Some aspects of Huang's "Empirical Mode Decomposition," from interpretation to applications

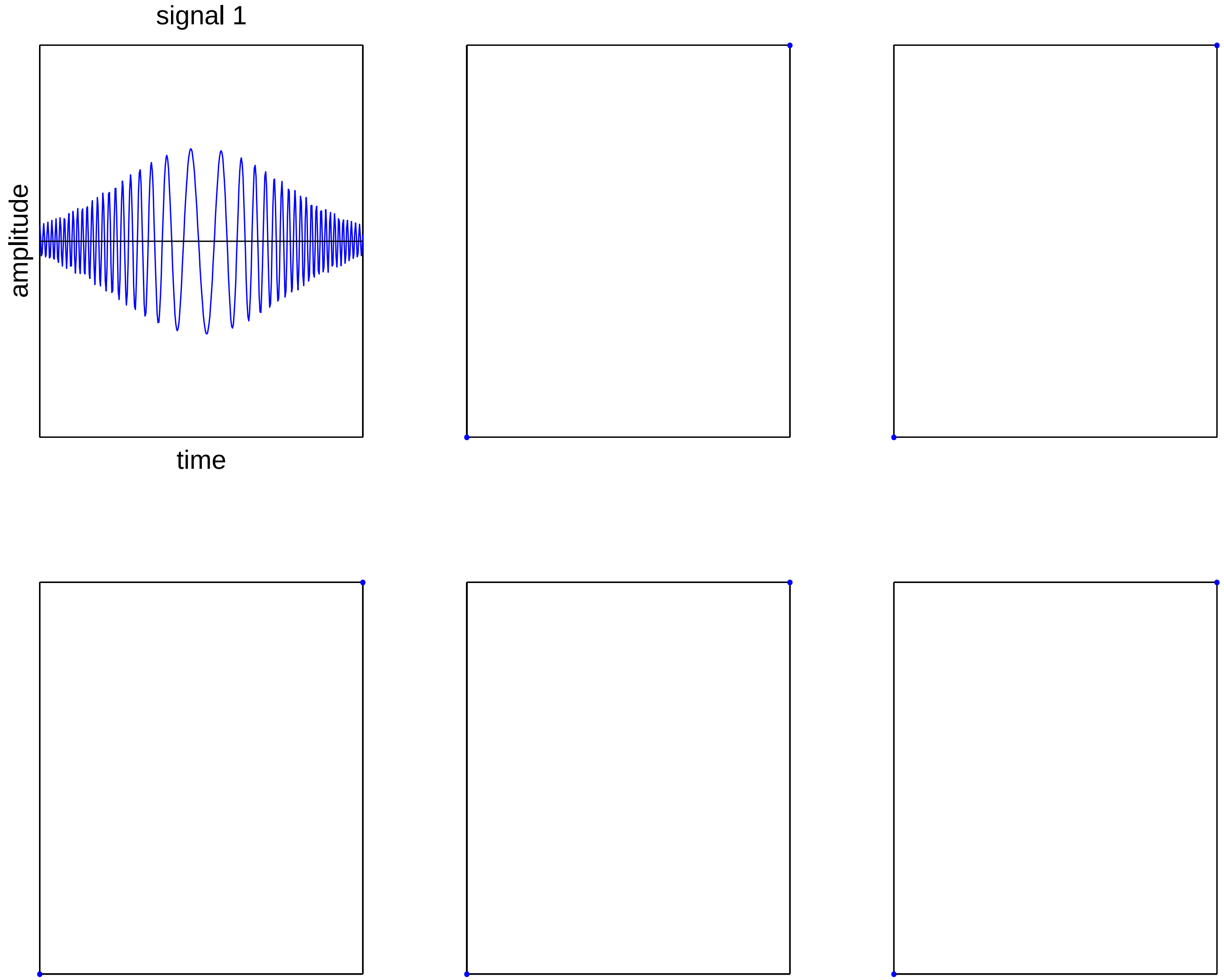
Patrick Flandrin*

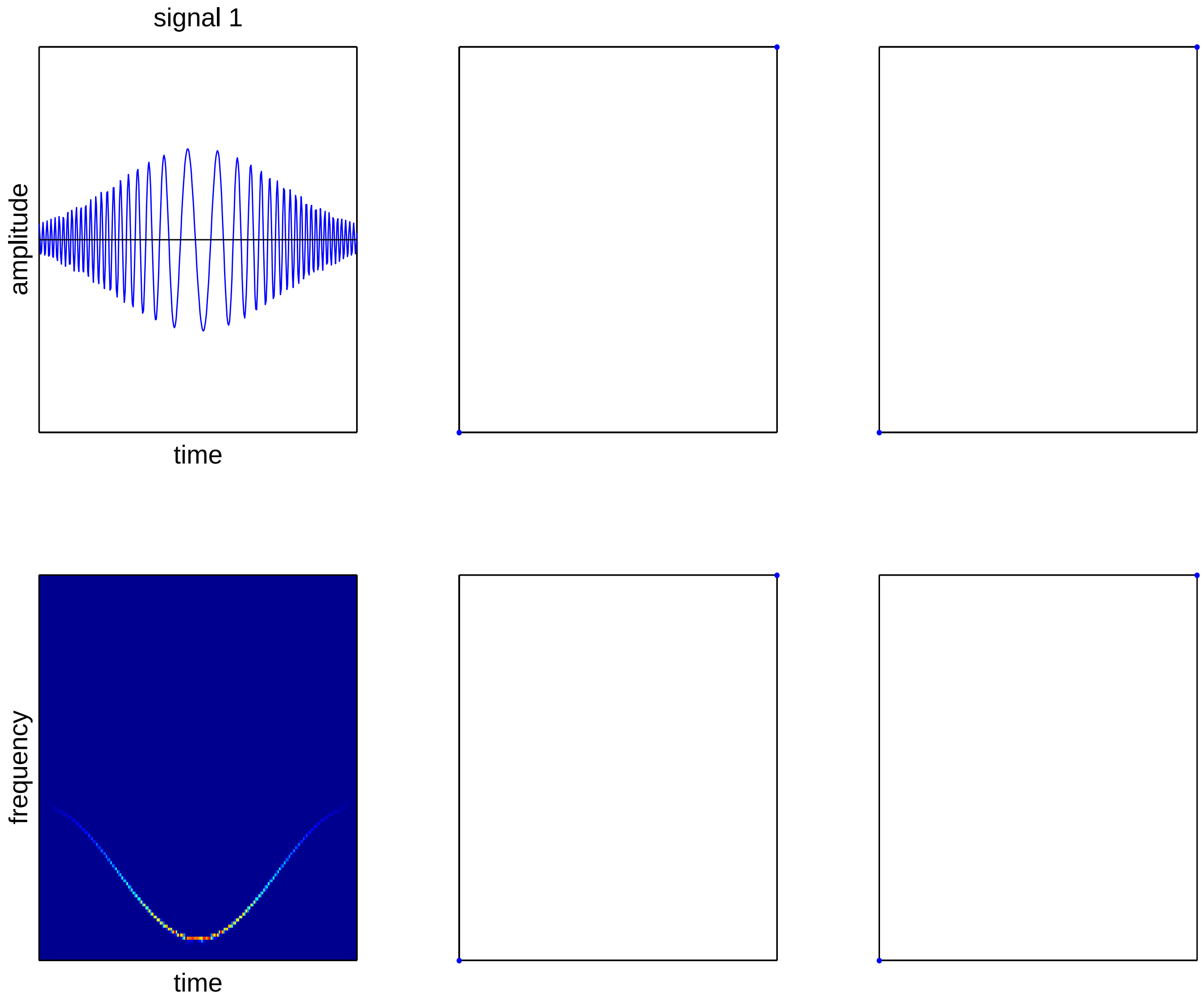
CNRS — École Normale Supérieure de Lyon, France

