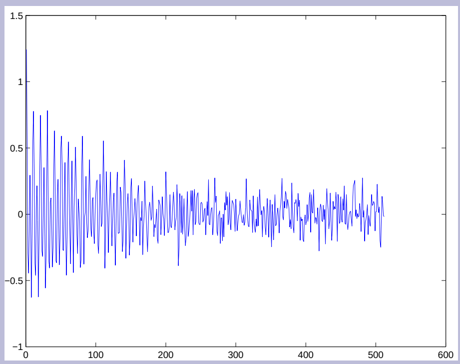
Advanced NMR Processing

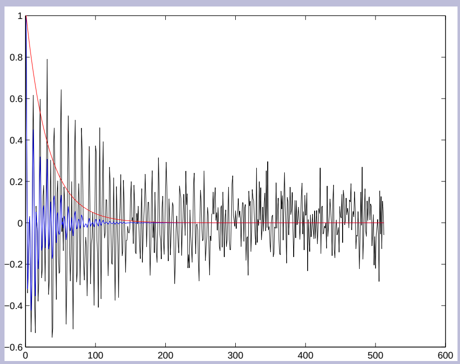
Ulrich Günther

EuroLabCourse "Advanced Computing in NMR Spectroscopy", Florence, Sept. 2001

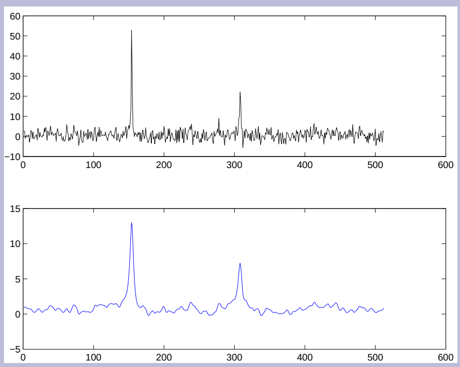
NMR Processing

Free induction decay (FID) 

↓

Apodization 

↓

FFT 

NMR Processing

- **Main processing steps**
  - Apodization
    - noise reduction
    - resolution enhancement
  - On-resonance water suppression
  - Zero-filling / linear prediction
  - Fourier transform

- **Post-processing**
  - Baseline correction
  - Denoising

FT - Convolution properties

A convolution is defined as

$$f_1(t) \star f_2(t) = \int_{-\infty}^{+\infty} f_1(\tau) f_2(t-\tau) d\tau$$

$$\mathcal{F}\{f_1(t) \star f_2(t)\} = \mathcal{F}\{f_1(t)\} \cdot \mathcal{F}\{f_2(t)\}$$

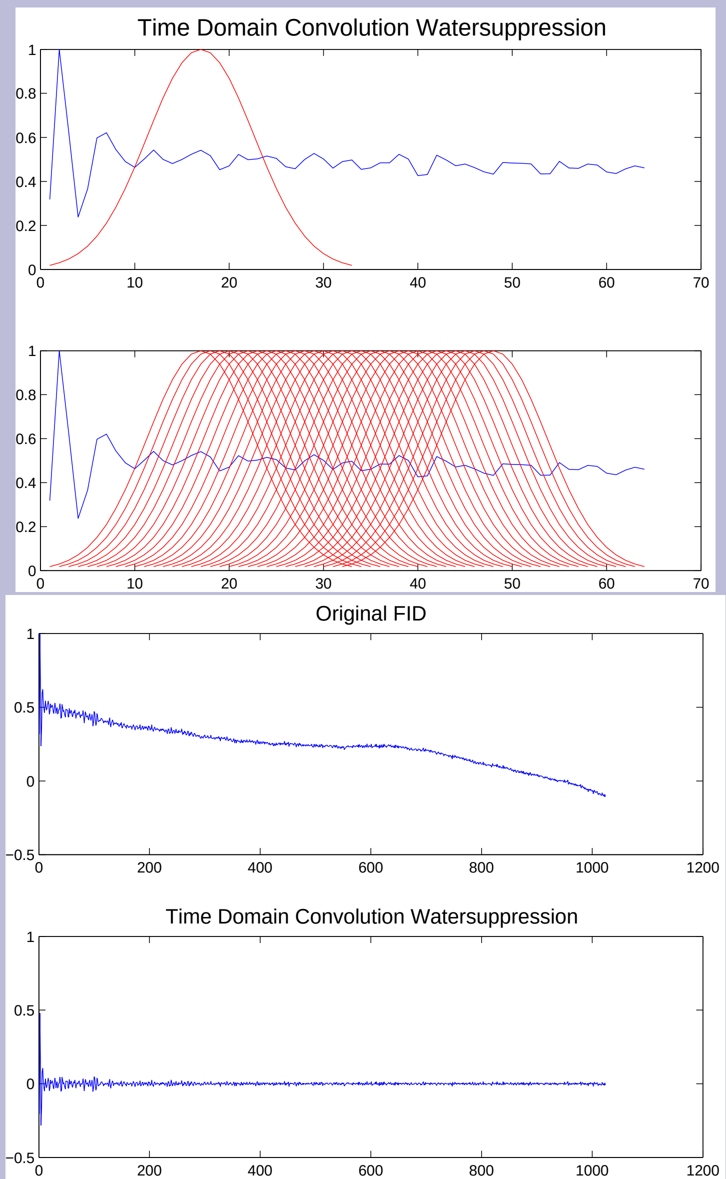
For discrete data this is equivalent to

$$DFT(\mathbf{da})_n = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} f_j A_{n-j}$$

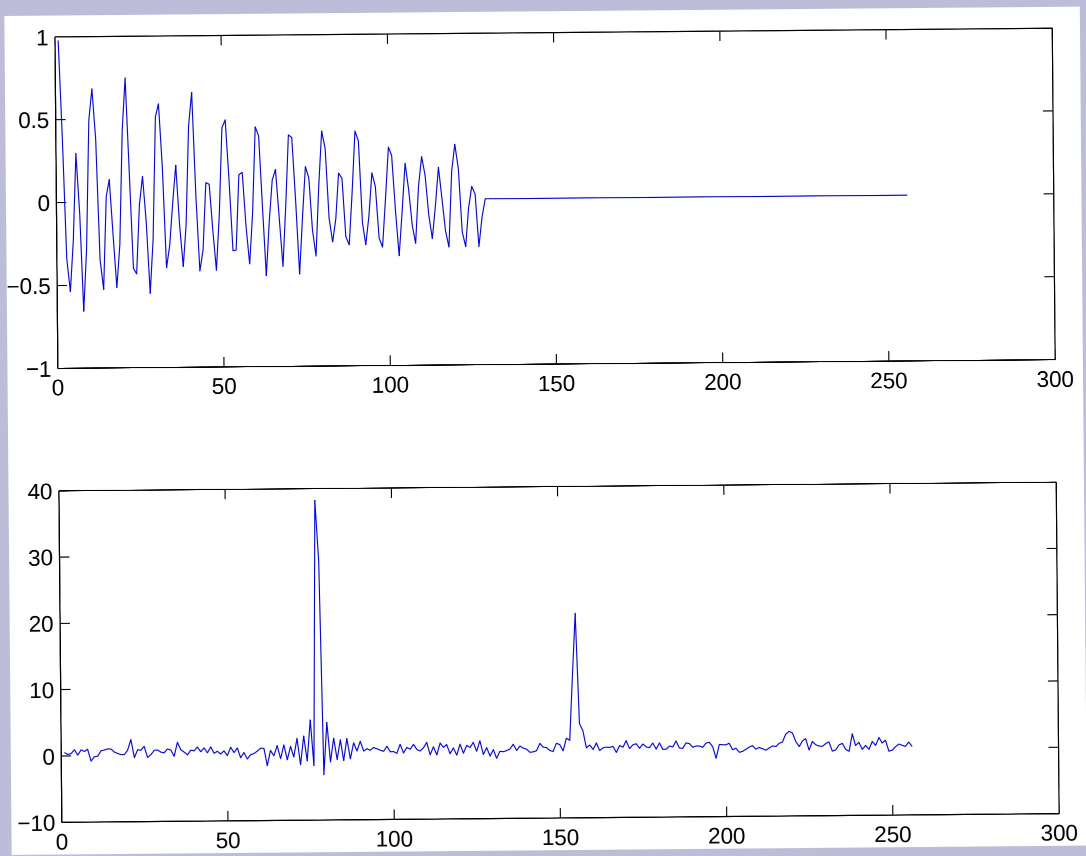
Special case: Convolution of a function with itself

$$\mathcal{F}\{f_1(t) \star f_1(t)\} = \mathcal{F}\{f_1(t)\} \cdot \mathcal{F}\{f_1(t)\}$$

