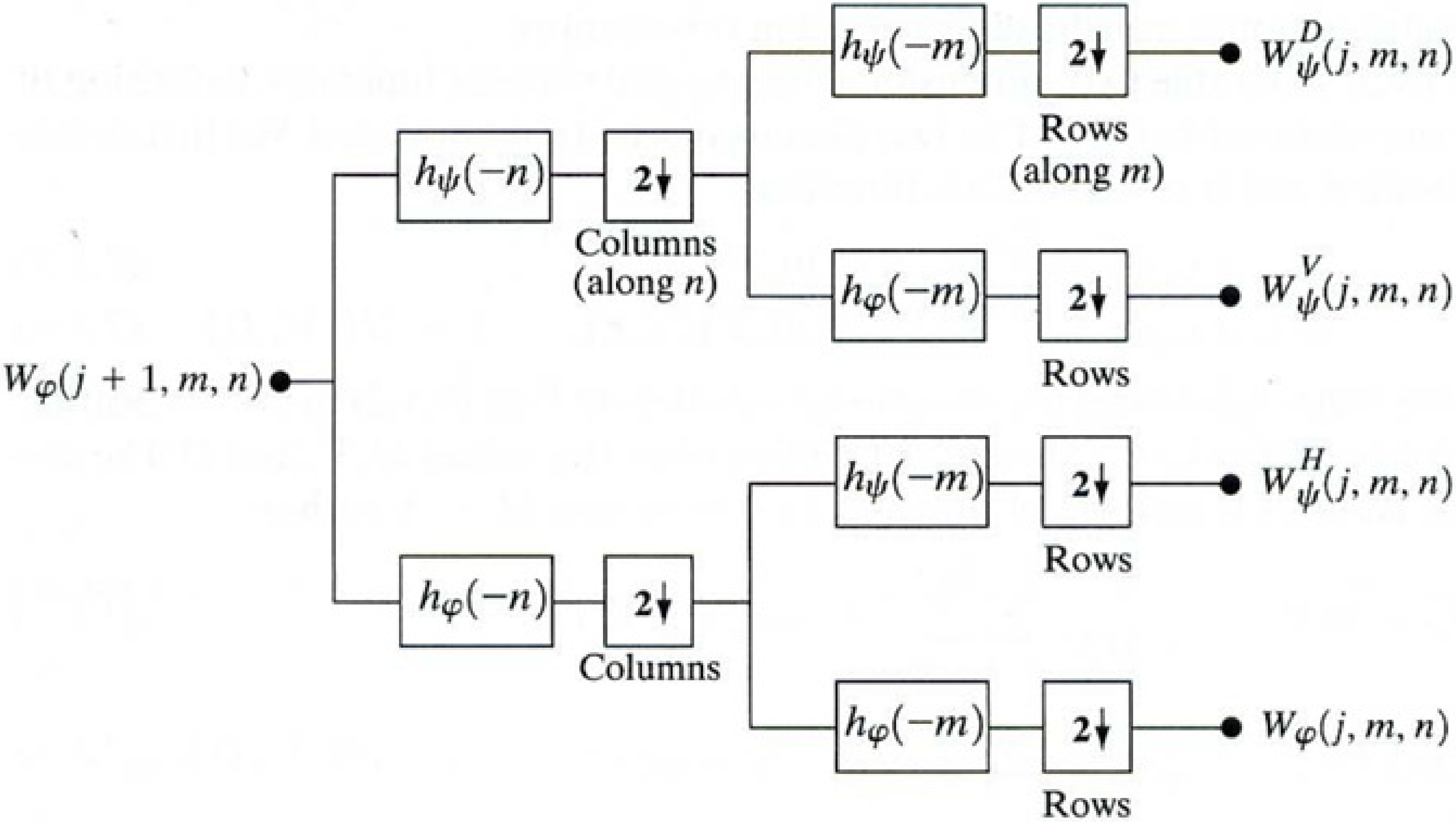
D.Donoho. **De-noising by soft thresholding**. IEEE Trans. on Information Theory, vol. 38(2), pp. 613--627, 1995.

Why Wavelets in Signal Processing

- Wavelets expansion is a meaningful signal representation "capturing significant part of the signal" at different resolution (denoising, compression, analysis)
- Wavelets decomposition is fast (convolution and resampling)
- Wavelets representation is not overcomplete