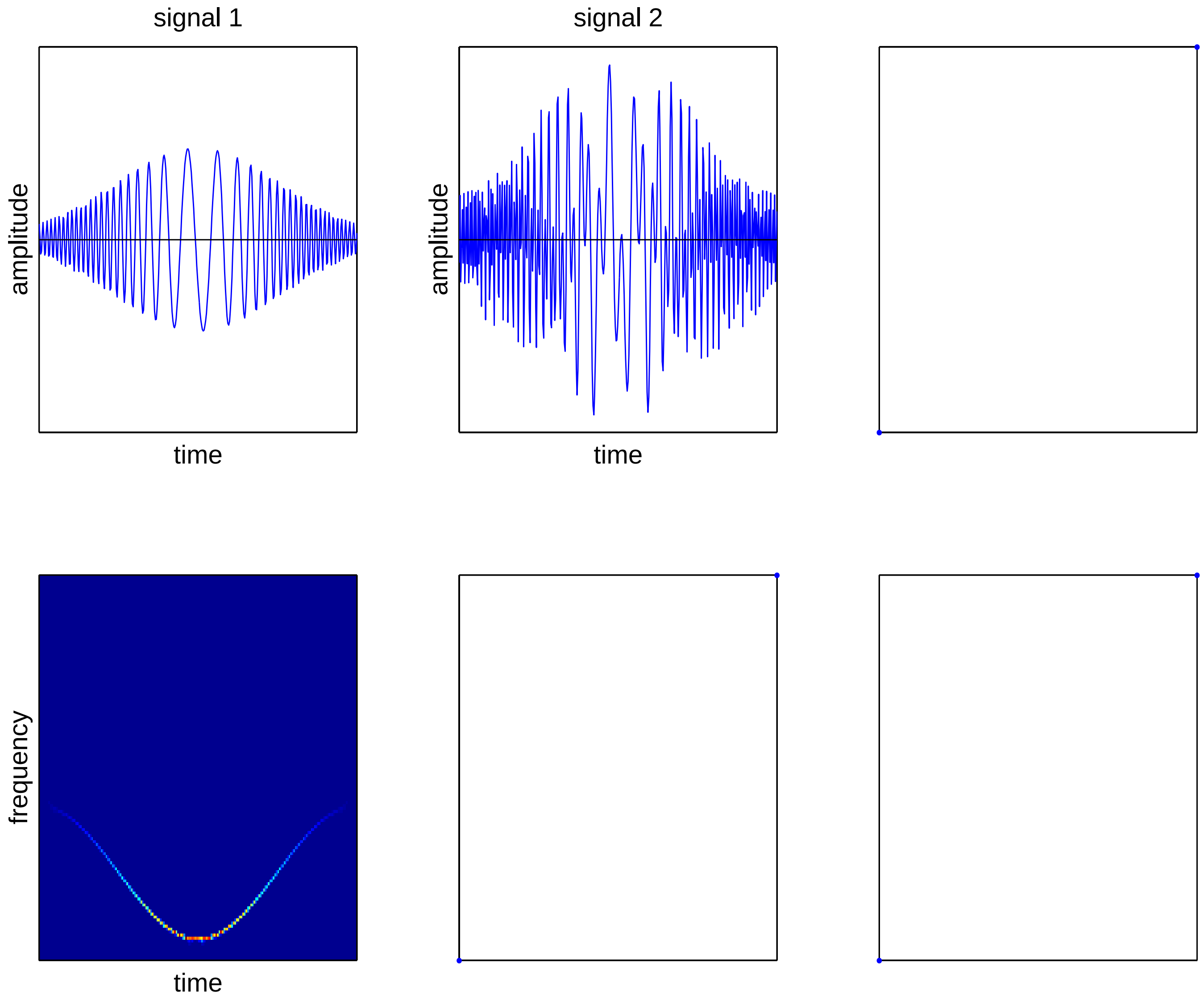
flandrin@ens-lyon.fr

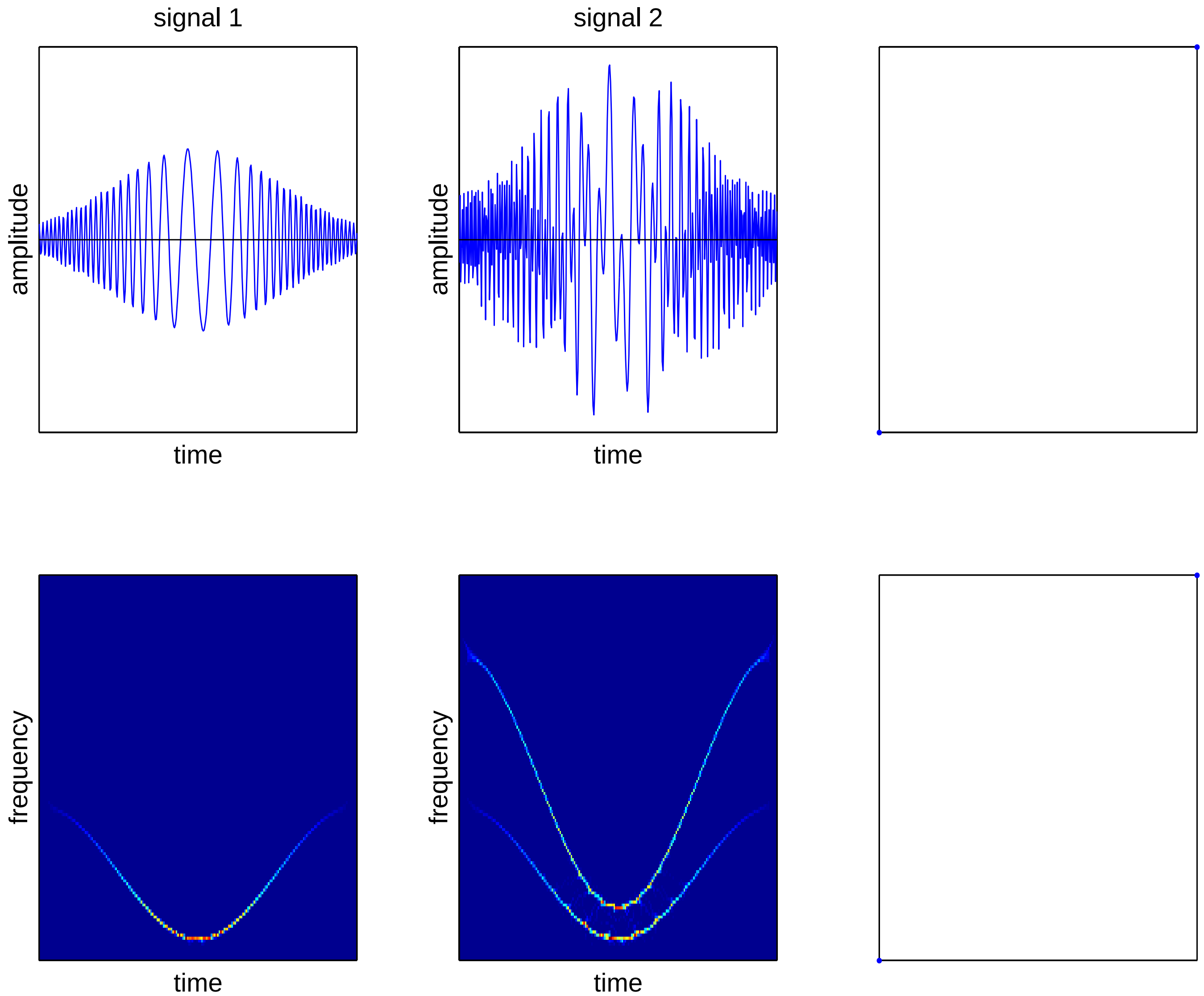
CHA'04 — Nashville (TN), May 2004

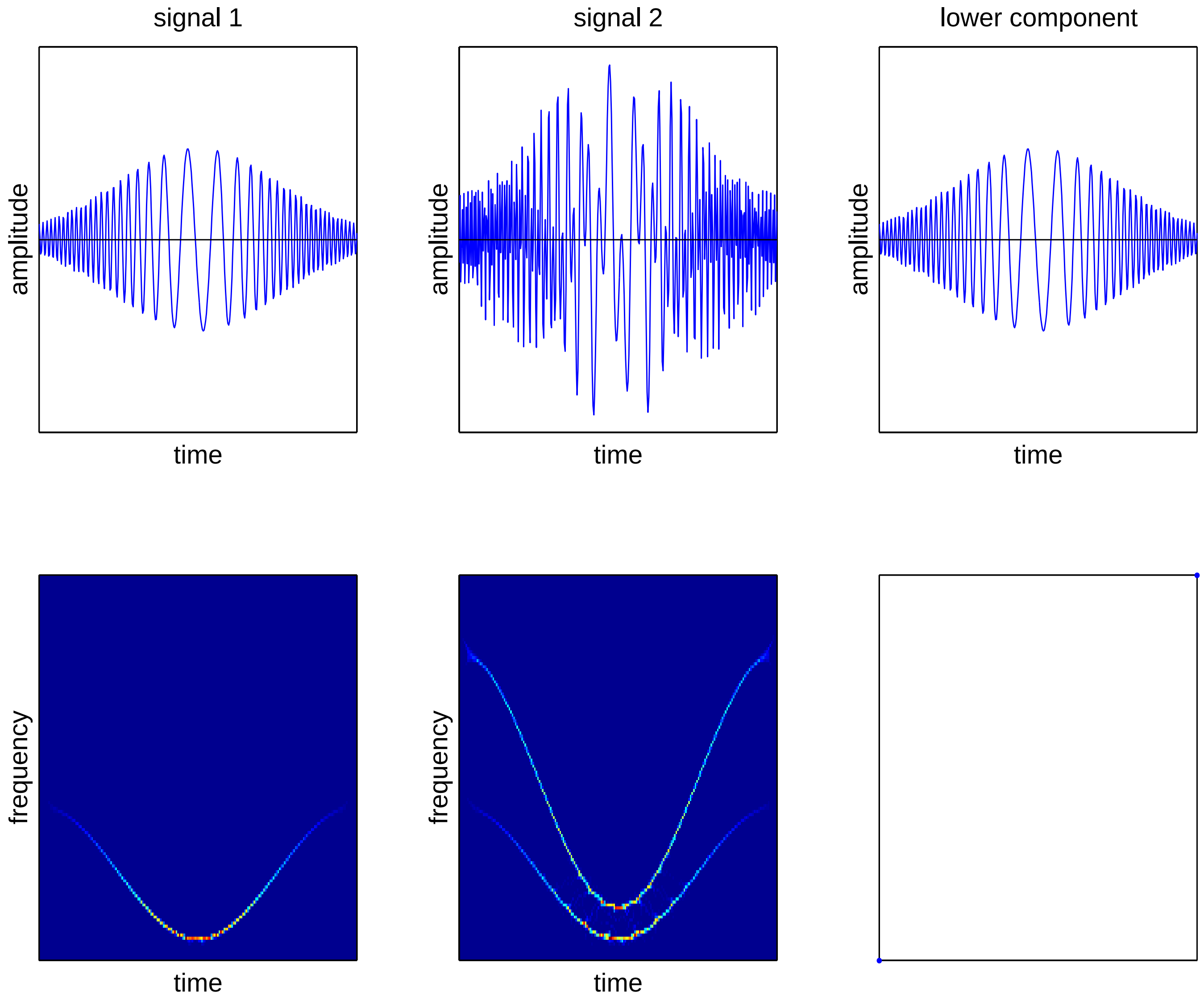
*joint work with
Paulo Gonçalvès (INRIAlpes & IST Lisbon) and Gabriel Rilling (ENS Lyon)