Water Suppression by Time-Domain Convolution

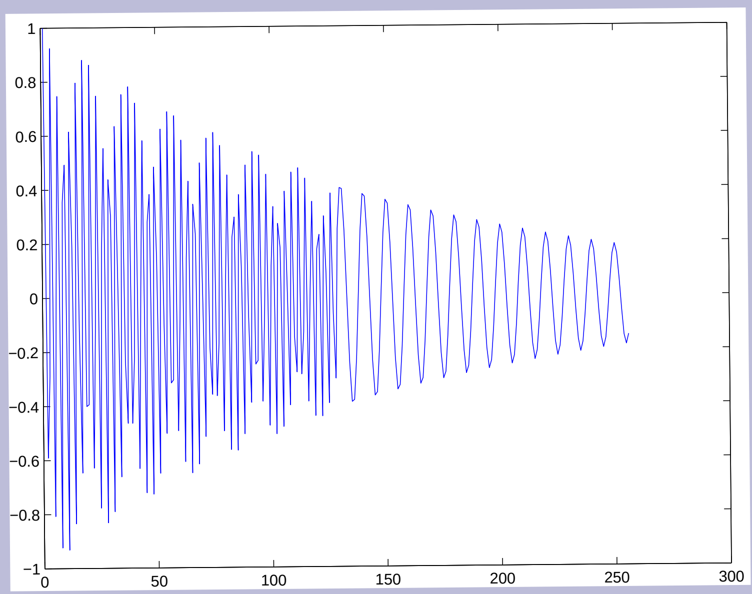


Zero-filling as extrapolation



Zerofilling may cause truncation artefacts

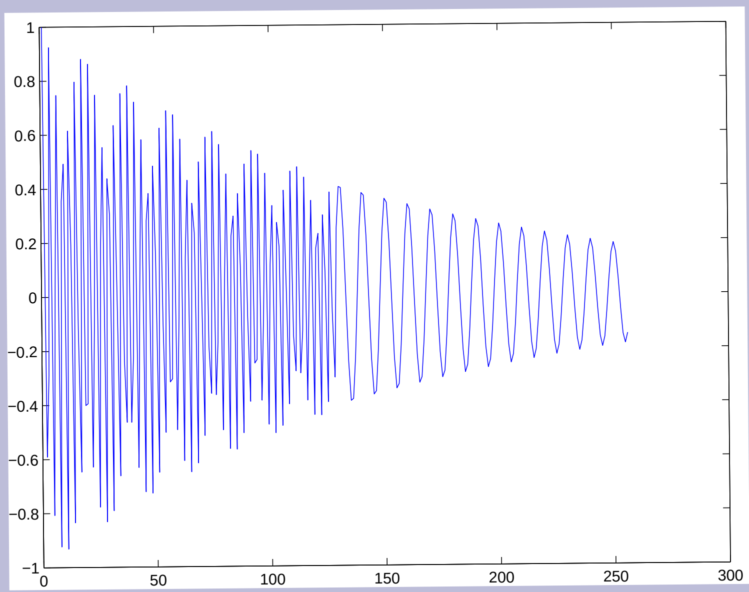
Fourier transform properties



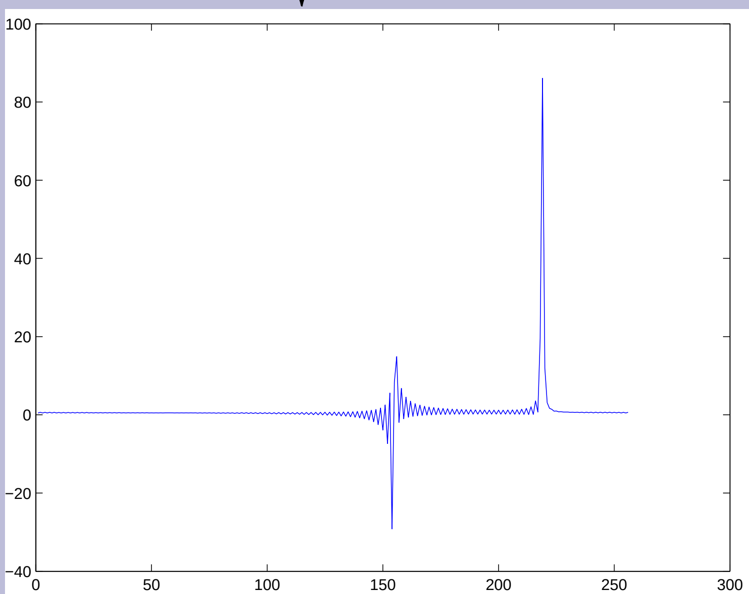
$\downarrow$ FFT

?

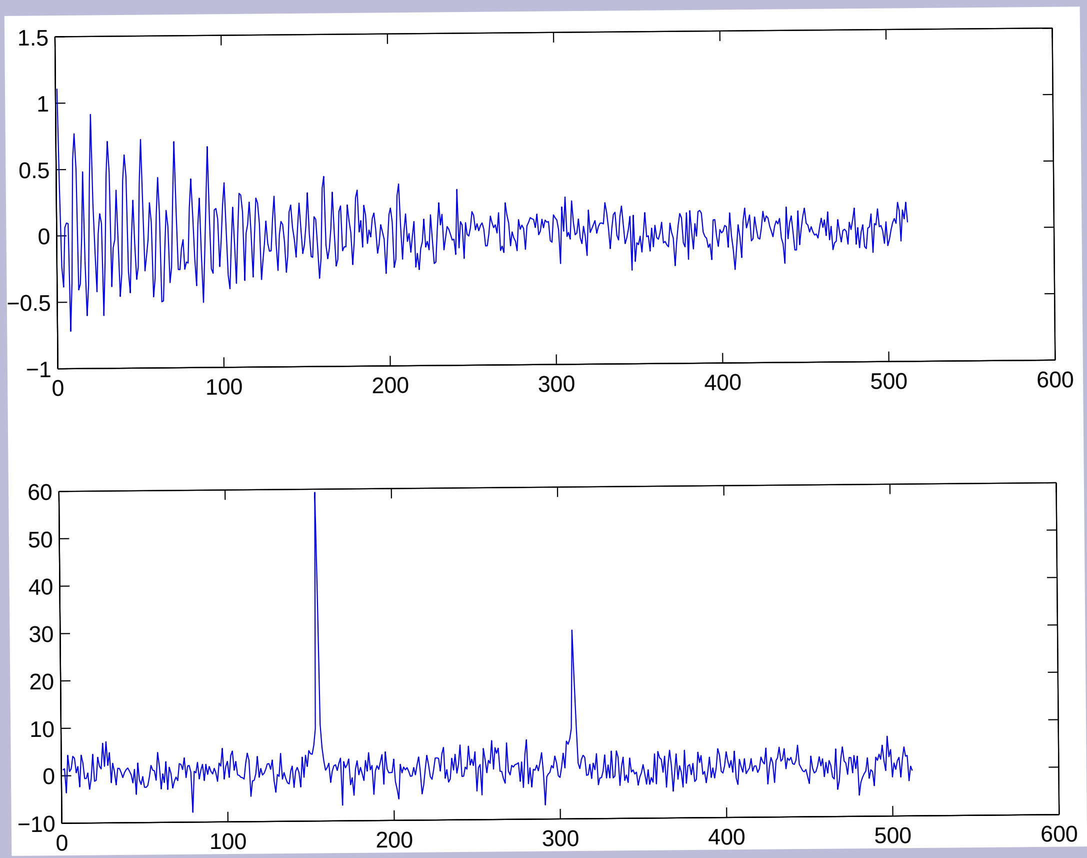
Fourier transform properties



↓ FFT



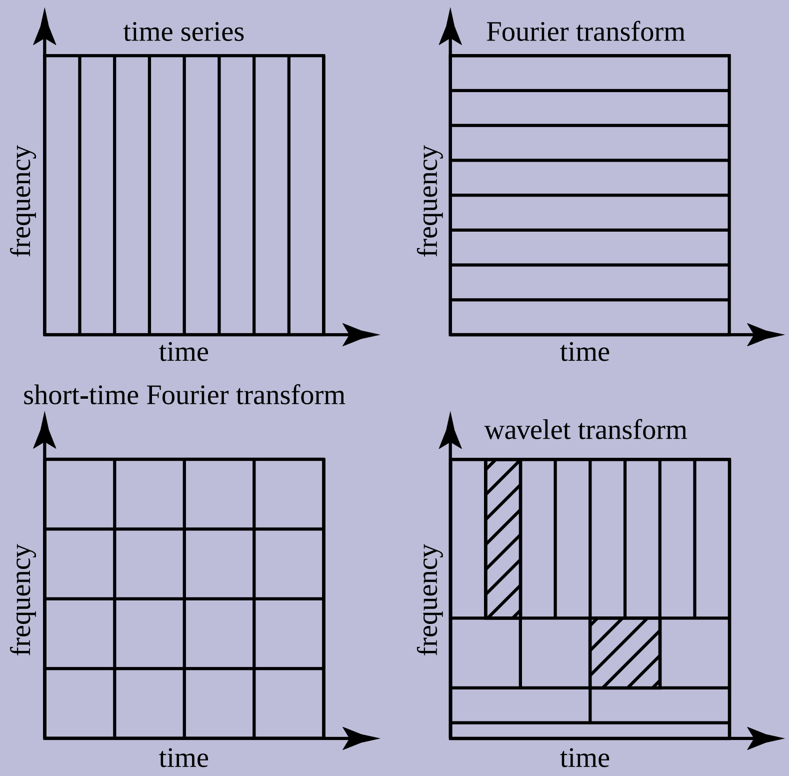
FFT and noisy data



Wavelet Transforms

- *Ondelettes* (wavelets) - Yves Meyer, 1980s
- *Wavelet* $\leftarrow$ "waving" above and below the x-axis $\Leftrightarrow$ integrate to zero.
- Wavelets **chop up data into frequency components**, and **analyze each frequency component with a resolution matched to its scale**.
- Applications:
  - data approximation (smoothing)
  - noise reduction
  - data compression (jpeg, mpeg, mp3)
  - time-frequency analysis
  - image analysis
- Fast dyadic wavelet transform (DWT) - N operations, compare: FFT $N \log(N)$.

Wavelet Transforms



- **FT**: high resolution in the time domain, each point contains information about all frequencies.

- **WFT**: divides the time-frequency plane in rectangular boxes. The resolution in time is increased at the expense of the frequency resolution.

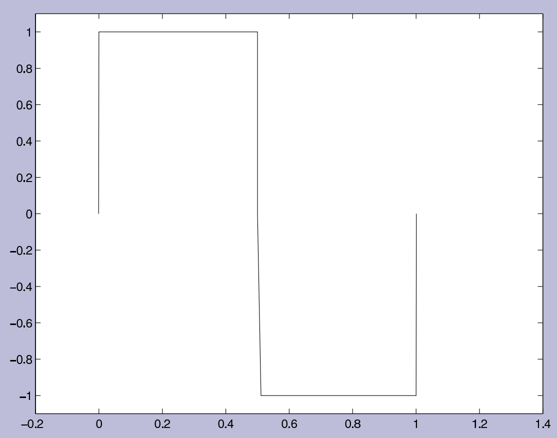
- **DWT**: scaling of the basis functions relative to their support.

The Haar System

A. Haar in 1909, appendix to a thesis.

The Haar function is a simple step function

$$\psi(x) = \begin{cases} 1, & x \in [0, \frac{1}{2}[\\ -1, & x \in [\frac{1}{2}, 1[\end{cases}$$

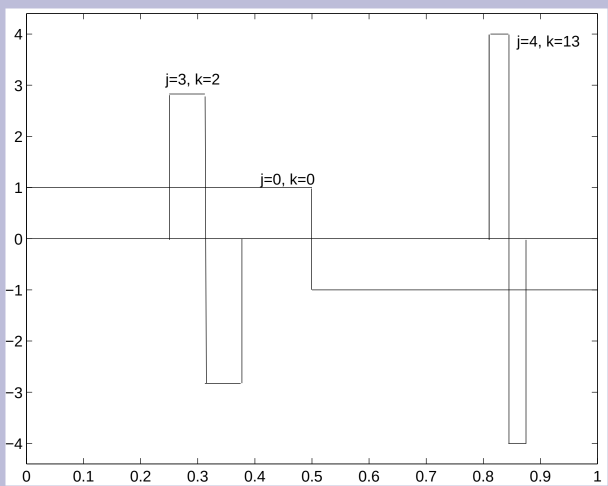


Dilatations and Translations

The Haar function ψ is used to define a *mother wavelet*. *Dilatations* and *translations* $\leftharpoondown$ other wavelets.

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k). \tag{1}$$

Integer translation indices k and dilatation indices j.



Properties of the Haar wavelet

Haar wavelets $\psi_{j,k}$ have a compact support

$$\operatorname{supp}\left(\psi_{j,k}\right) = \left[k2^{-j}, (k+1)\,2^{-j}\right[\tag{2}$$

i.e. they are zero outside this interval.

For each Haar wavelet $\psi_{j,k}$ the integral

$$\int_{-\infty}^{\infty} \psi_{j,k}(x)dx = 0.$$

I.e. the area above the x-axis is equal to the area below the x-axis.