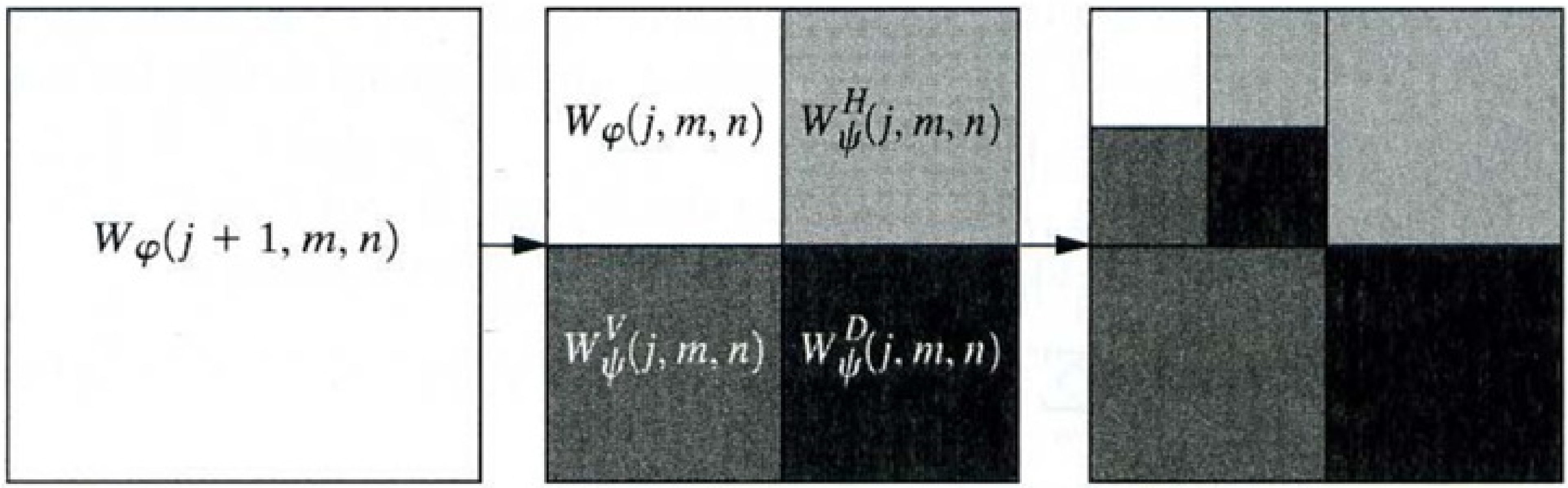
2D Fast Wavelet Transform

- The decomposition is performed separately along image rows and columns



2D Fast Wavelet Transform

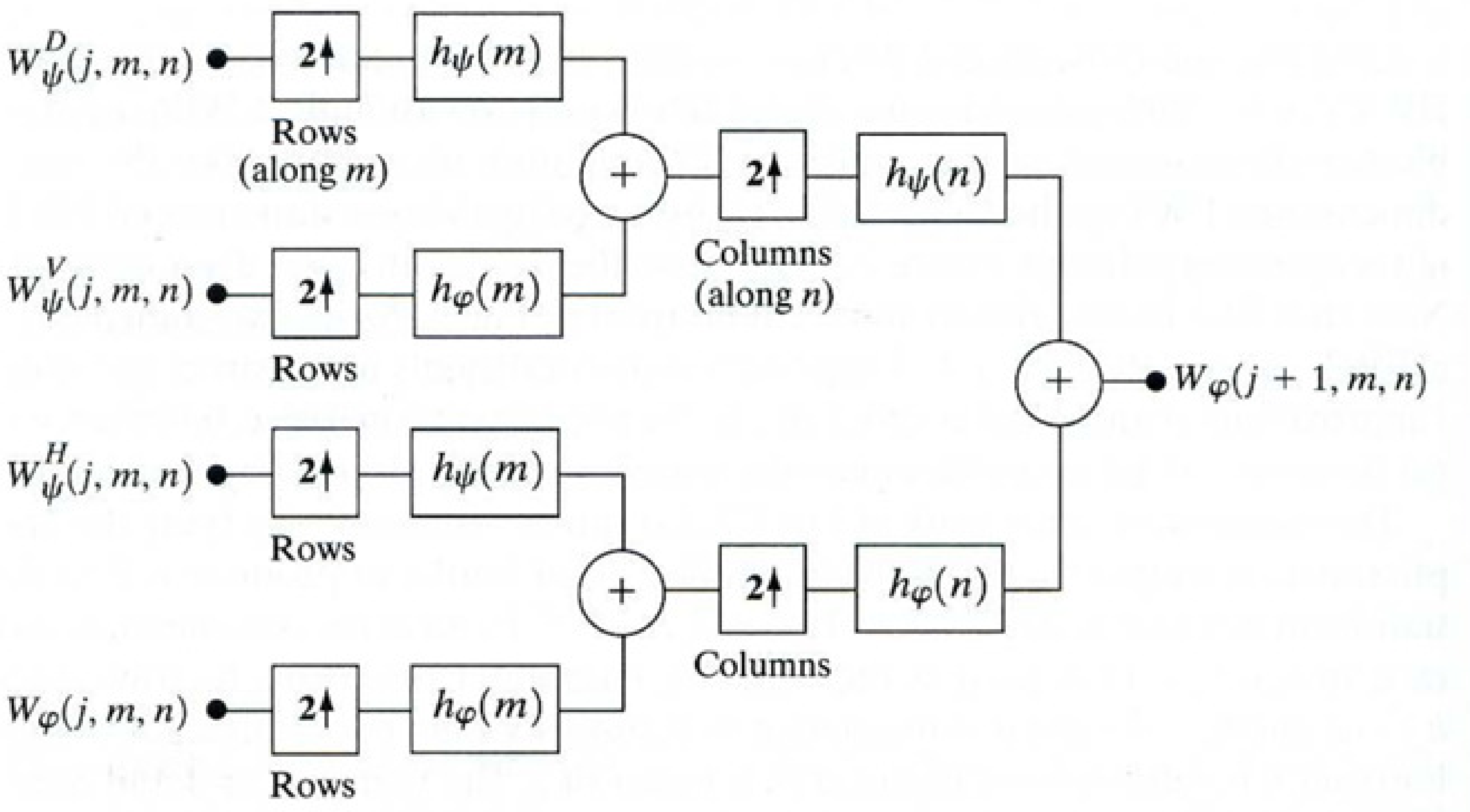
- The decomposition is performed separately along image rows and columns



- and possibly iterated

2D Fast Wavelet Transform

- The reconstruction is defined similarly



an Example