signal 1
amplitude
time

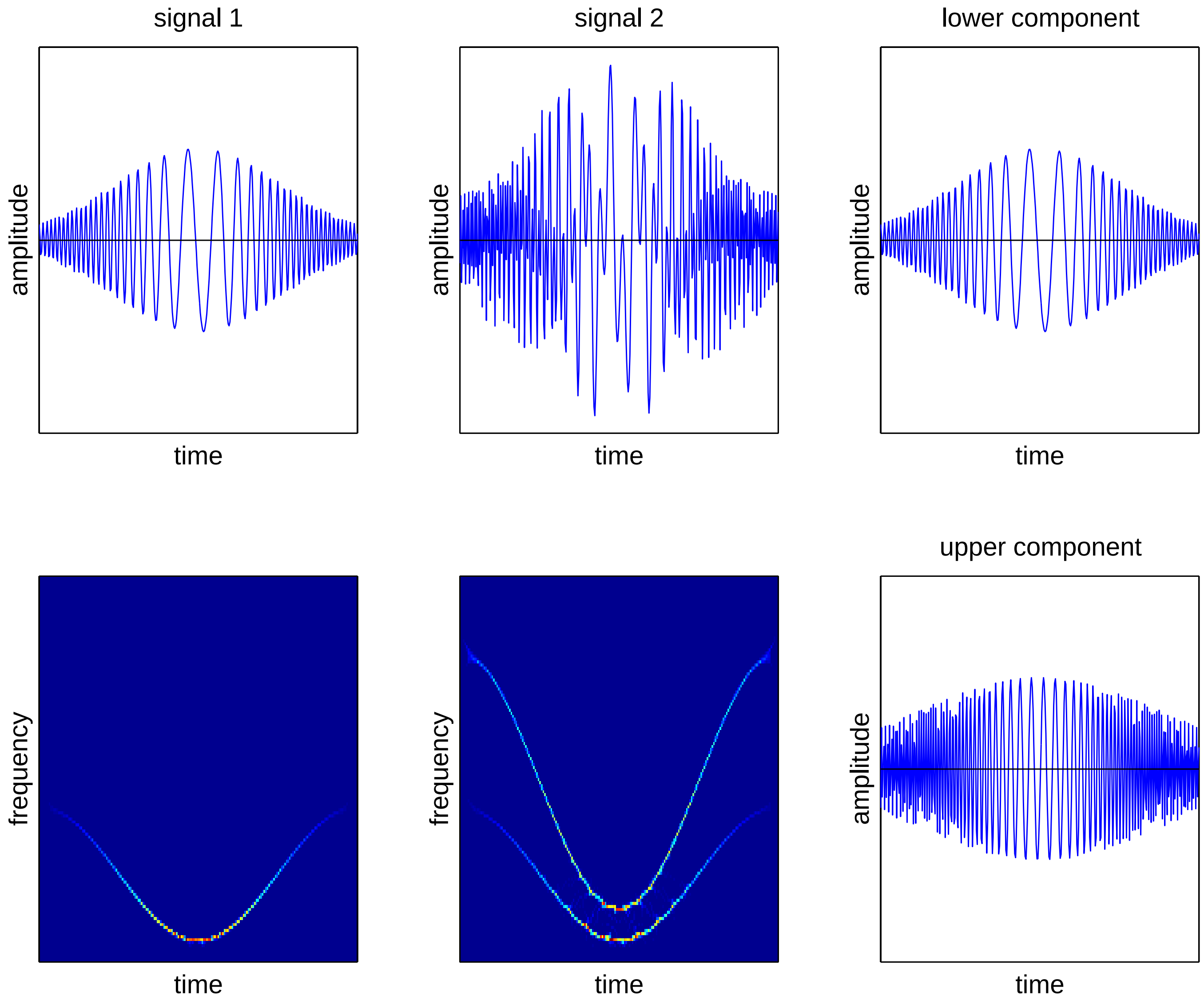

signal 1
amplitude
time
frequency
time

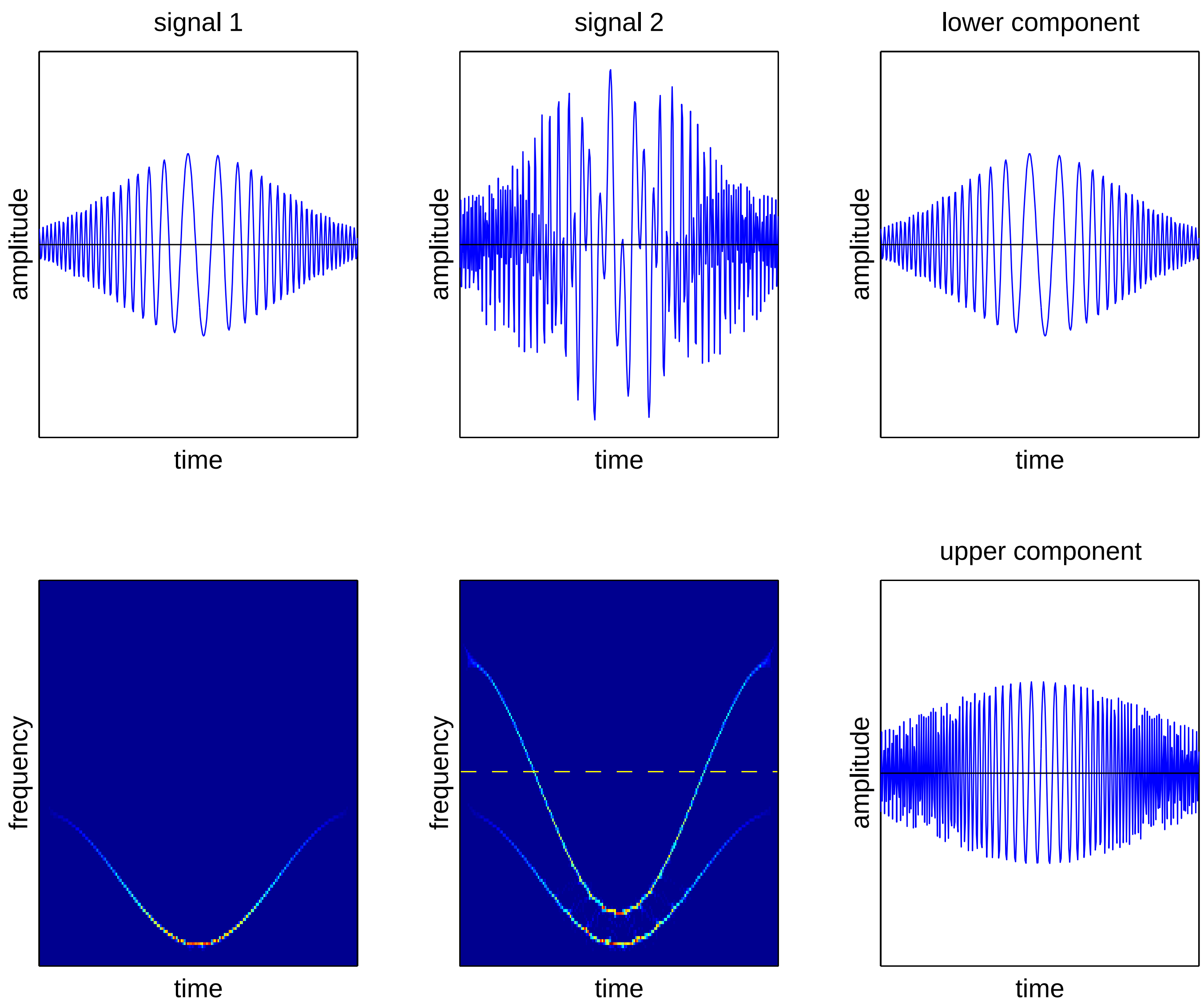

signal 1
signal 2
amplitude
time
amplitude
time
frequency
time

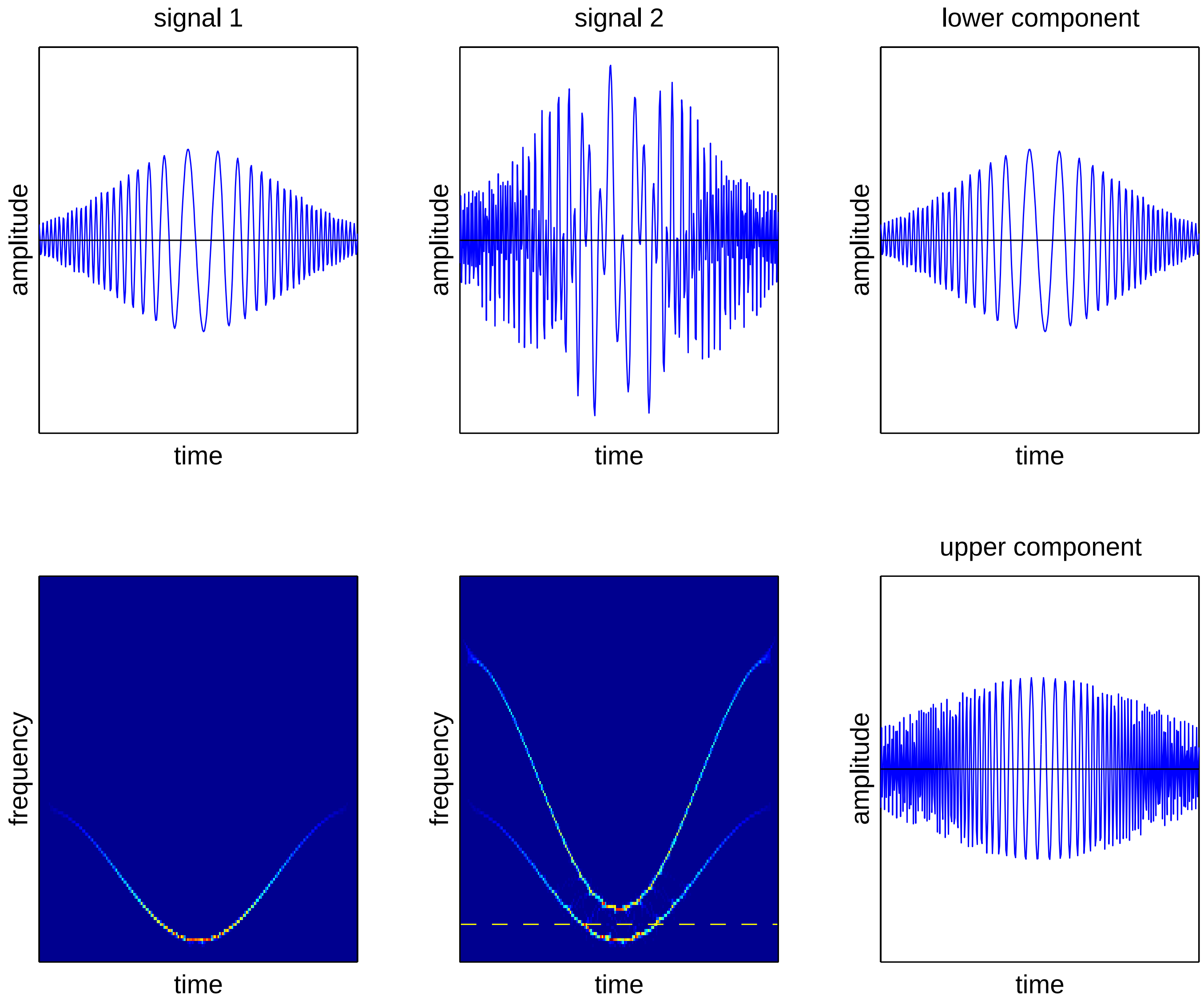

signal 1
signal 2
amplitude
time
amplitude
time
frequency
time
frequency
time

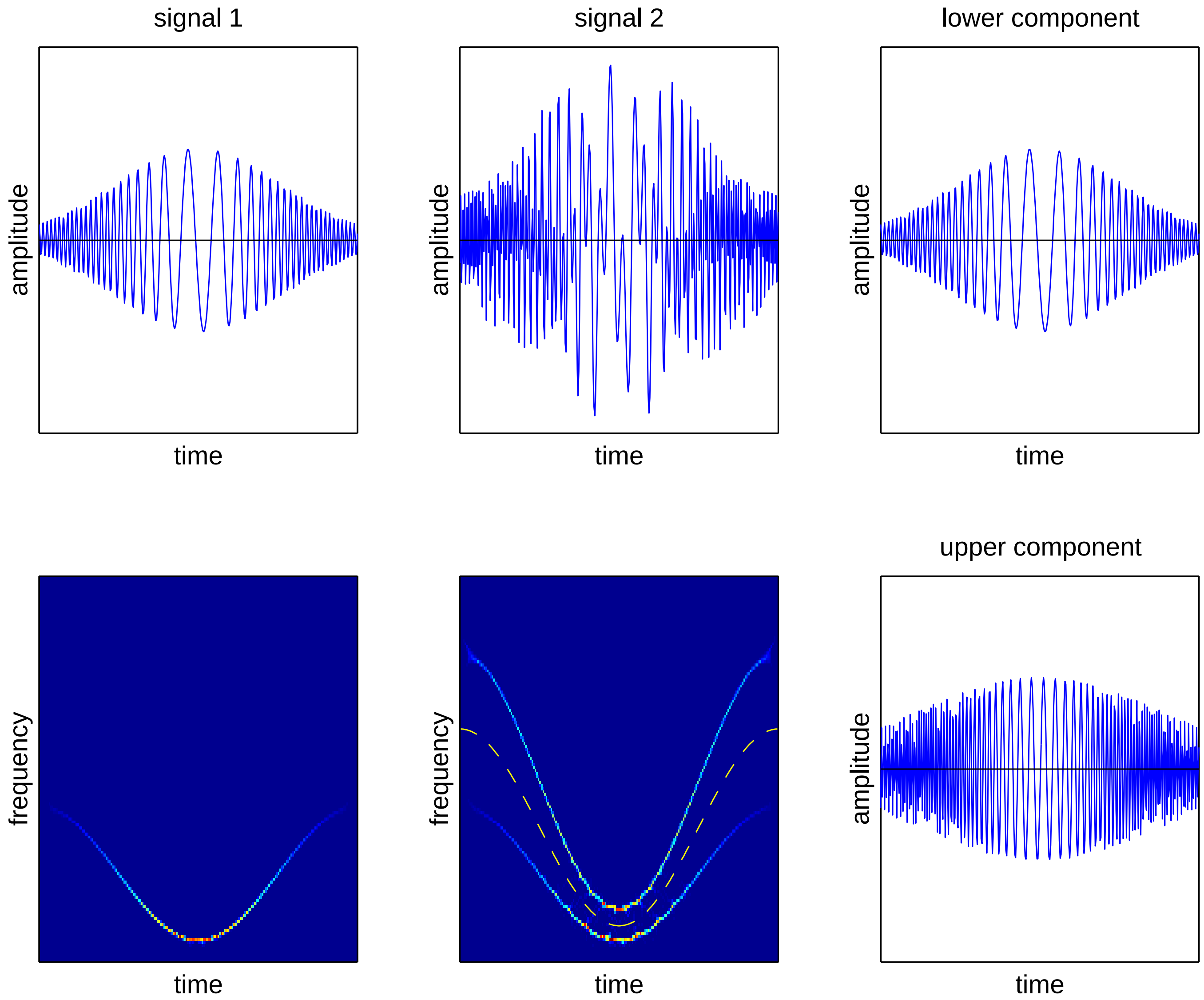


signal 1
signal 2
lower component
amplitude
time
amplitude
time
amplitude
time
frequency
time
frequency
time


signal 1
signal 2
lower component
amplitude
time
upper component
frequency


signal 1
amplitude
time
frequency
time
signal 2
amplitude
time
frequency
time
lower component
amplitude
time
upper component
amplitude
time


signal 1
amplitude
time
signal 2
amplitude
time
lower component
amplitude
time
frequency
time
frequency
time
upper component
amplitude
time


signal 1
signal 2
lower component
upper component
amplitude
time
frequency

Empirical Mode Decomposition

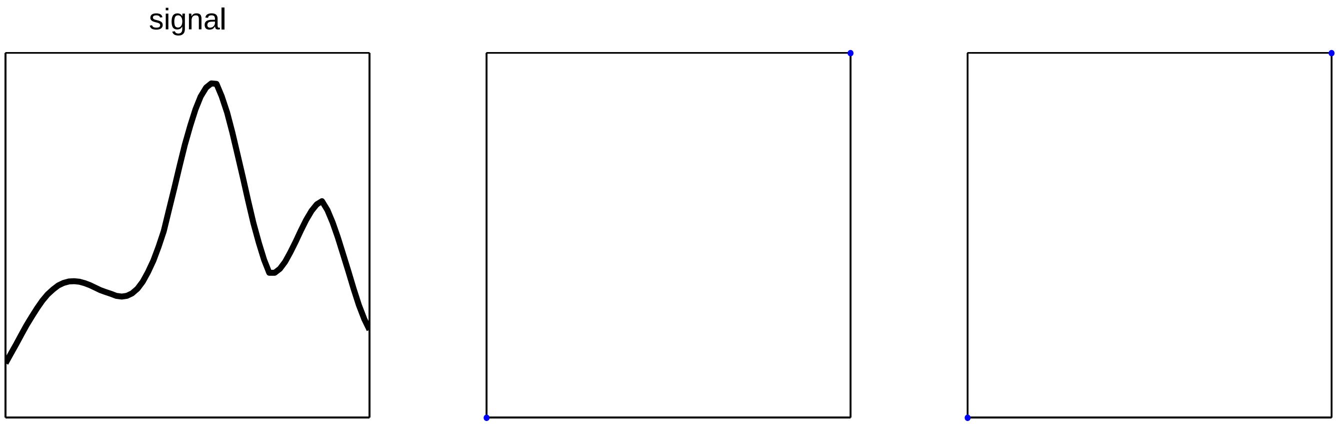
Problem — Given an observation $x(t)$, get a representation of the form:

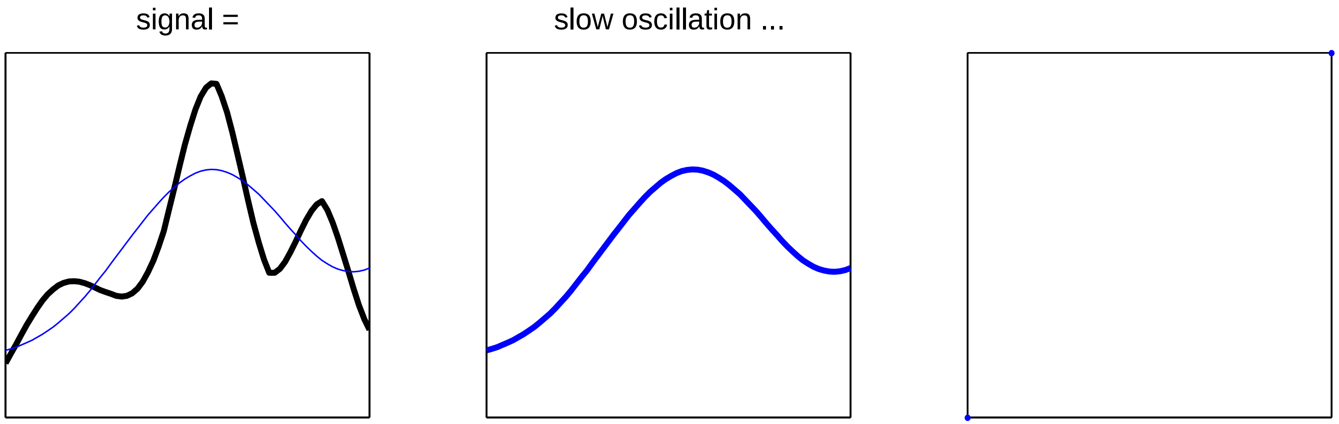
$$x(t) = \sum_{k=1}^{K} a_k(t)\, \psi_k(t),$$

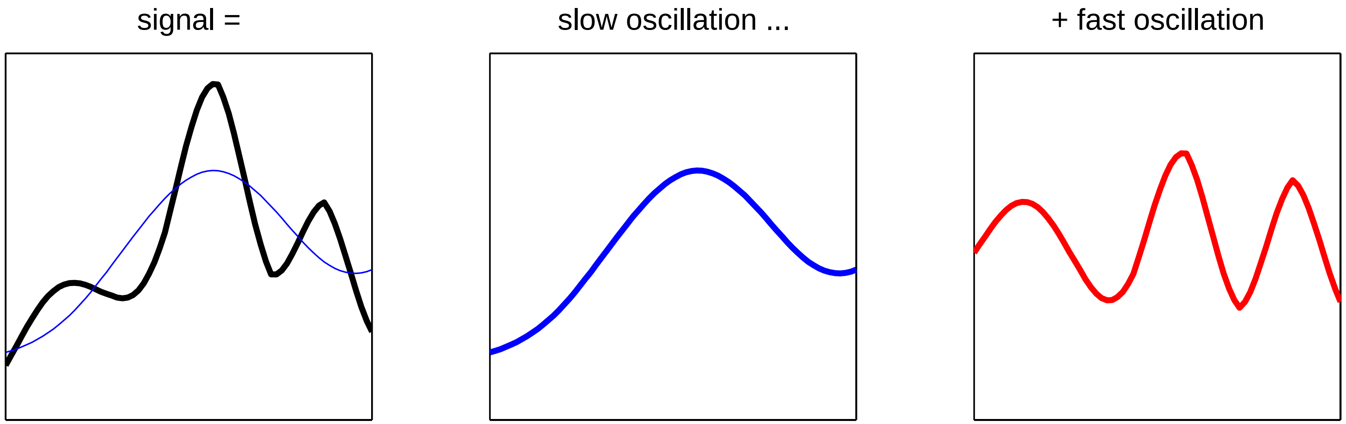
where $a_k(t)$ refer to "amplitude modulations" and $\psi_k(t)$ to "oscillations".

Intuitive idea — "signal = fast oscillations superimposed to slow oscillations".

Implementation (Huang *et al.*, '98) — (1) identify (locally) the fastest oscillation; (2) subtract to the original signal; (3) iterate on the residual.


signal


signal =
slow oscillation ...


signal =
slow oscillation ...
+ fast oscillation

Huang's algorithm

1. identify local maxima and local minima

2. deduce an upper envelope and a lower envelope by interpolation (cubic splines)

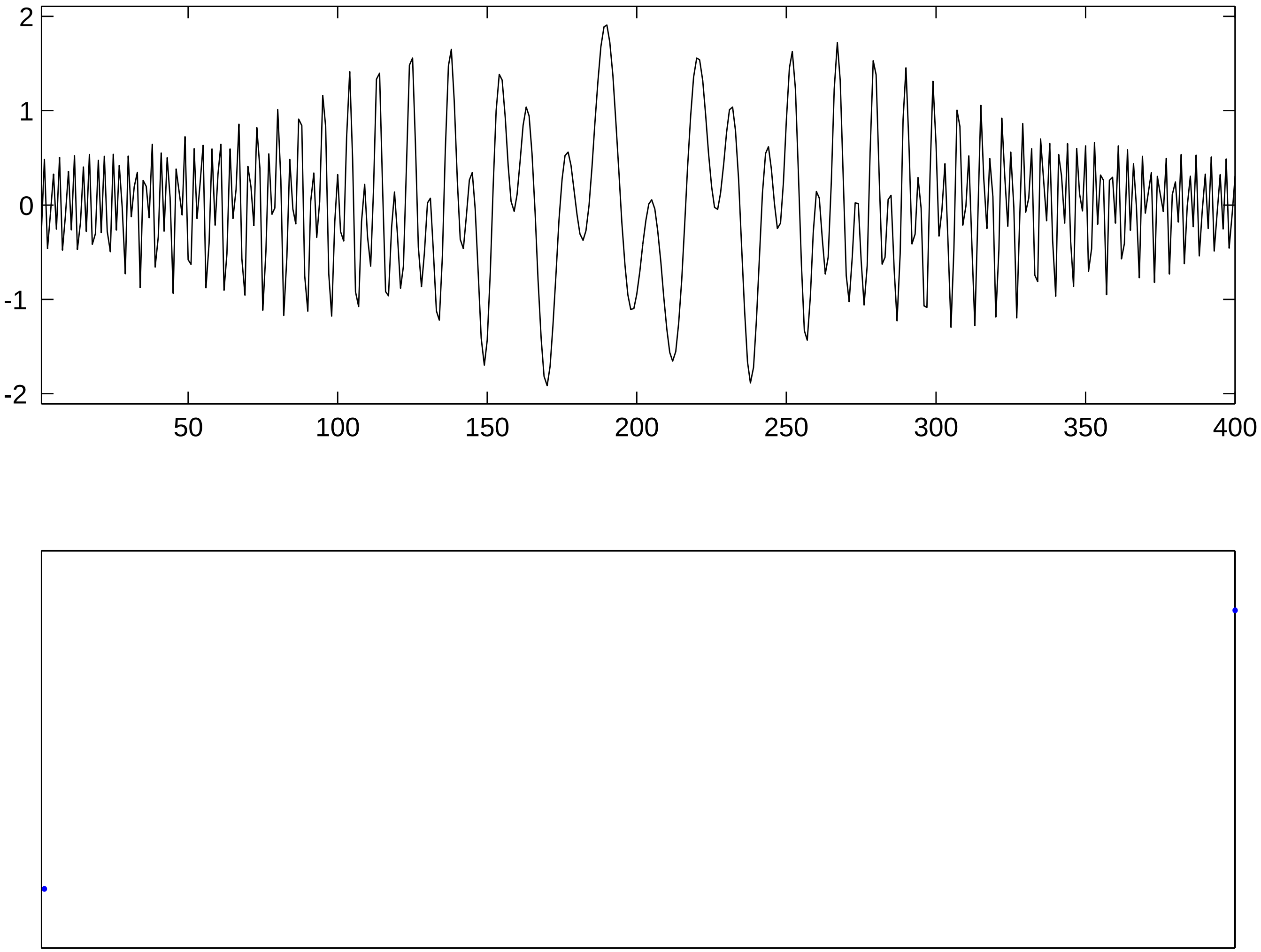
   (a) subtract the mean envelope from the signal

   (b) iterate until "mean envelope $= 0$" (*sifting*)

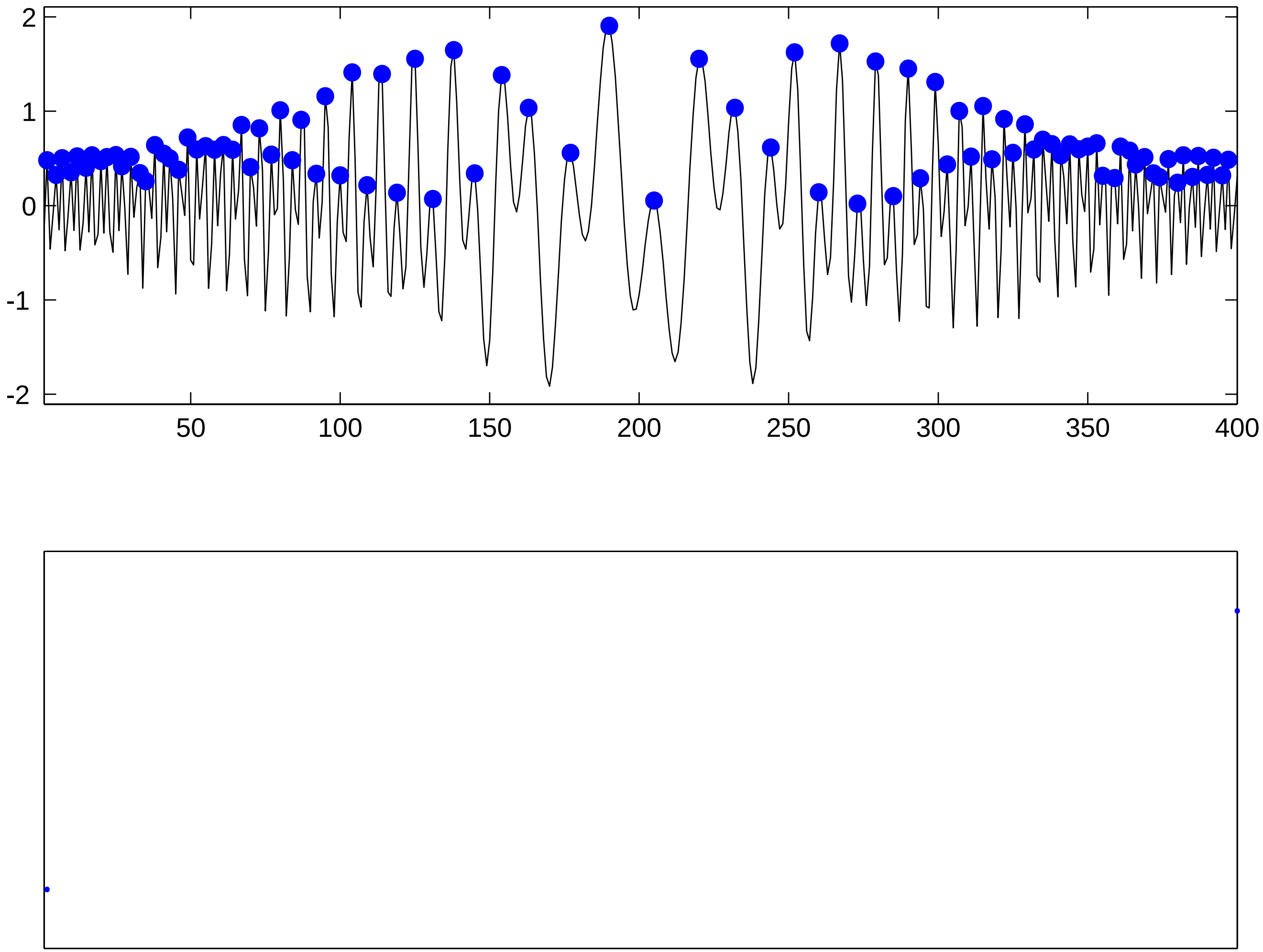
3. subtract the obtained mode from the signal