Orthonormal System

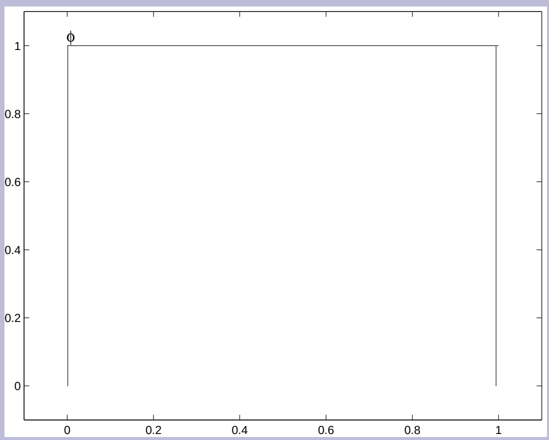
$\{\psi_{j,k}, j,k \in Z\}$ constitutes a complete orthonormal basis in L_2

$\Longleftrightarrow$ **Any square integrable function can be approximated arbitrarily well by a linear combination of these basis functions.**

$L_2(\mathbb{R})$ = space of complex valued functions with a finite L_2-norm:

$$||F||_2 = (\int_{\infty}^{\infty} |f(x)|^2 dx)$$

Father Wavelet



- For the Haar wavelet the father wavelet is a simple step function:

$$\phi = \begin{cases} 1, & x \in [0,1[\\ 0, & x \notin [0,1[\end{cases} .$$

- Required for fast DWT and wavelet construction (father wavelet $\rightarrow$ mother wavelet).

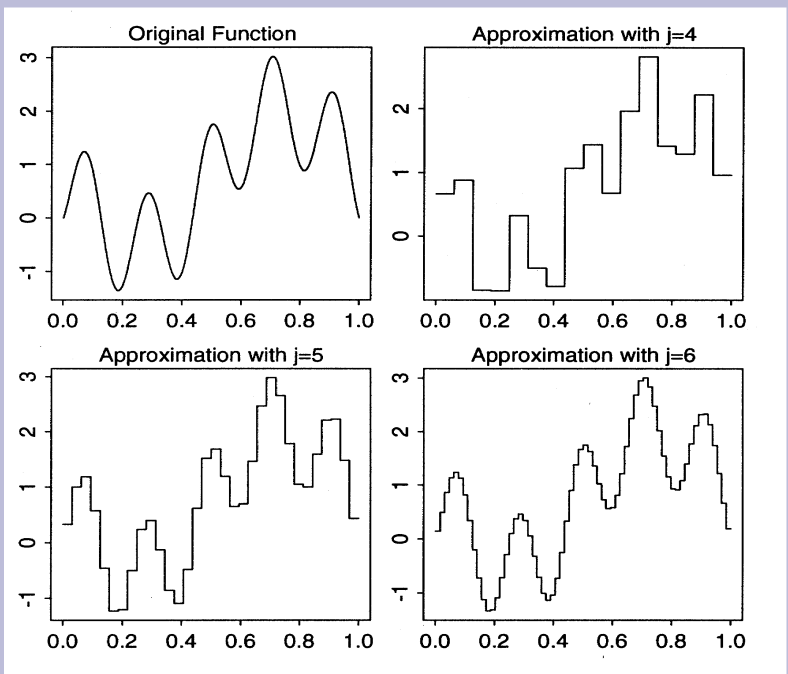
- With the scaling function ϕ we can expand our original set to $\{\phi_{j_0,k}, \psi_{j,k},\ j \geq j_0, k \in \mathbb{Z}\}$. The combined set is again an orthonormal basis in L_2.

Signal Approximation in the Haar System

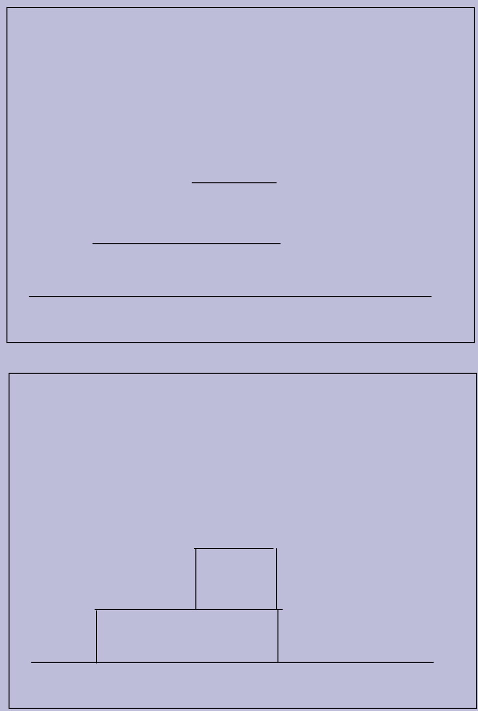
A signal can be represented as

$$f(x) = c_{00}\phi(x) + \sum_{j=0}^{n-1}\sum_{k=0}^{2^j-1} c_{j,k}\psi_{j,k}(x) \qquad (3)$$

$c_{j,k}$ are wavelet coefficients,
$\psi_{j,k}$ are wavelets derived from a mother wavelet ψ
ϕ is a scaling function (father wavelet)



Step Function $\rightarrow$ Wavelet Transform



Approximations of a function with increasing resolution by a step function.

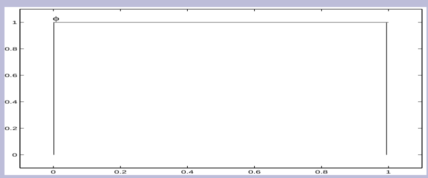
- Any L_2-function can be approximated by a simple step function.
- The approximation converges for infinitely small steps.
- Step functions do not constitute an orthonormal basis for L_2.

Multiresolution Analysis (MRA)

- G. Mallat
- general framework to construct wavelet bases
- start from a father wavelet
- derive an *orthonormal* mother wavelet
- and wavelet subspaces suitable to approximate function with increasing resolution

MRA 1

Start with the scaling function (father wavelet)



$$\phi_{0,k} = \phi(x-k)$$

Set $\{\phi_{o,k}\}$ = orthonormal basis for a reference space V_0.

The functions in V_0 have the form

$$f(x) = \sum_k c_k \phi(x-k)$$

Define a function space V_0, functions in $[k, k+1[$

$$V_0 = \{f(x) = \sum_k c_k \phi(x-k)\}.$$

MRA 2

Starting from V_0 we define linear spaces

$$V_1 = \{h(x) = f(2x) \; : \; f \in V_0\}$$

$$\vdots$$

$$V_j = \{h(x) = f(2^j x) \; : \; f \in V_0\}$$

V_1: all functions constant on $[\frac{k}{2}, \frac{k+1}{2}[$.

Set $\{\phi_{1,k}\}$ is an orthonormal basis in V_1

with

$$\phi_{1,k}(x) = \sqrt{2}\phi(2x - k)$$

.

MRA 3

$V_0 \longrightarrow V_1 \longrightarrow V_j$

Basis functions of V_j are $\phi_{j,k} = 2^{j/2}\phi(2^j x - k)$.
ϕ generates a sequence of **spaces** $\{V_j,\ j \in \mathbb{Z}\}$ **which are nested**:

$$V_0 \subset V_1 \subset \ldots \subset V_j \subset \ldots$$

$$V_j \subset V_{j+1},\ j \in \mathbb{Z}.$$

If in addition every square integrable function can be approximated by functions in

$$\bigcup_{j \geq 0} V_j$$

than $\{V_j,\ j \in \mathbb{Z}\}$ is a MRA!

MRA 4

$\{V_j,\ j \in \mathbb{Z}\}$ is not a basis in L_2!

To obtain a basis we must **orthogonalize** it.

$\Longrightarrow$ find a W_0 for which

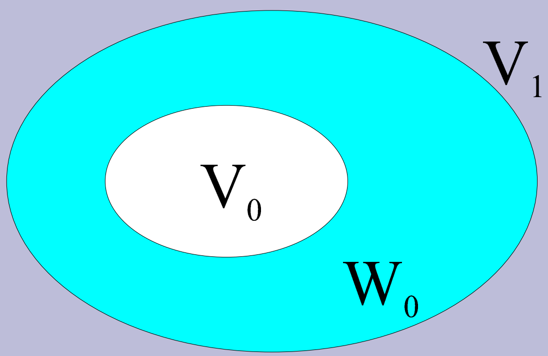
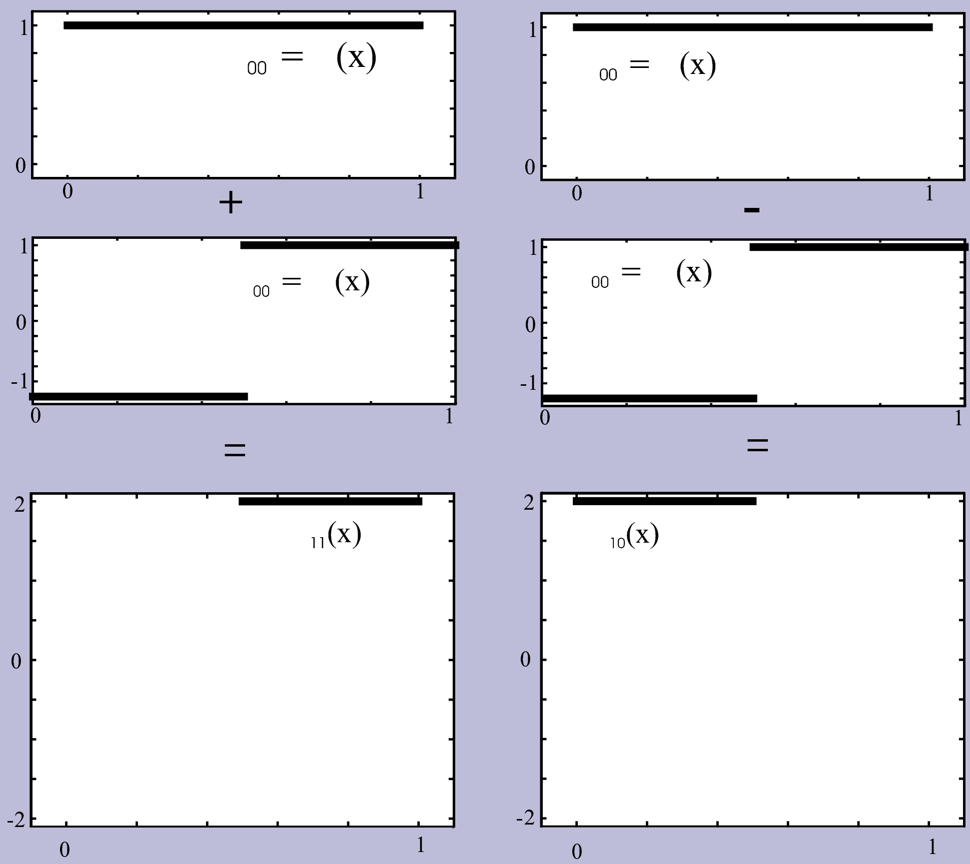


Figure 1: $V_1 = V_0 \bigoplus W_0$.