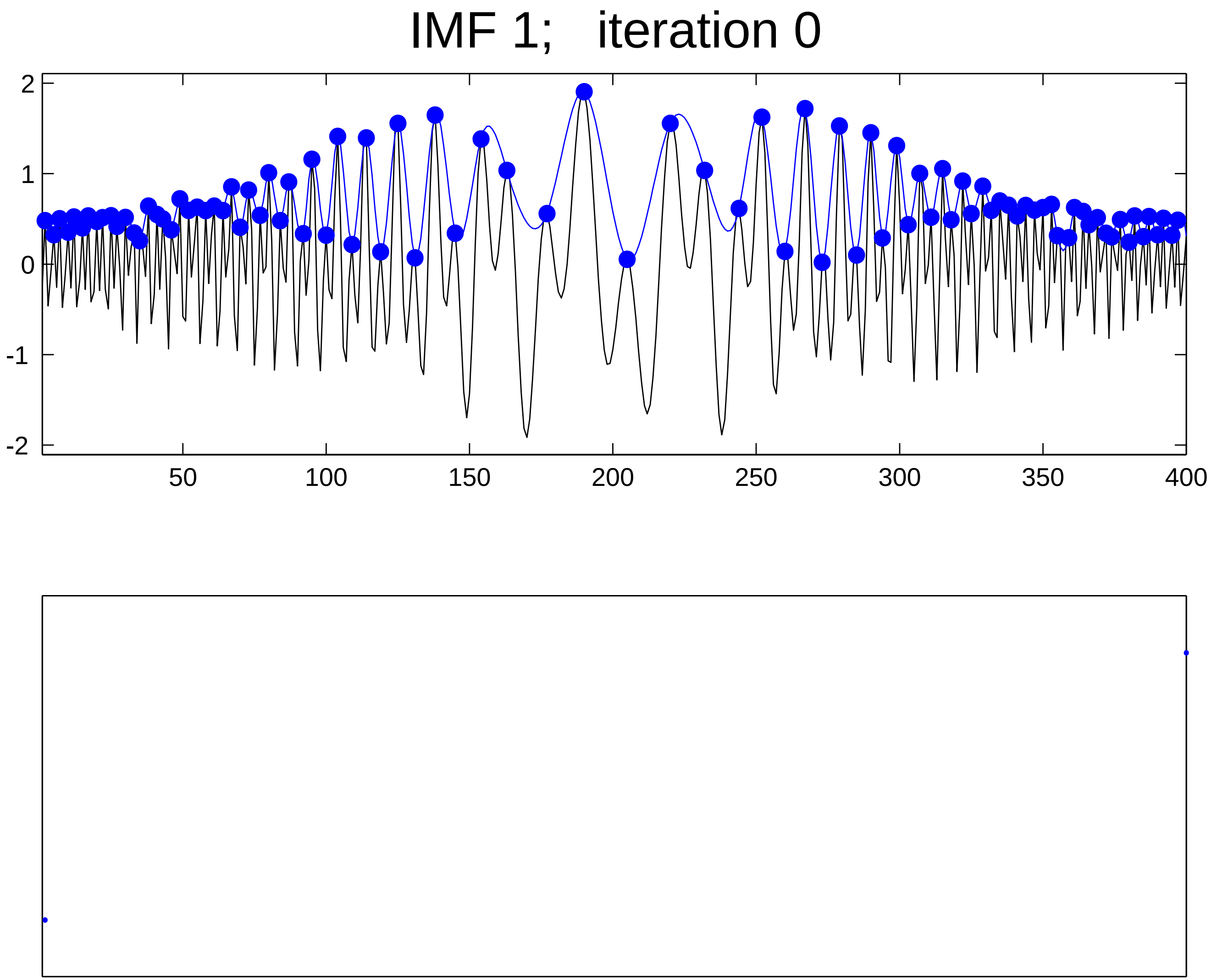
4. iterate on the residual

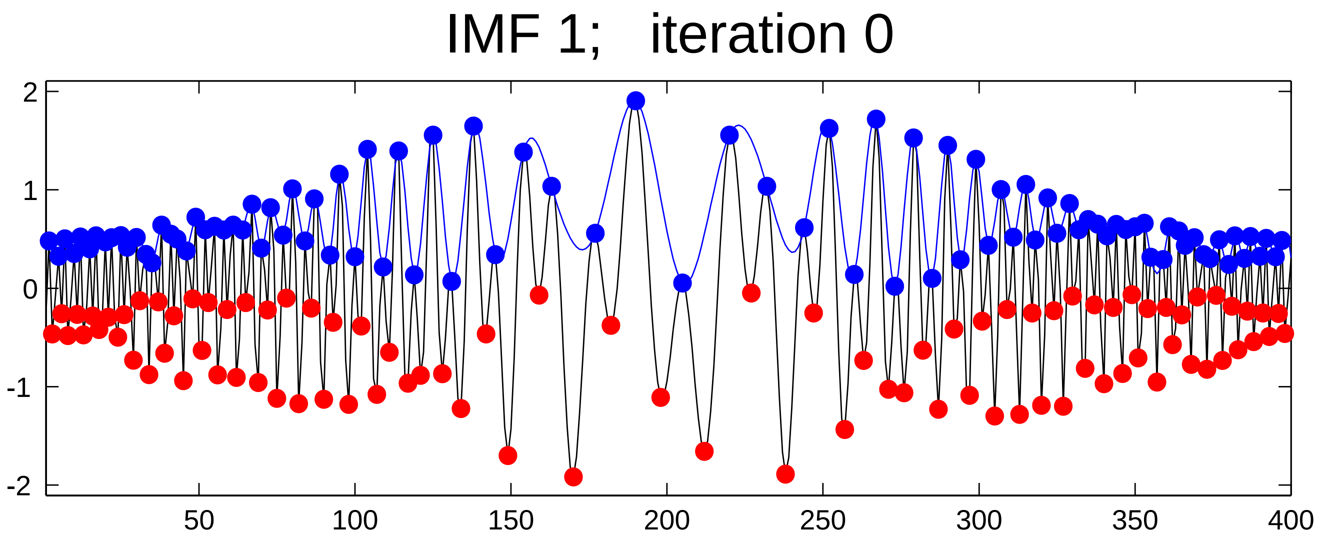
IMF 1;   iteration 0



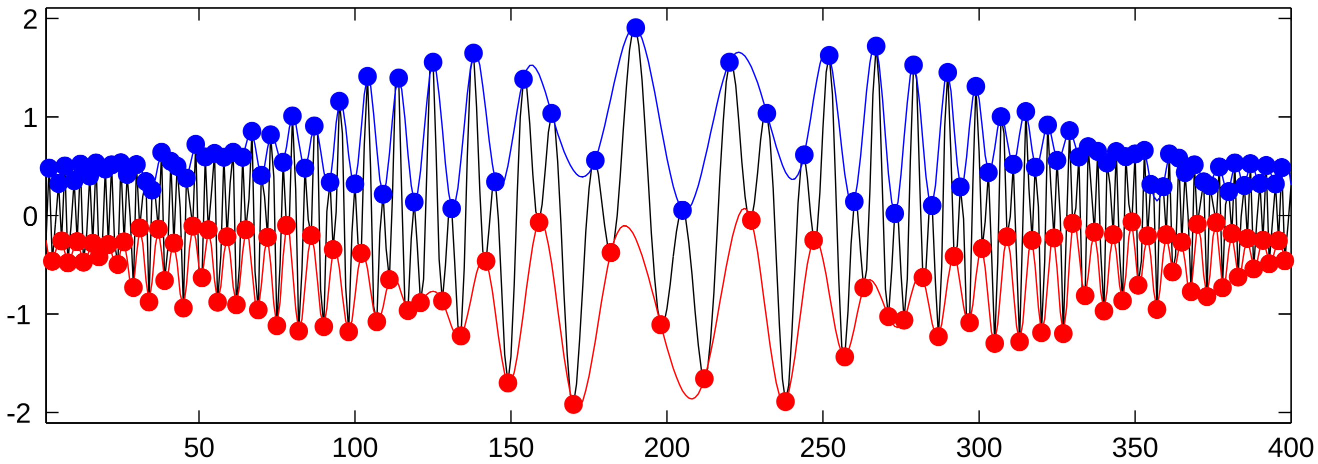
IMF 1;   iteration 0



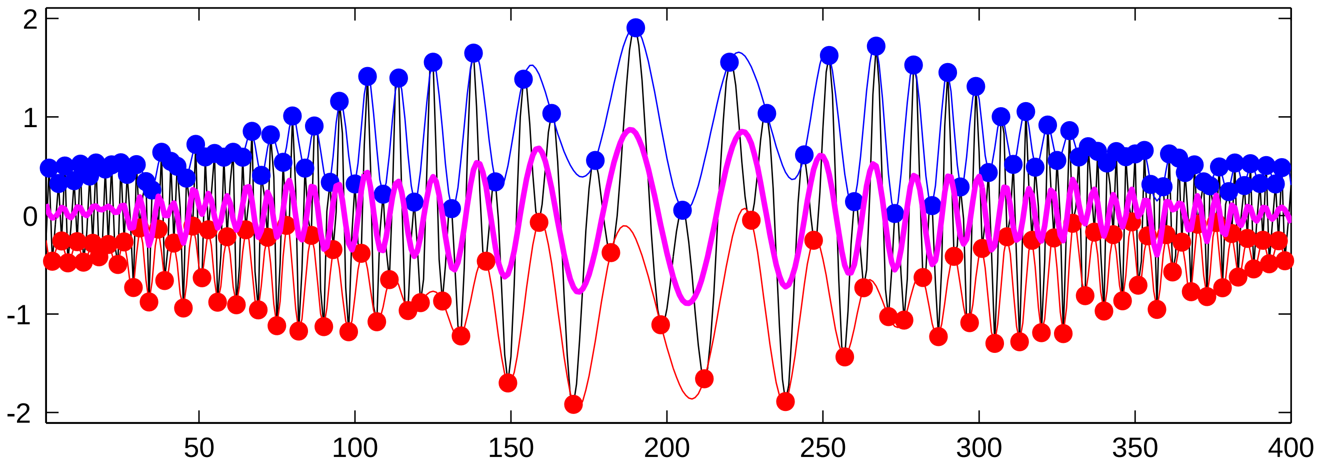

IMF 1; iteration 0
2
1
0
-1
-2
50
100
150
200
250
300
350
400

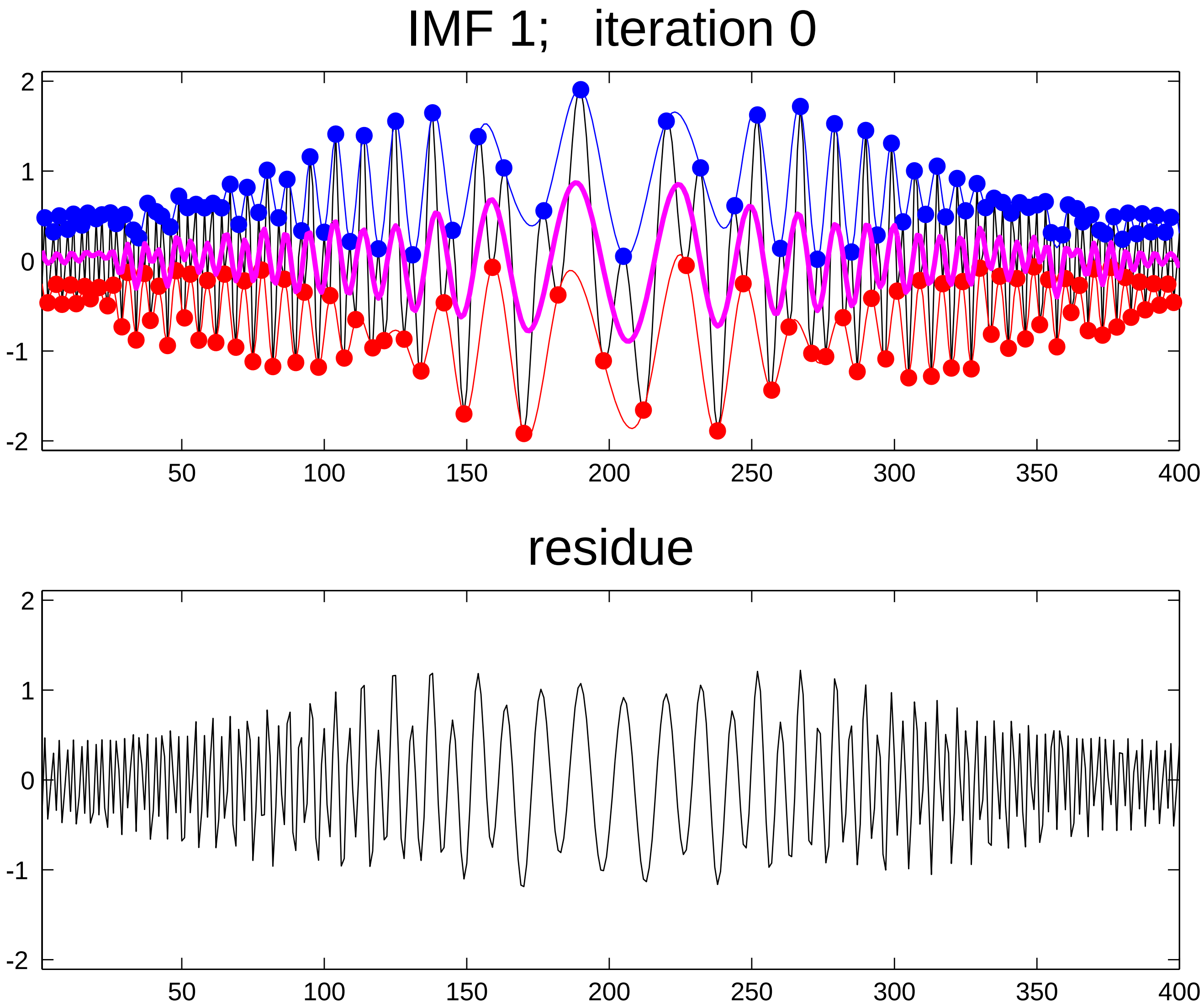