$$V_1 = V_0 \bigoplus W_0.$$

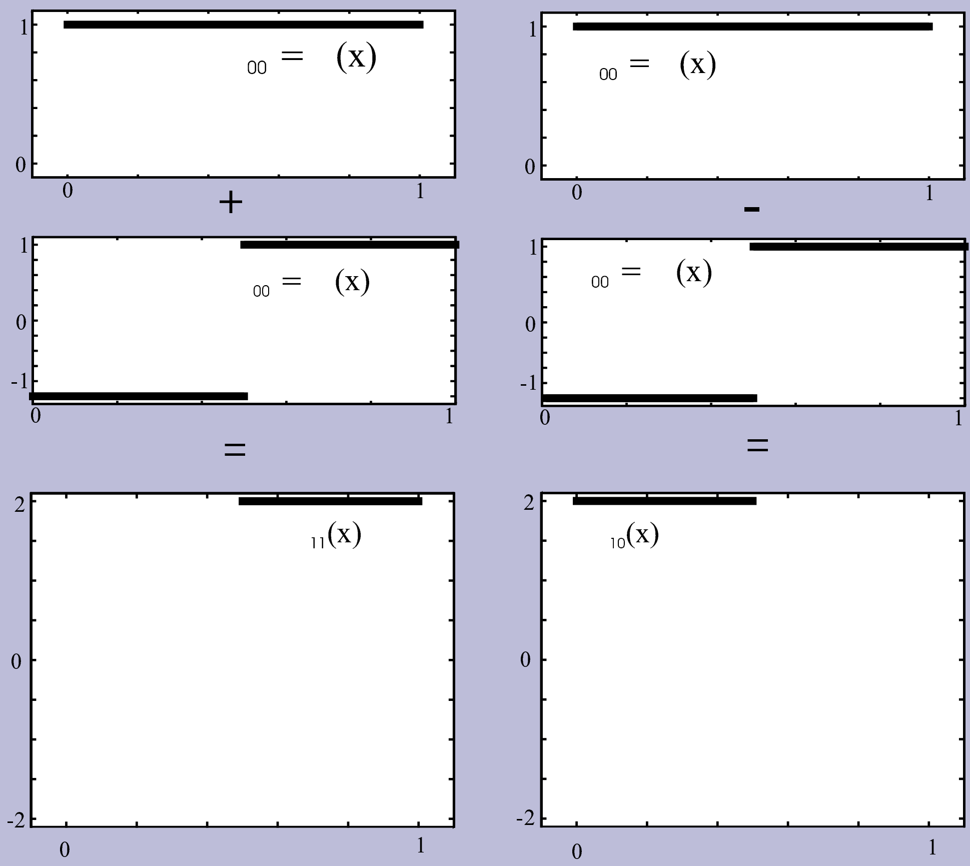
W_0 is the orthogonal complement of V_0 in V_1.

MRA 6



ϕ_{1k} must be a linear combination of ϕ_{0k} and ψ_{0k}:

$$\phi_{10}(x) = \sqrt{2}\phi(2x) \quad = \sqrt{2}\{\phi_{00}(x) - \psi_{00}(x)\}/2$$
$$= \frac{1}{\sqrt{2}}\{\phi_{00}(x) - \psi_{00}(x)\}$$



$$\phi_{11}(x) = \sqrt{2}\phi(2x-1) \quad = \frac{1}{\sqrt{2}}\{\phi_{00}(x) + \psi_{00}(x)\}/2 \quad .$$

MRA 7

Repeat the same process for higher values of j.

$\longrightarrow$ consecutive summation of subspaces

$$\begin{aligned} V_{j+1} &= V_j \oplus W_j \\ &= V_j \oplus W_{j-1} \oplus W_j \\ &= V_0 \oplus W_0 \oplus W_1 \oplus \ldots \oplus W_j \\ &= V_0 \oplus \bigoplus_{l=0}^{j} W_l \end{aligned}$$

Finally this sum of nested spaces spans :

$$L_2(\mathbb{R}) = V_0 \oplus \bigoplus_{l=0}^{j} W_l.$$

$$L_2(\mathbb{R}) = V_0 \oplus \bigoplus_{l=0}^{j} W_l.$$

$\Longrightarrow$ every square integrable function can be represented as the series

$$f(x) = \sum_k \alpha_{0,k}\phi_{0,k}(x) + \sum_{j=0}^{\infty} \sum_k \beta_{j,k}\psi_{j,k}(x).$$

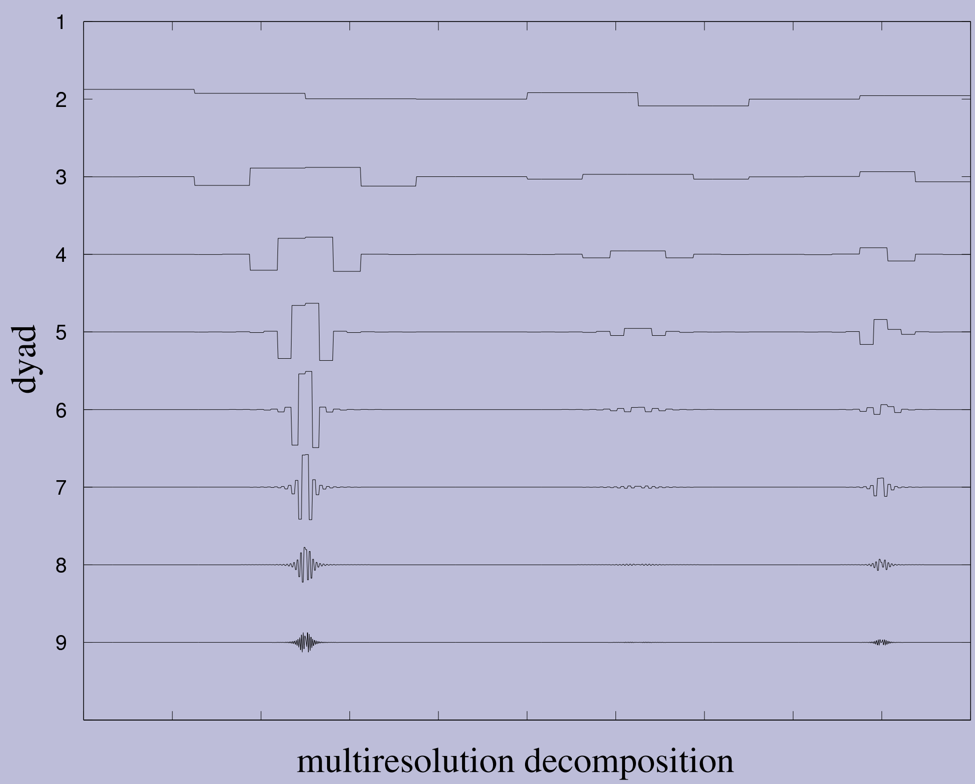
- The father wavelet ϕ creates a MRA

- $\alpha_{0,k}$ and $\beta_{0,k}$ = coefficients for mother and father wavelets

- f provides a location in time and frequency
  - The location in time is determined by k
  - the location in frequency by j (the larger j, the higher the frequency related to $\psi_{j,k}$).

$$f(x) = \alpha_{00}\phi(x) + \sum_{j=0}^{n-1}\sum_{k=0}^{2^j-1} c_{j,k}\psi_{j,k}(x)\,, \tag{4}$$

Summation for $f(x)$ stops at 2^{j-1} for finite data sets $N = 2^{j-1}$ =number of data points.



Thresholding

Noise reduction = Search for the largest "true" wavelet coefficients.

Hard thresholding function: "keep or kill" selection:

$$\delta_{\mathrm{hard}}(x) = \begin{cases} x, & \text{if } |x| > \lambda \\ 0, & \text{if } |x| < \lambda \end{cases}$$

Soft thresholding:

$$\delta_{\mathrm{soft}}(x) = \begin{cases} x - \lambda, & \text{if } x > \lambda \\ 0, & \text{if } |x| \leq \lambda \\ x + \lambda, & \text{if } x < -\lambda \end{cases}$$

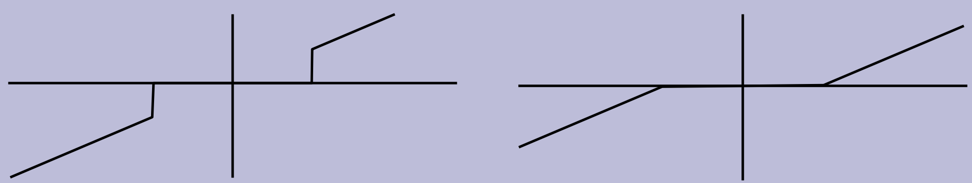


Figure 2: Hard and soft thresholding.

Universal threshold selection for white noise:

$$\lambda = \sigma \sqrt{2 \log n} / \sqrt{n}$$

Denoising Procedure

Signal (FID)

$\downarrow$ DWT

Wavelet coeffs.

$\downarrow$ Thresholding

$\downarrow$ IDWT

Denoised signal

Smoother Wavelet Bases

Must form an orthonormal basis for $L_2(\mathbb{R})$.

Minimize the wavelet coefficients for smooth functions.

Number constraints during the design of wavelets determines number of vanishing moments.

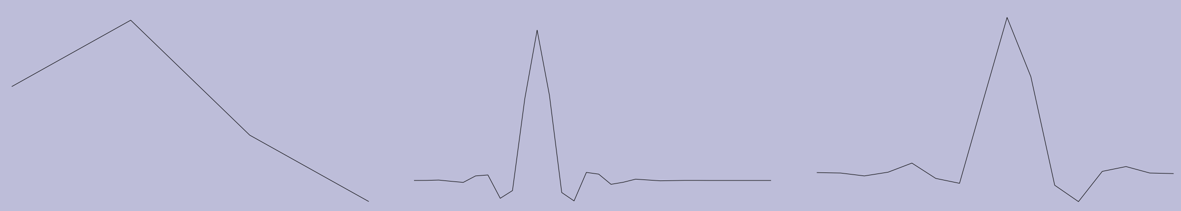


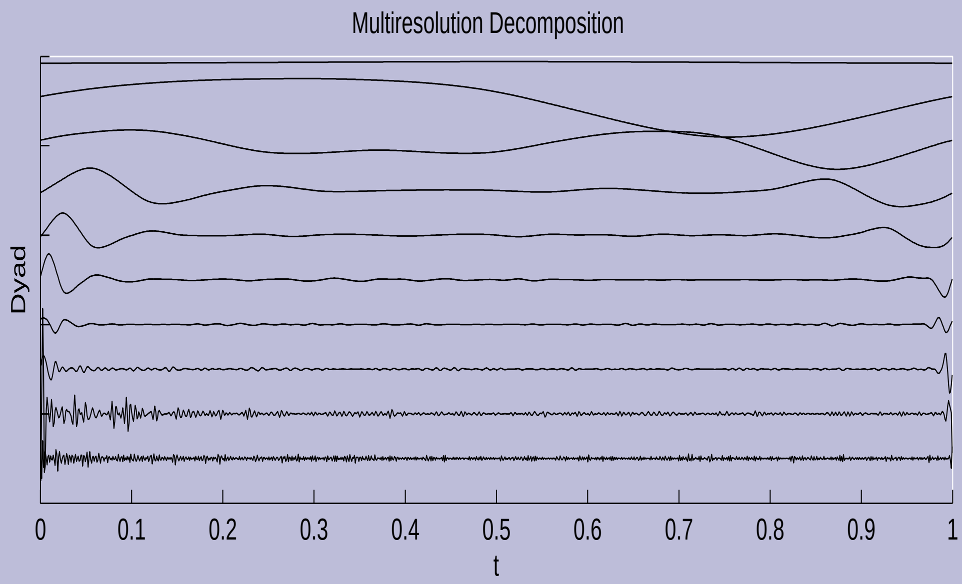
Figure 3: Types of wavelets from left to right: Daubechies (4), Coiflet (5) and Symmlet (8).

- Daubechies wavelets: vanishing moments for mother but not for father wavelets, asymmetric.

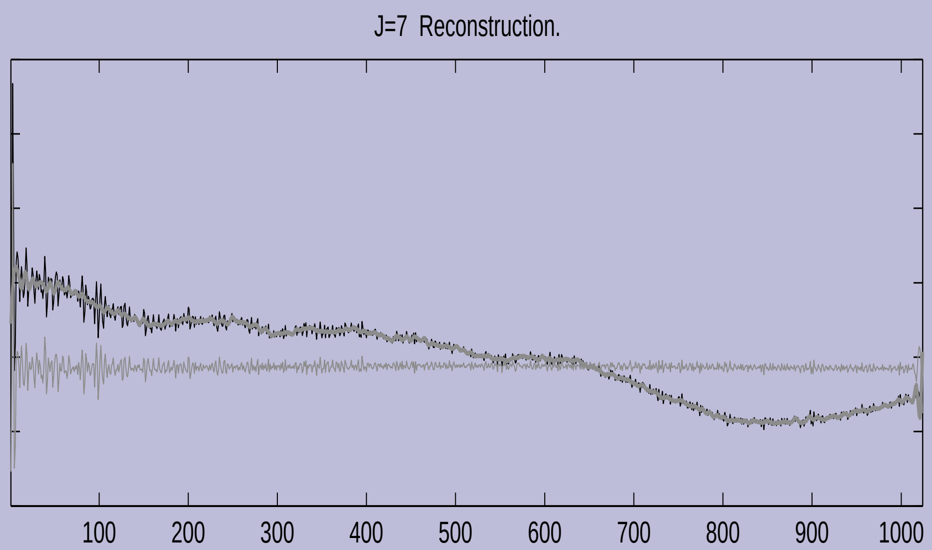
- Coiflets: additional vanishing moments for father wavelets.

- Symmlets are as close to symmetry as possible.

WAVEWAT

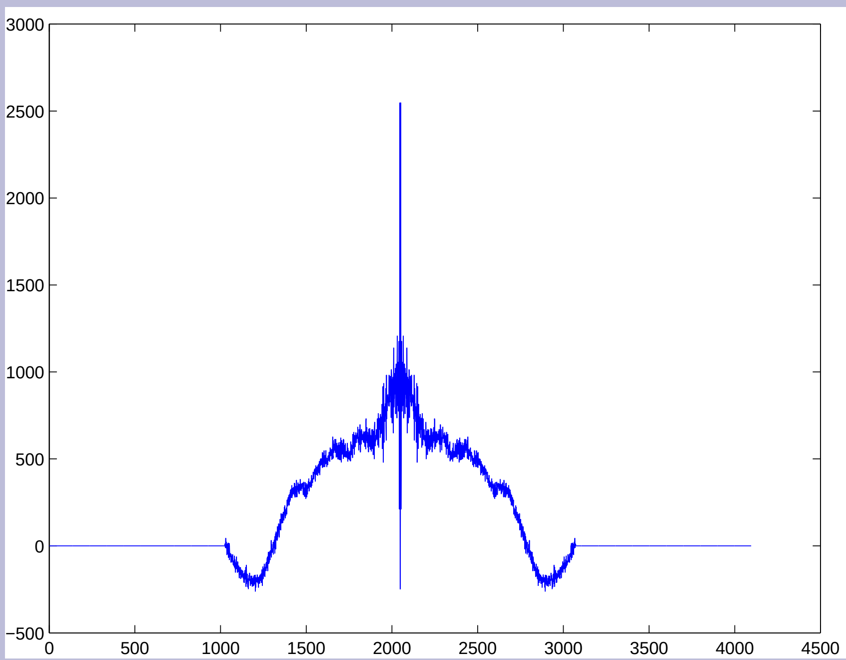


Multiresolution Decomposition
Dyad
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
t



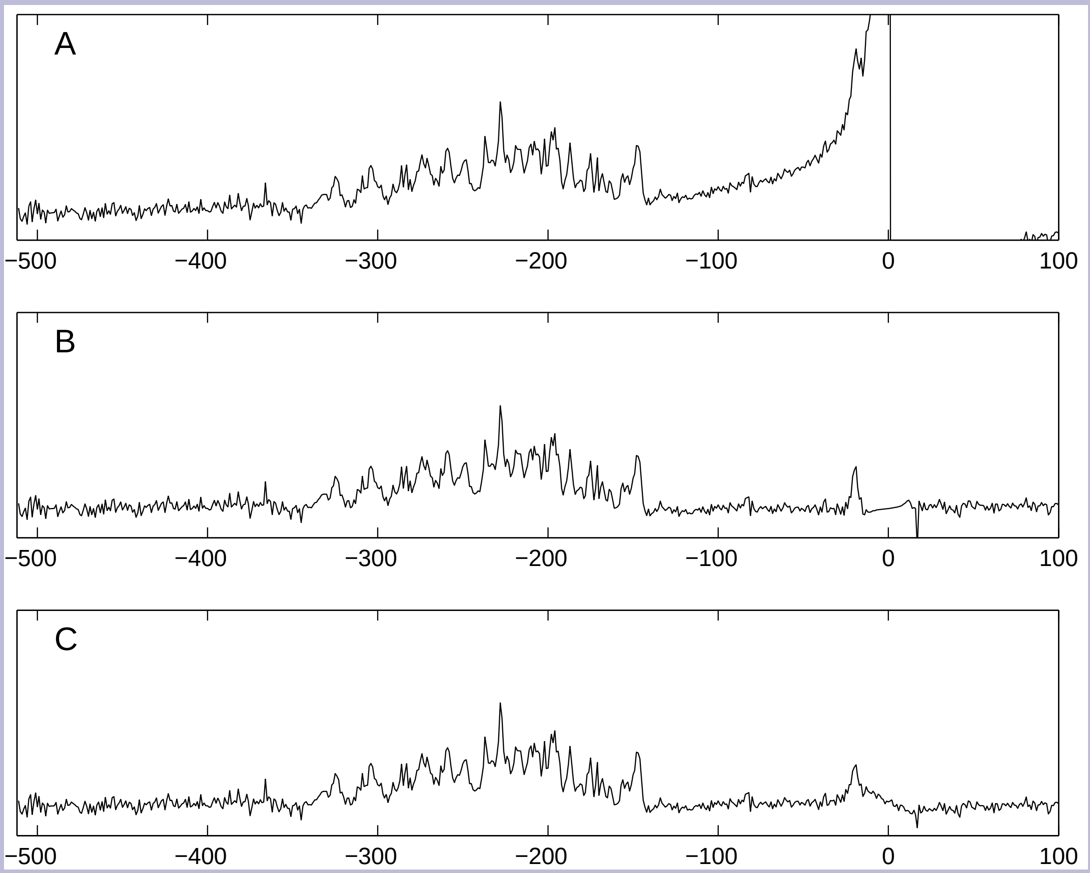
J=7 Reconstruction.
100
200
300
400
500
600
700
800
900
1000

Mirror reflections of FID to avoid edge effects

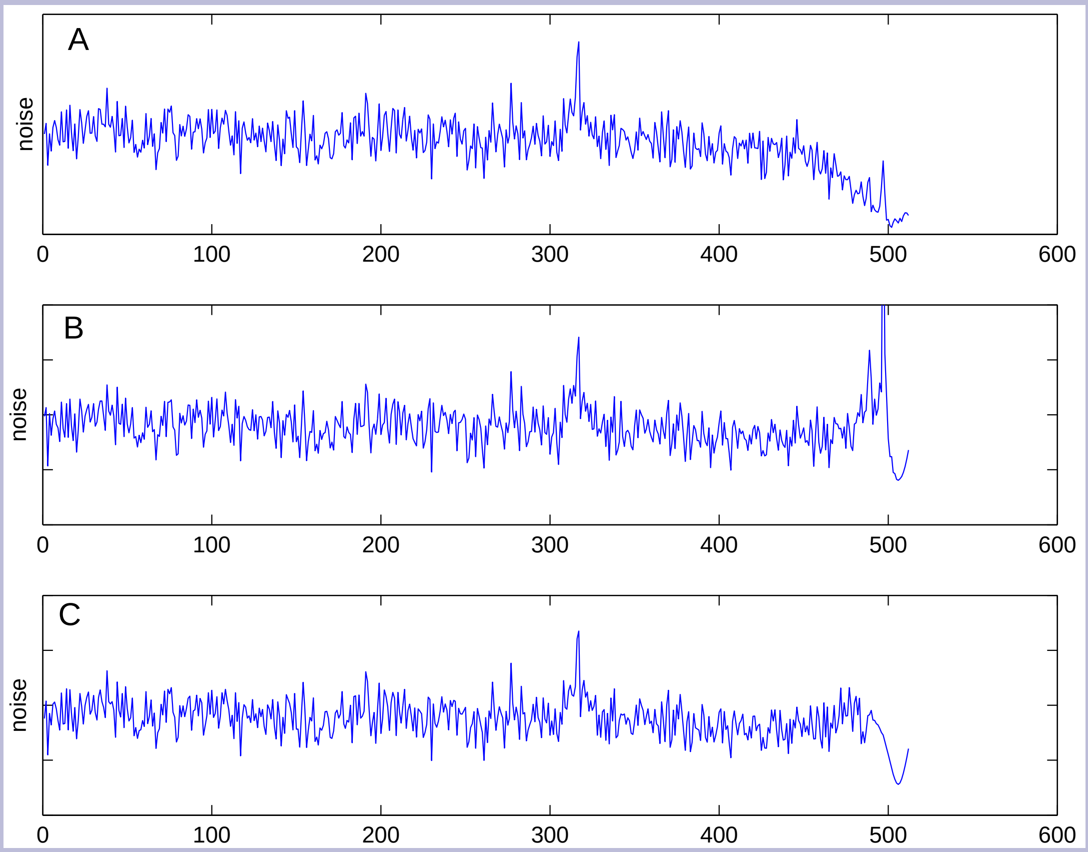


Use ZF to increase number of dyadic levels!