IMF 1; iteration 0
2
1
0
-1
-2
50
100
150
200
250
300
350
400

IMF 1;   iteration 0



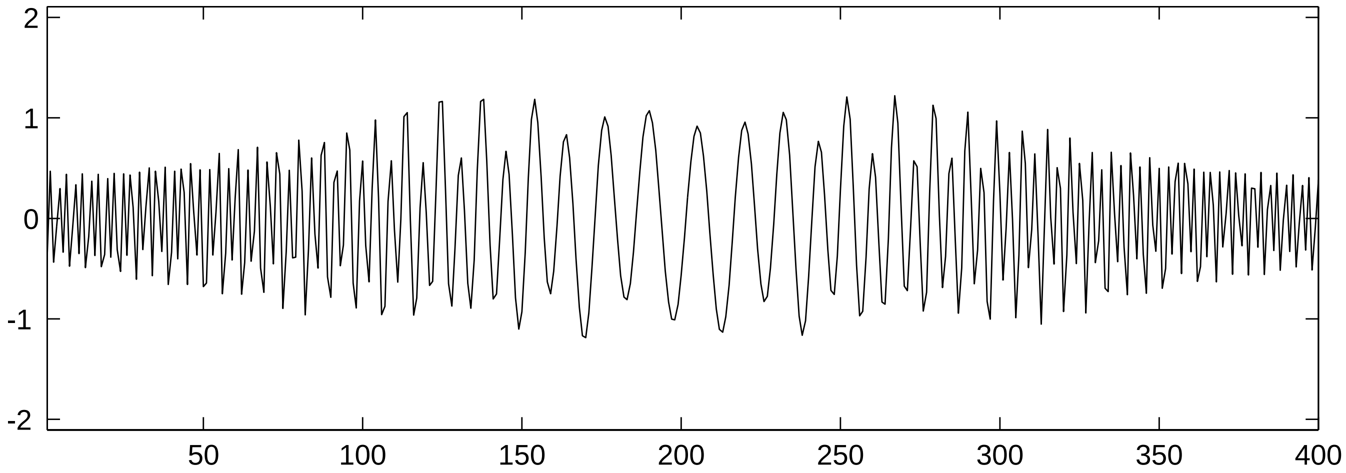
IMF 1;   iteration 0

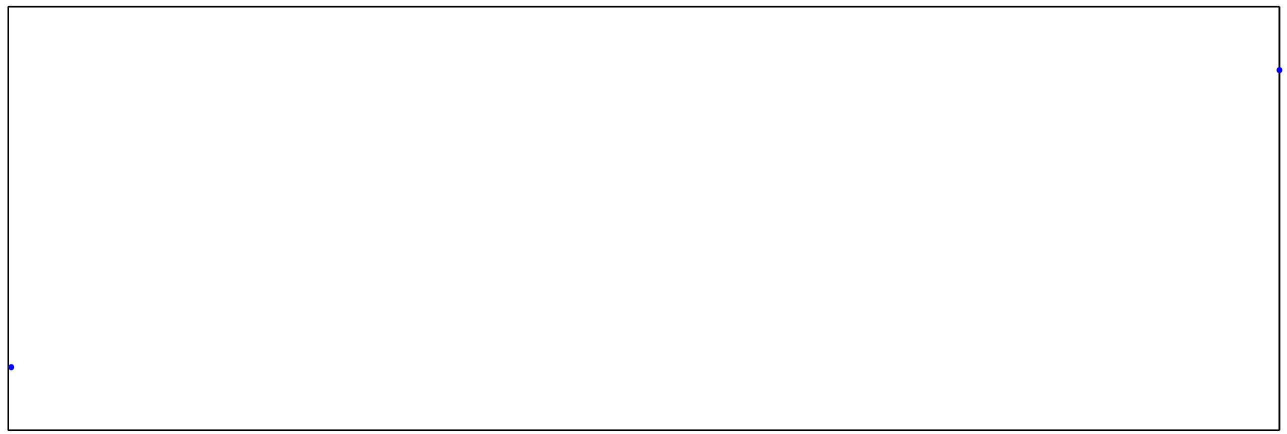




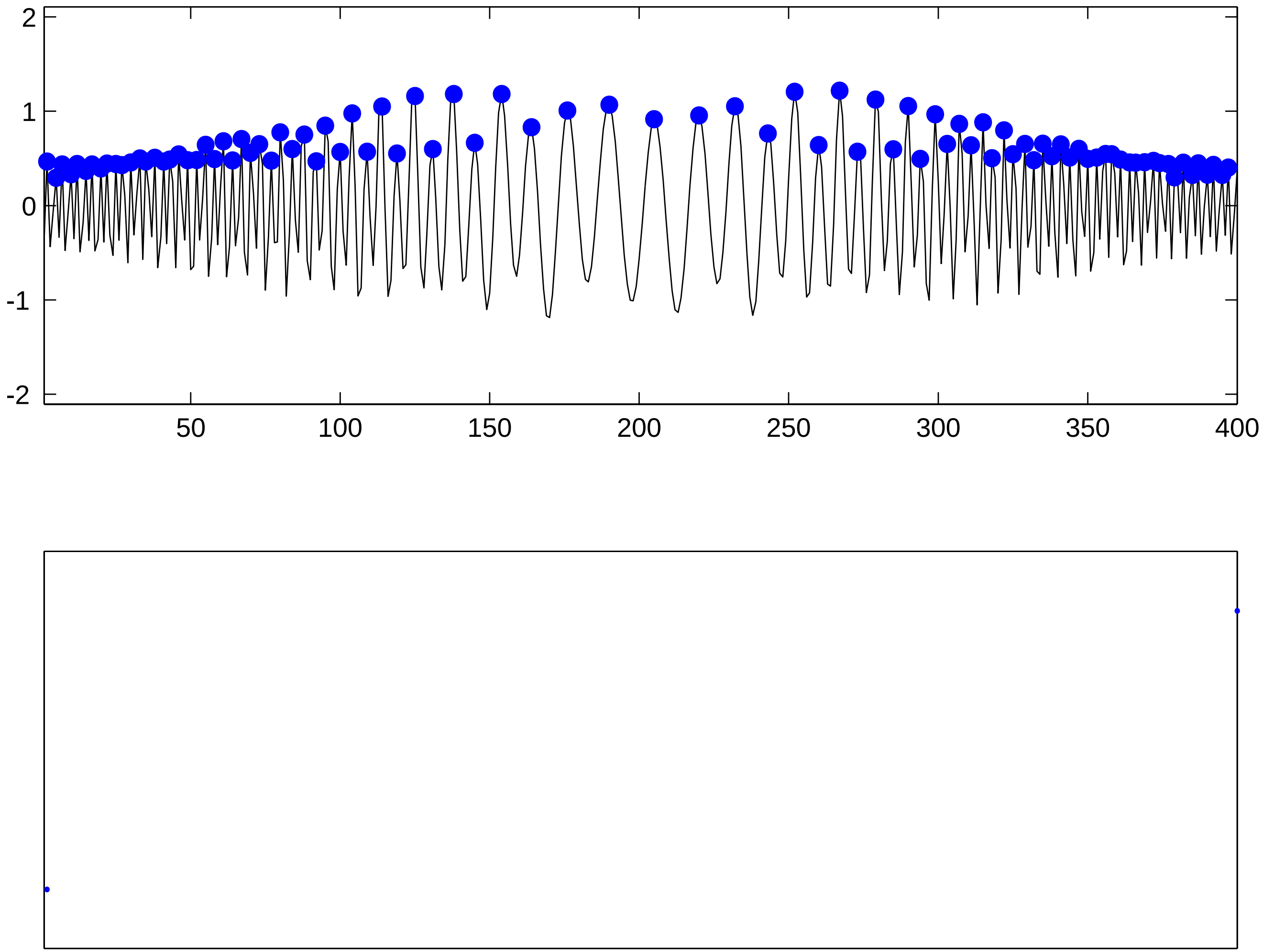
IMF 1; iteration 0
2
1
0
-1
-2
50
100
150
200
250
300
350
400
residue
2
1
0
-1
-2
50
100
150
200
250
300
350
400