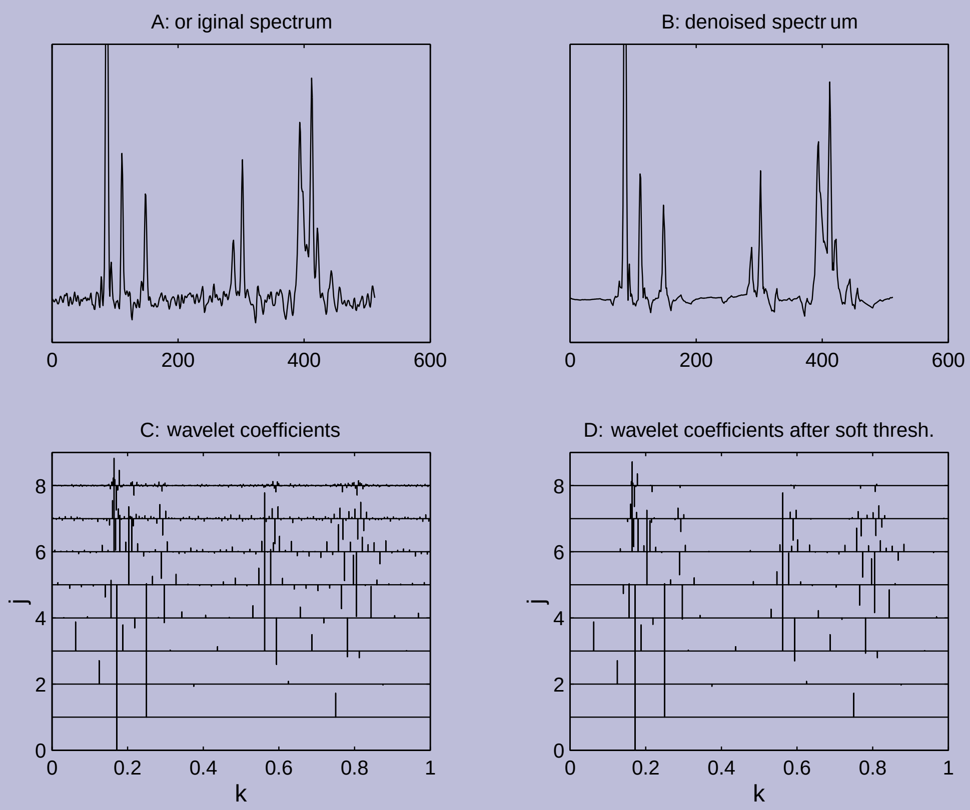
WAVEWAT vs. TD Convolution Difference



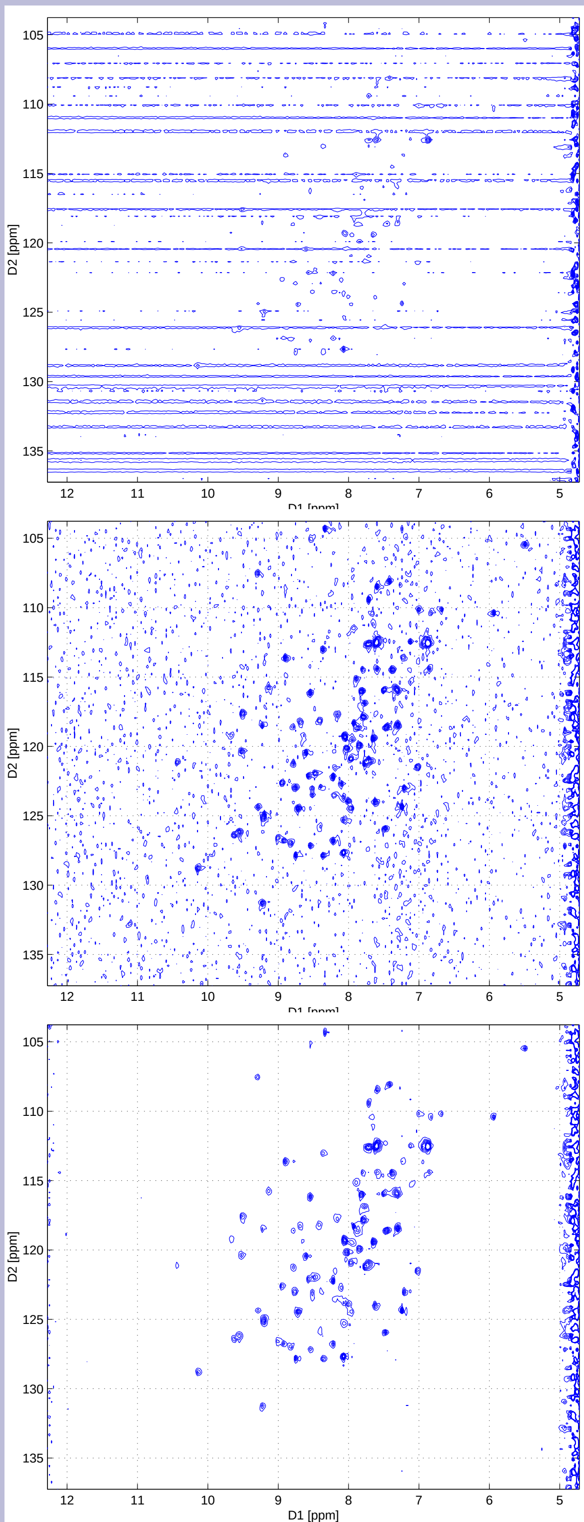
Noise Levels

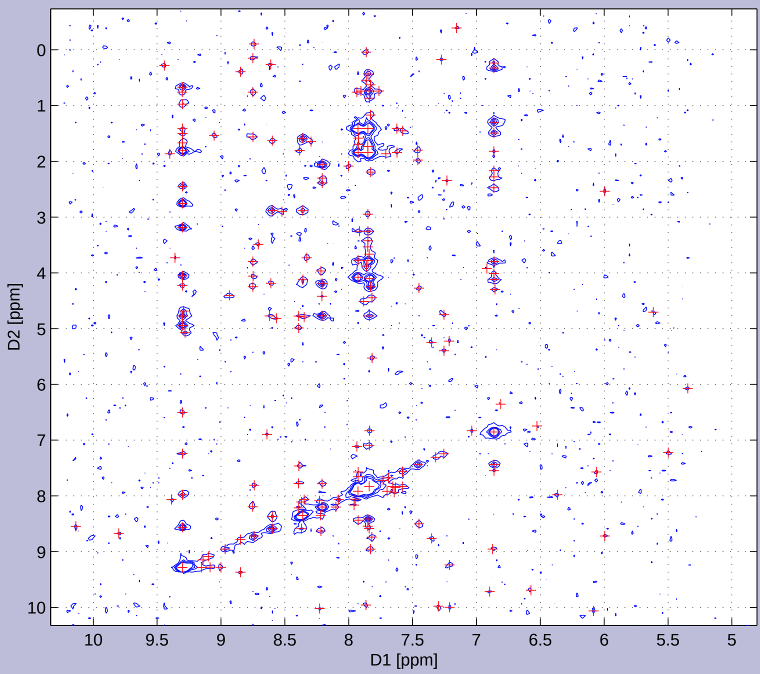


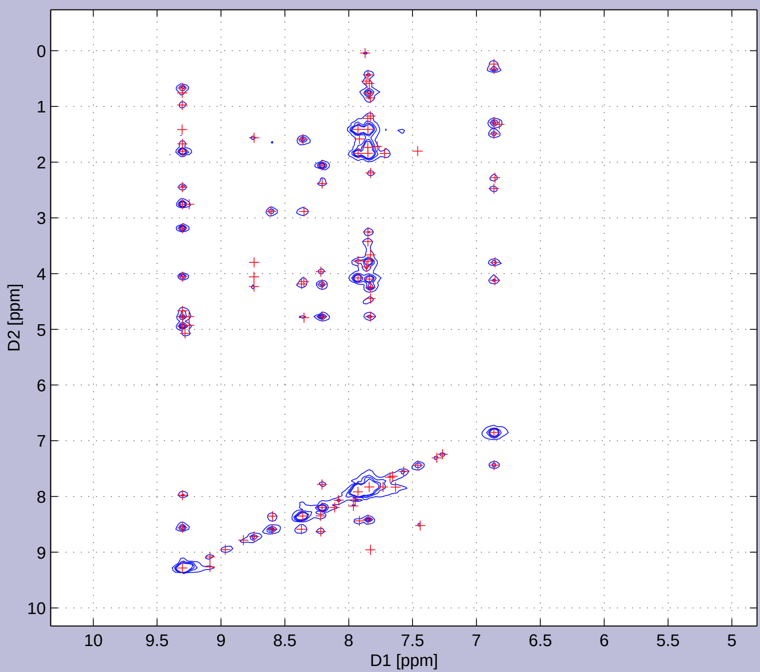
Noise Suppression



Wavelet shrinkage using a Daubechies (4) wavelet and soft thresholding.


D2 [ppm]
105
110
115
120
125
130
135
12
11
10
9
8
7
6
5
D1 [ppm]


D2 [ppm]
D1 [ppm]
0
1
2
3
4
5
6
7
8
9
10
10
9.5
9
8.5
8
7.5
7
6.5
6
5.5
5


D2 [ppm]
D1 [ppm]
0
1
2
3
4
5
6
7
8
9
10
10
9.5
9
8.5
8
7.5
7
6.5
6
5.5
5

SVD

$$\mathbf{H} = \mathbf{U_{LL}} \mathbf{S_{LM}} \mathbf{V}^{\dagger}_{\mathbf{MM}}.$$

$\mathbf{U}$ and $\mathbf{V}$ are $(L \times L)$ and $(M \times M)$ unitary matrices

$\mathbf{S}$ is $(L \times M)$ in size and contains the singular values.

Singular values$\Longleftrightarrow$ signal components in $\mathbf{H}$.

Large signal components $\Longleftrightarrow$ large singular values.

SVD noise suppression

FID:

$$f = (s_0, s_1, \ldots, s_{N-1})$$

derive Hankel type matrix:

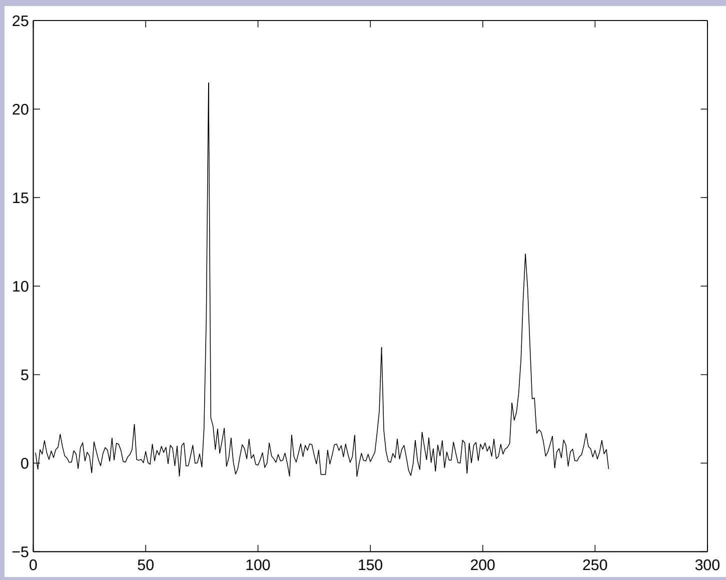
$$\mathbf{H} = \begin{pmatrix} s_0 & s_1 & s_2 & \cdots & s_{M-1} \\ s_1 & s_2 & s_3 & \cdots & s_M \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ s_{L-1} & s_L & s_{L+1} & \cdots & s_{N-1} \end{pmatrix}.$$

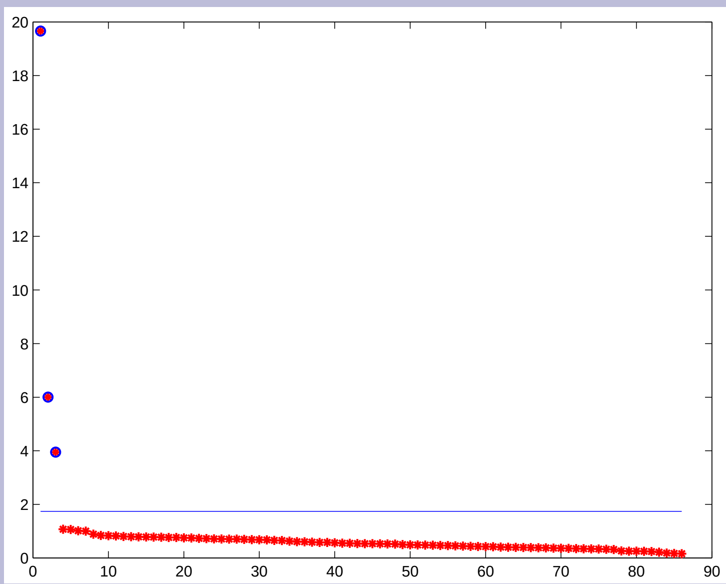
No noise: Rank of $\mathbf{H}$ is equivalent to the number of signals in the FID.

Noise: $\rightarrow$ full rank.

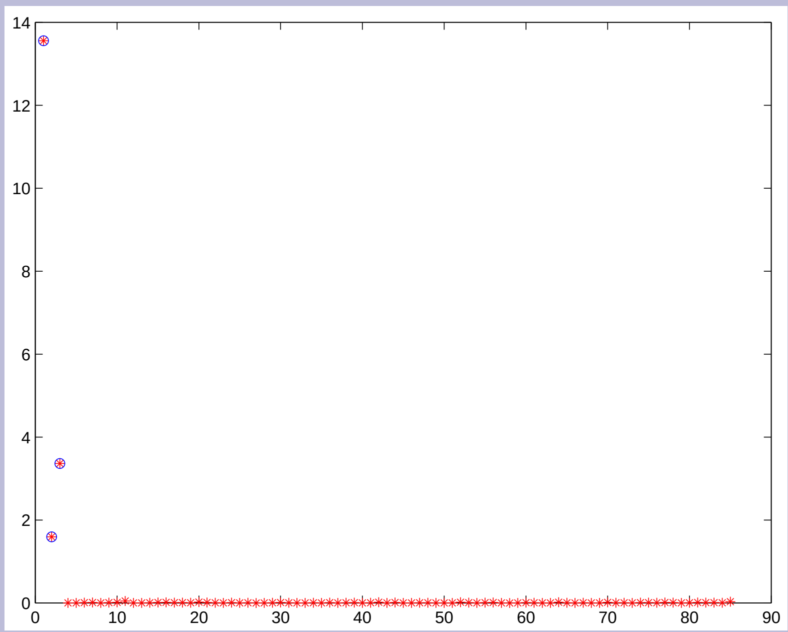
Periodic components in the FID will be represented singular values.

$O((L-M)^2M)$ process!

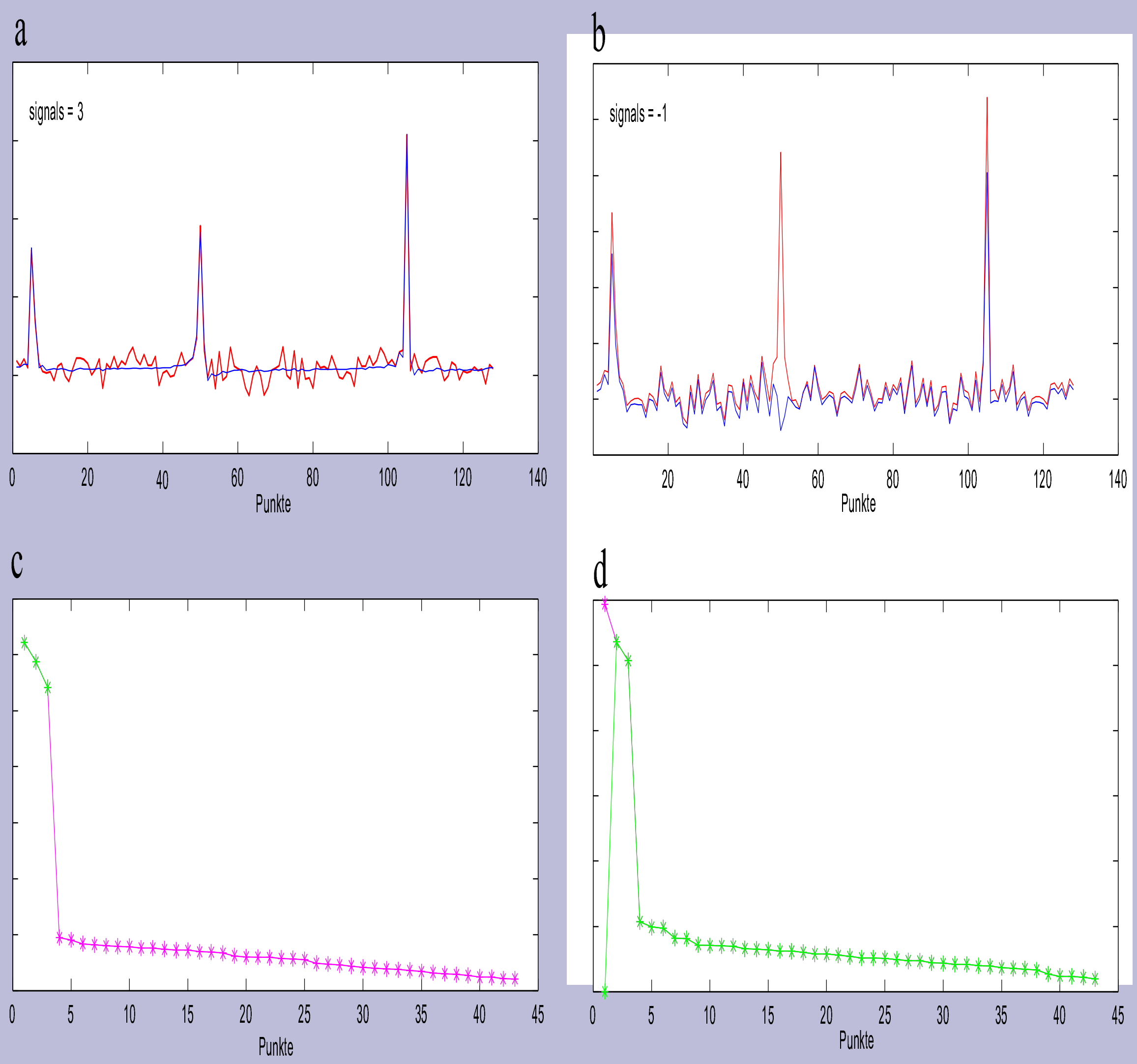

25
20
15
10
5
0
−5
0
50
100
150
200
250
300


20
18
16
14
12
10
8
6
4
2
0
0
10
20
30
40
50
60
70
80
90

Differential plot of the singular values



SVD Applications



Linear prediction

LP-SVD: by Kumaresan and Tufts modified by Porat and Friedlander.

Forward linear prediction: describe a data point by a linear combination of K preceding points:

$$x_n = \sum_{k=1}^{K} a_k x_{n-k}, \tag{5}$$

Backward linear prediction model by a linear combination of the K following points:

$$x_n = \sum_{k=1}^{K} b_k x_{n+k}. \tag{6}$$

$$x_n = \sum_{k=1}^{K} b_k x_{n+k}. \tag{7}$$

Equation 5 in matrix form:

$$x = a \cdot \mathbf{Y}$$

inversion of $\mathbf{Y} \rightarrow$ matrix coefficients a:

by SVD:

$$\mathbf{Y} = \mathbf{U}\Lambda\mathbf{V}^{\dagger}$$

$\rightarrow \mathbf{Y}'$ is obtained by

$$\mathbf{Y}^{\dagger} = \mathbf{V}\Lambda^{-1}\mathbf{U}^{\dagger}.$$

$$a = \mathbf{V}\Lambda^{-1}\mathbf{U}^{\dagger}x.$$

Prediction coefficients $a_k \to$ frequencies and damping by calculating the roots of the polynomial

$$z^K - a_1 z^{K-1} - \ldots - c_K = 0.$$

The z^K must be damped exponentials!

The signal x_n is described by

$$x_n = \sum_{k=1}^{K} A_k z_k^{n-1}. \quad (8)$$

A_k = encode phase and amplitude of each signal

determined by an additional SVD of $\mathbf{Z}$.

$$x = A\mathbf{Z}.$$

Forward-backward LP

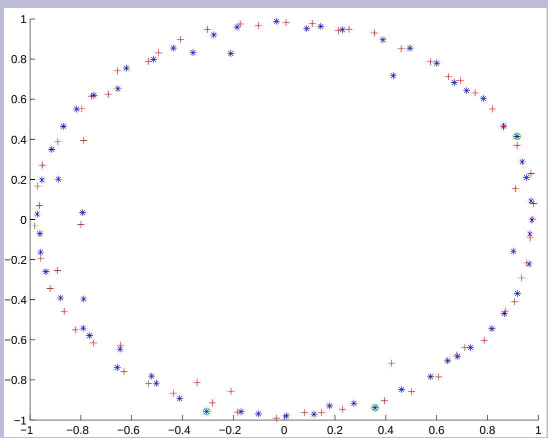
Calculate A_k LP coefficients for forward and backward LP

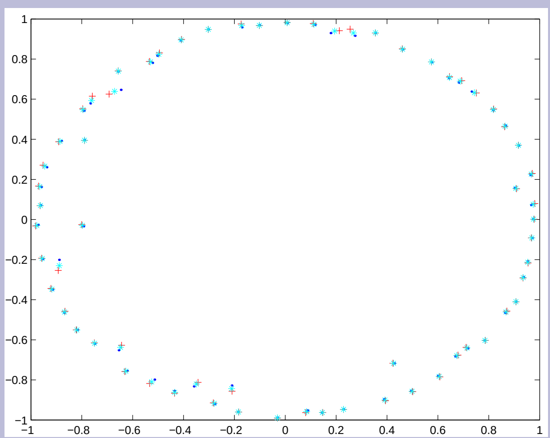
$\rightarrow$ two sets of polynomial roots