IMF 1;   iteration 1

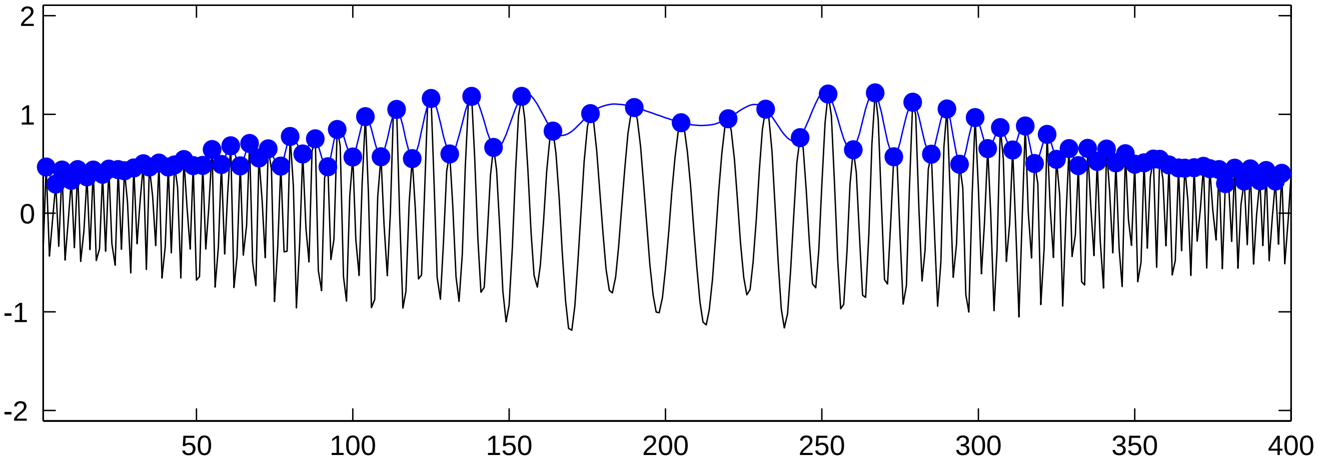


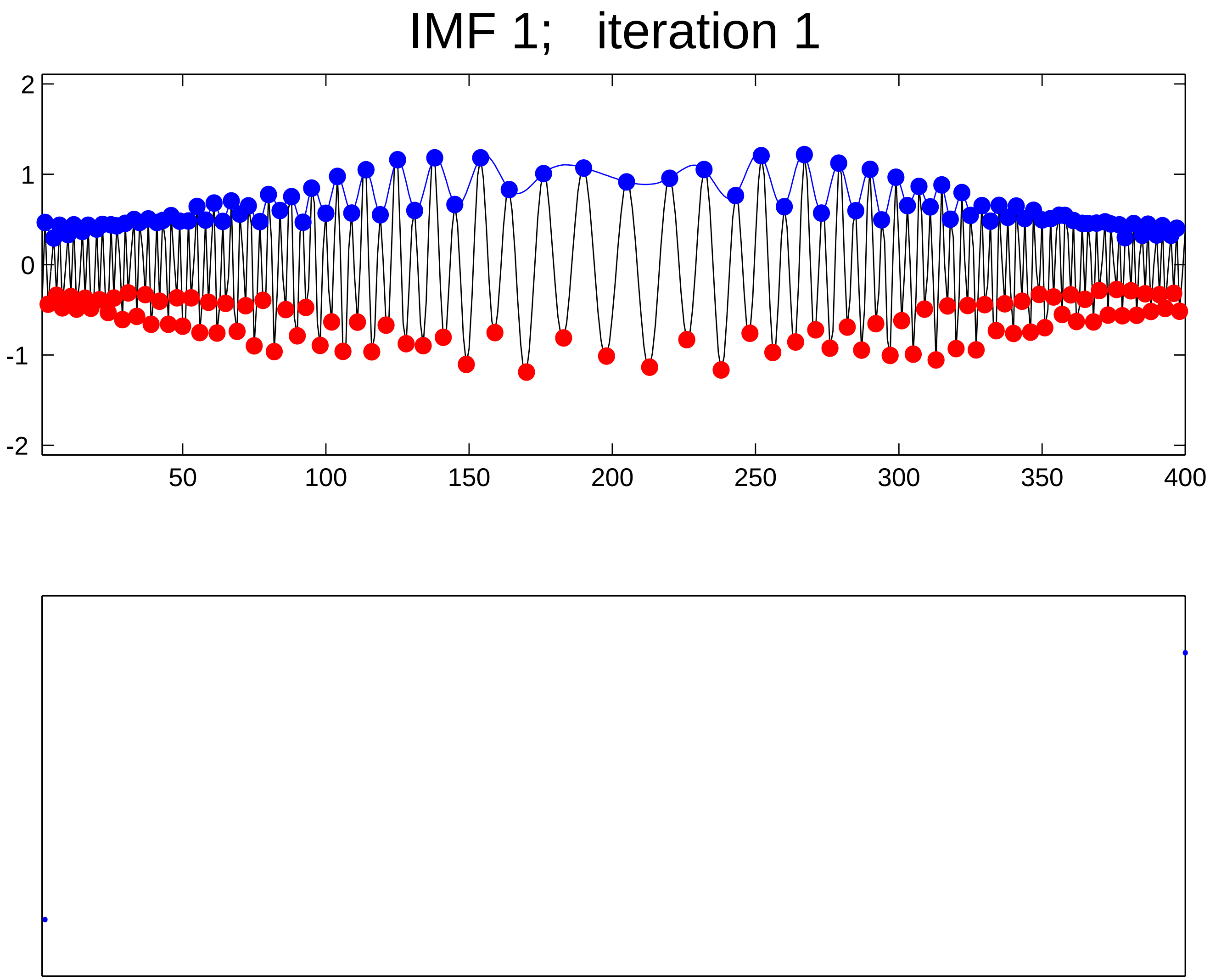


IMF 1;   iteration 1

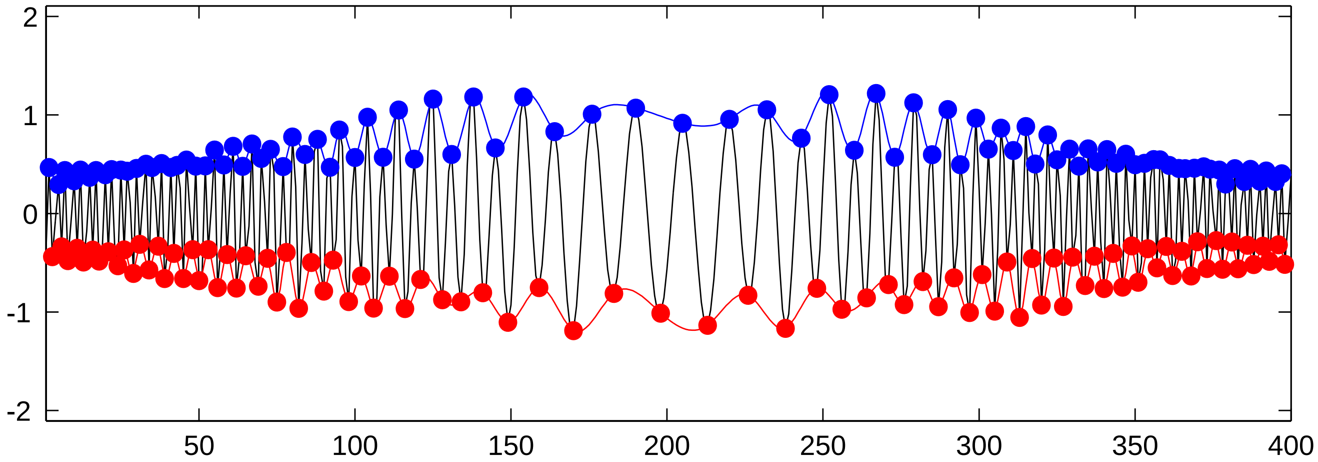


IMF 1; iteration 1

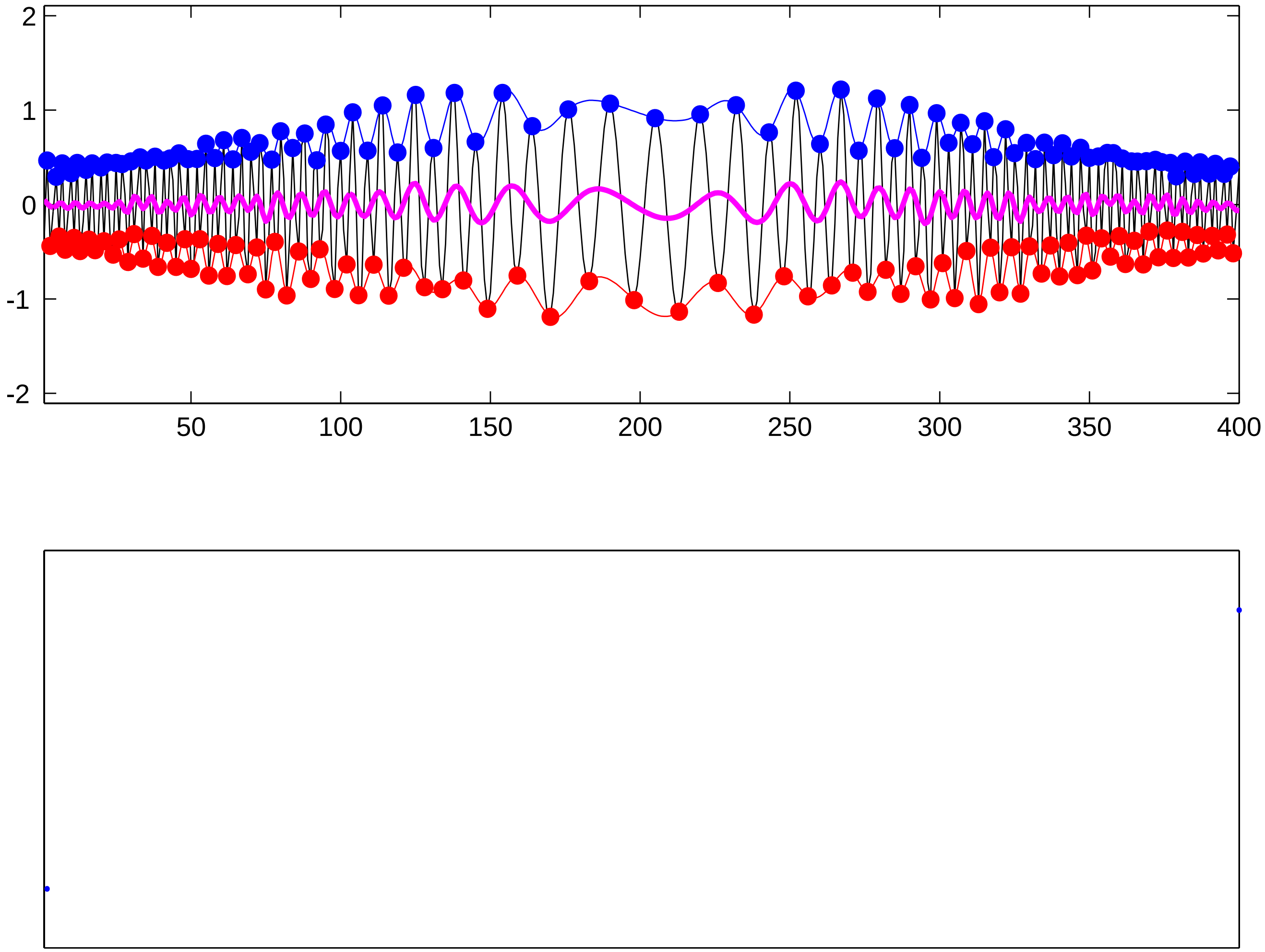



IMF 1;   iteration 1
2
1
0
-1
-2
50
100
150
200
250
300