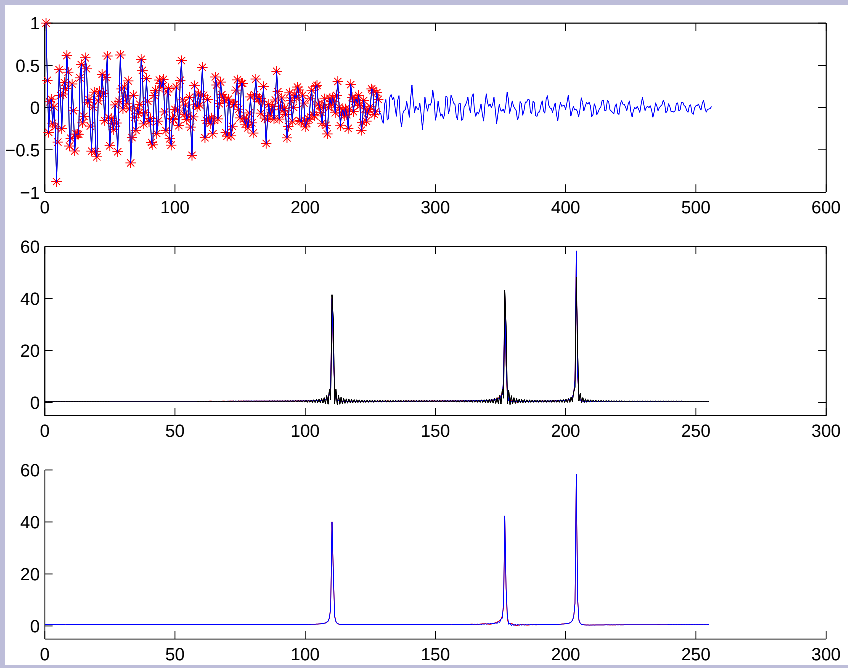
Forward linear prediction: complex roots inside the unit circle, outside: noise / exp. incr. signal

Backward linear prediction: inside the unit circle

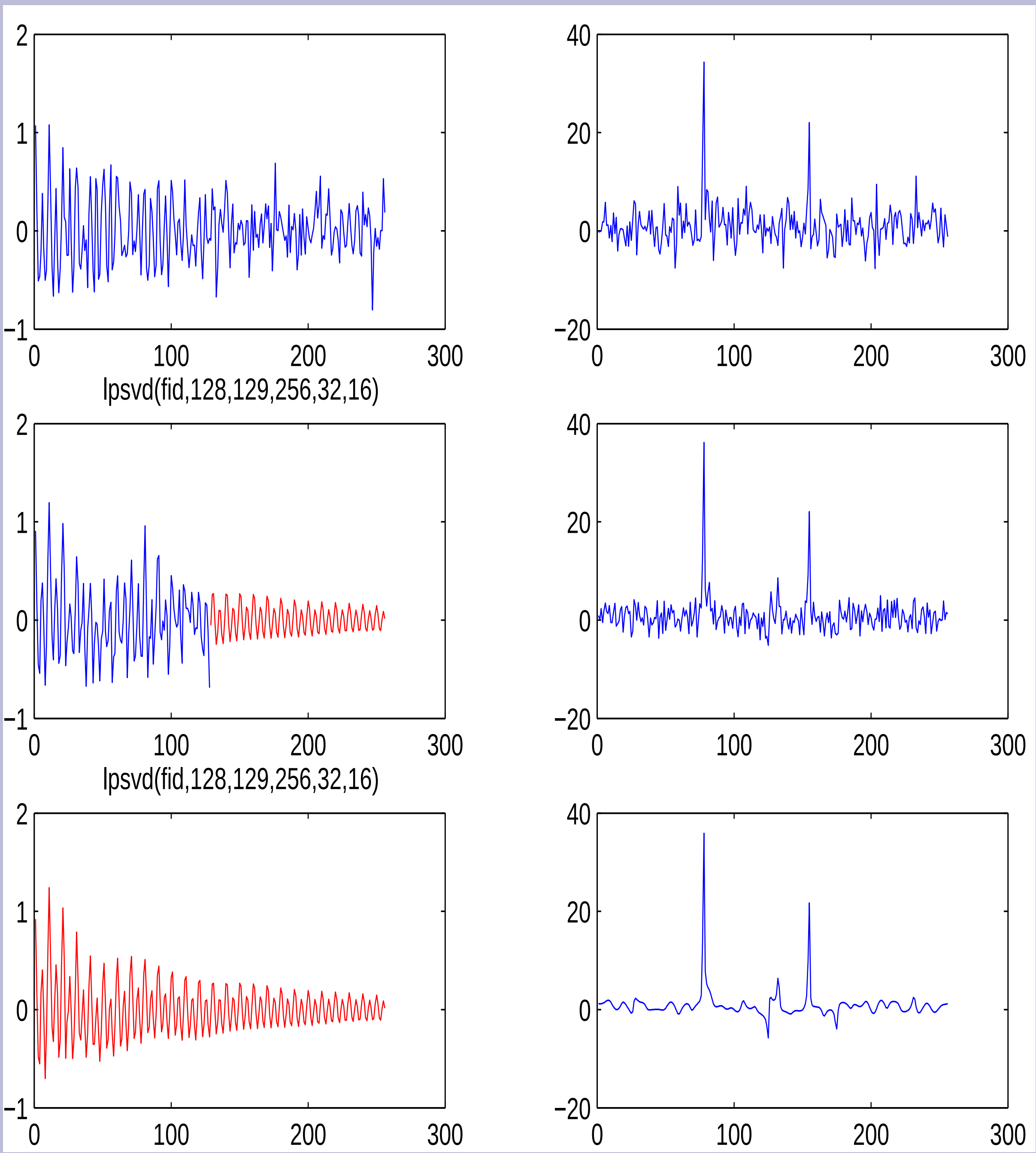
Root reflection: replace z_k by $z_k/|z_k^2| \Longrightarrow$ noise realted signals are eliminated



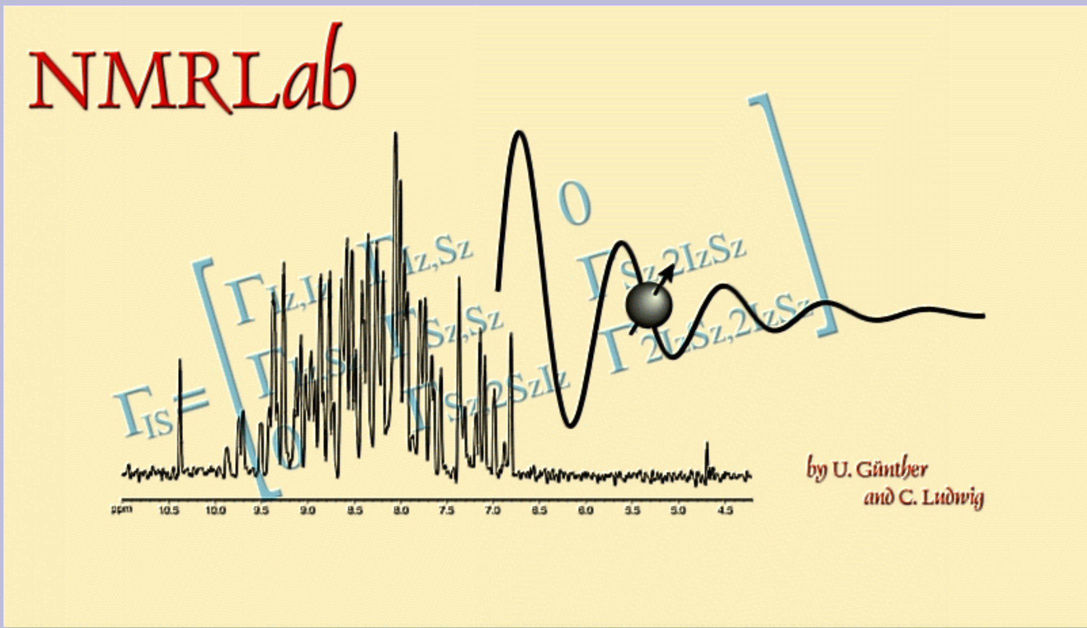




1
0.5
0
−0.5
−1
0
100
200
300
400
500
600
60
40
20
0
0
50
100
150
200
250
300
60
40
20
0
0
50
100
150
200
250
300



S/N enhancement by returning the complete predicted FID (lpsvd(fid,128,**1,256**,32,16))



- Standard processing: FFT, apodization functions, phase correction, baseline correction (FLATT and other algorithms), linear prediction (lpsvd with root reflection), Cadzow algorithm

- Export to NMRVIEW, AURELIA

- Peak picking: Export of peak lists to DYANA (NOAH)

- SAR by NMR evaluation: Principal component analysis, clustering

- Advanced wavelet algorithms
  - Noise reduction (2D and 3D)

– Wavelet based suppression of solvent signals
– Save data in wavelet format (50-80% reduced data size)

NMRLab

Table 1: NMRDAT structure

NMRDAT	**field name**	
	NAME	Dataset name. Used for saving sul
	SER	Converted SER file.
	MAT	Processed data matrix.
	ACQUS	Acquisition parameters (3D).
	PROC	Processing parameters (3D).
	DISP	Display parameters.

NMR processing functions in NMRLAB.	
Function	**Description**
`hft`	Hilbert Transform
`fft`*	Fast Fourier transform
`ift`*	Inverse fast Fourier transform
`dft`	Fourier transform of BRUKER digital filtered data
`rft`	Real Fourier transform for TPPI-type data
`smo`	smooth = polynomial sovent filter [**?**]
`sol`	solvent filter by time-domain convolution [**?**]
`wavewat`	WAVEWAT water suppression
`wdwf2`	Window functions (gm, em, sine bell, cubic sine bell)
`baseline2`	Different algorithms for baseline correction

Control functions in NMRLAB.	
Function	**Description**
nmrlab	MATLAB script to setup parameters for NM
re	read raw data from disk
relist	read series of experiments
readser	read BRUKER ser files.
readacqus	read parameters from BRUKER ser file
snd	show data sets and sizes NMRDAT
uiphase	interactive phase correction
uicont	interactive contour plotting*
sartitr	analyze series of 2D NMR spectra (e.g. SA
edp	edit proocessing parameters
edd	edit display parameters
browse	browse, edit, save, export and load NMRD
xfb	process two dimensional data
xfall	process series of 2D data sets (e.g. SAR
tf	process three dimensional data
absc	2D/3D post processing baseline correctio
phasend	2D/3D post processing phase correction
denoise	2D/3D post processing wavelet denoising
xyztranspose	transpose 3D structures
makespc	utility to create synthetic spectra

*MATLAB $\geq$ 5.3 will automatically activate graphical tools and

For MATLAB 5.0-5.2 use the zoom and plotedit commands ins

Table 3:

[ht]

Wavelet Shrinkage Parameters in NMRLAB		
qmf_type	wavelet type (quadrature mirror filter)	Haar, *Coiflet*, *Daubec* Vaidyar
par	QMF paramater	Coiflet: (*3*).Dau 4,6,8,*1* Symml (*8*).
thr_type	Type of shrinkage	*hard*, SURE, MinMax
L	Low-Frequency cutoff for shrinkage.	must b N = nu
normalize	normalize noise *	
WT_type**		*period* fully tra
thr	threshold value	manual *versal* = numb points.
*2D and 3D version of normalization has been implemented in		
**Other wavelet transforms (i.e. the Meyer wt) are available in		
Parameters which yield good results for NMR spectroscopy ha		

[ht]

Type of LP	algorithm	parameter ve
`lpsvd`	SVD based linear prediction	[FID, start,sto
`prony`	PRONY's method	[FID,NB,NA]
`stmbc`	Steiglitz McBride	[FID,NB,NA,N
`lpc`	LPC	[FID,N]