350
400

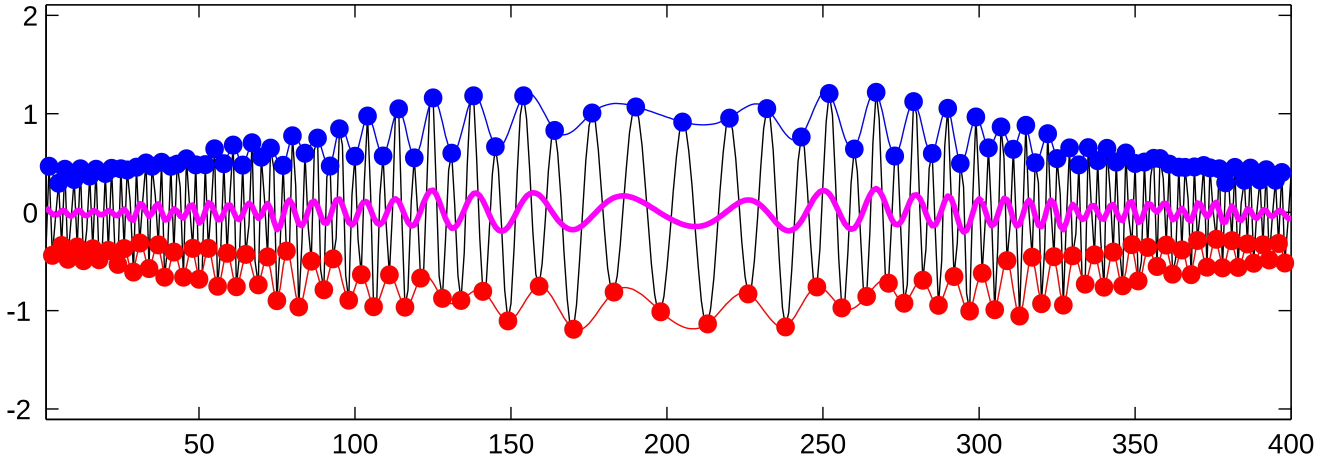
IMF 1;   iteration 1



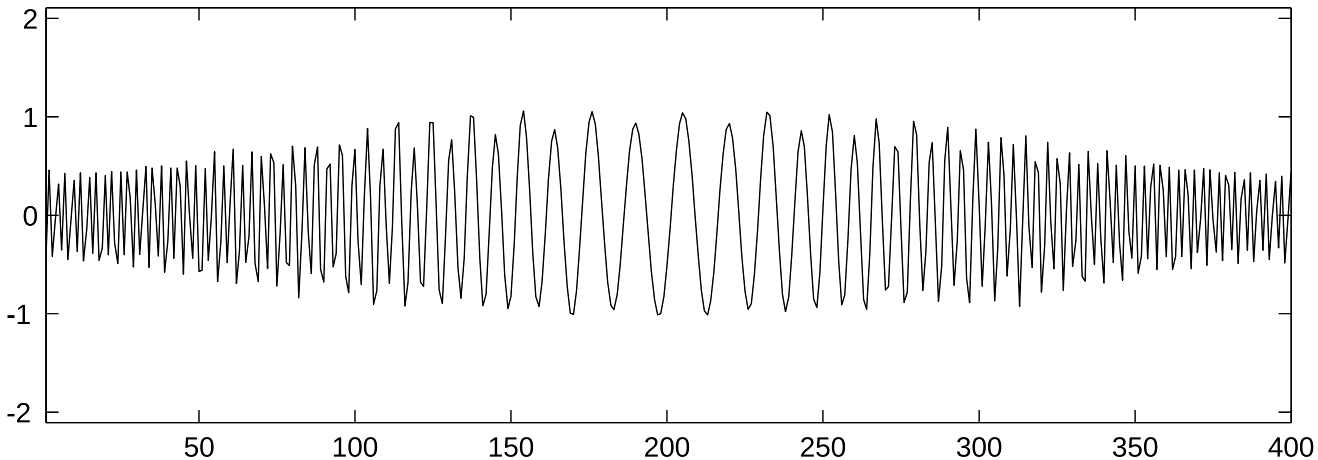
IMF 1;   iteration 1



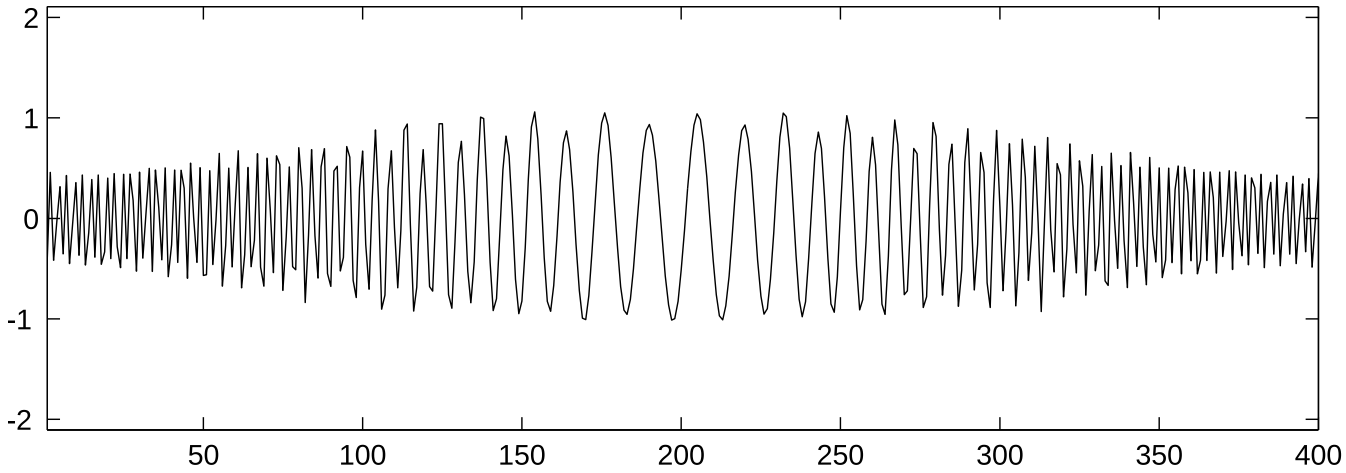
IMF 1; iteration 1

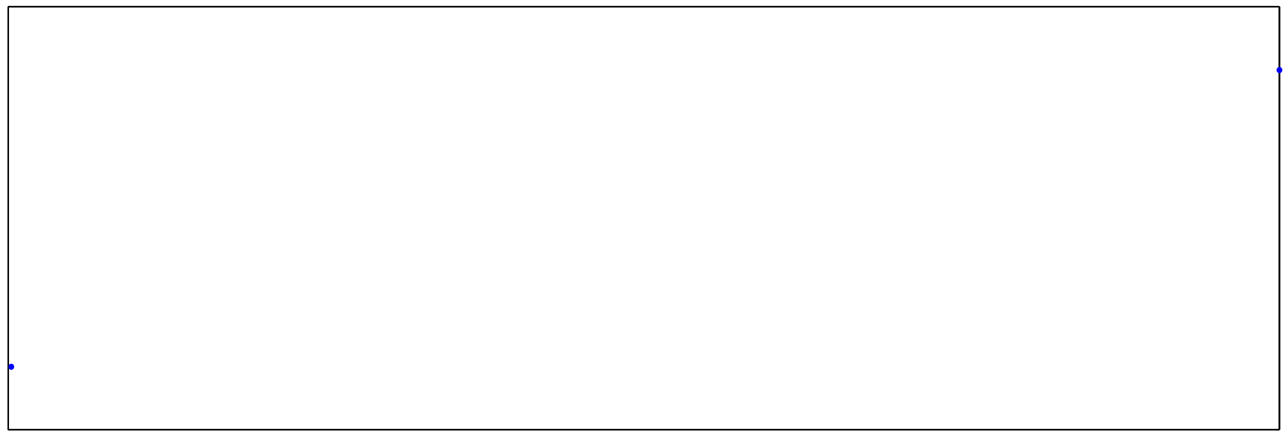


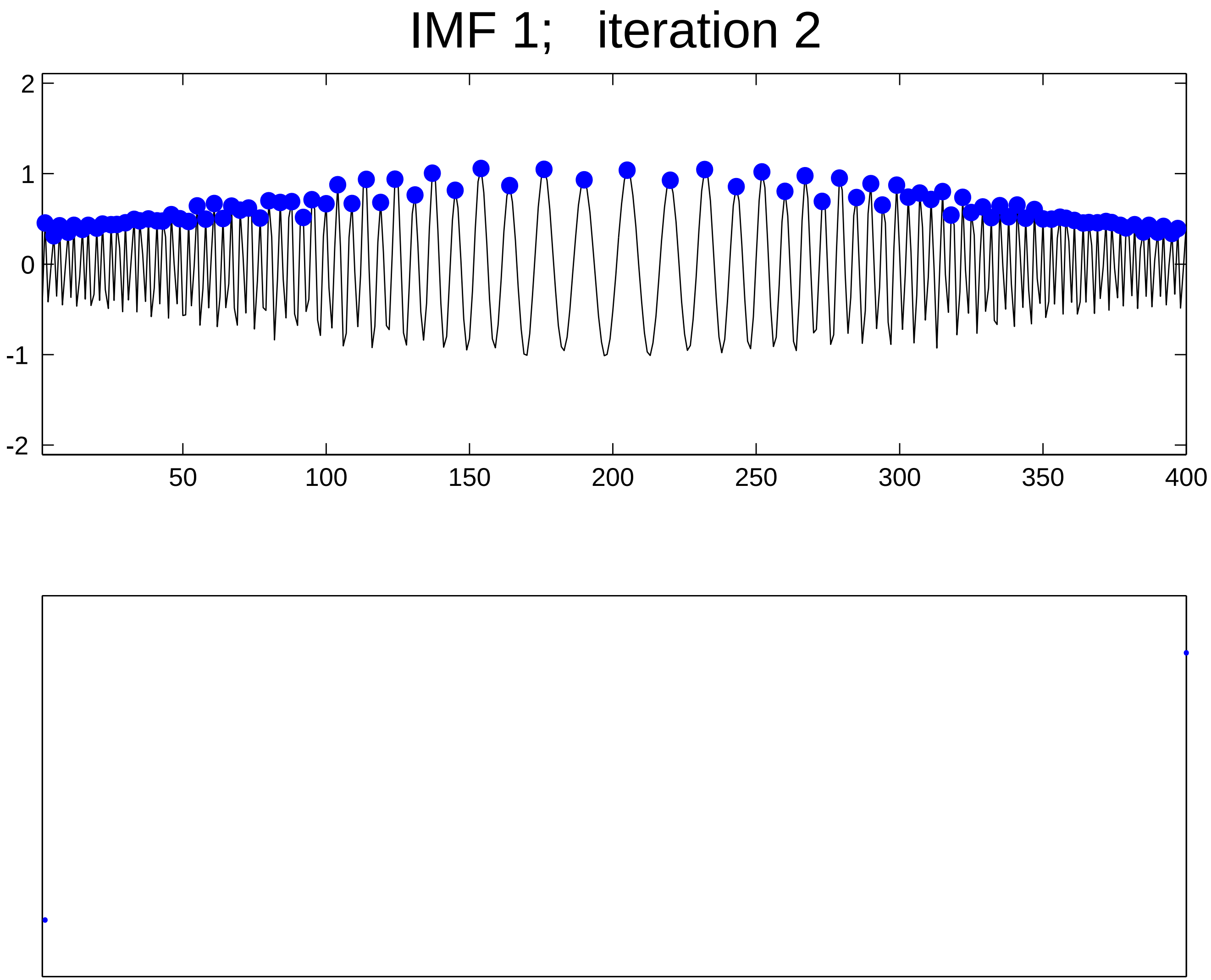
residue



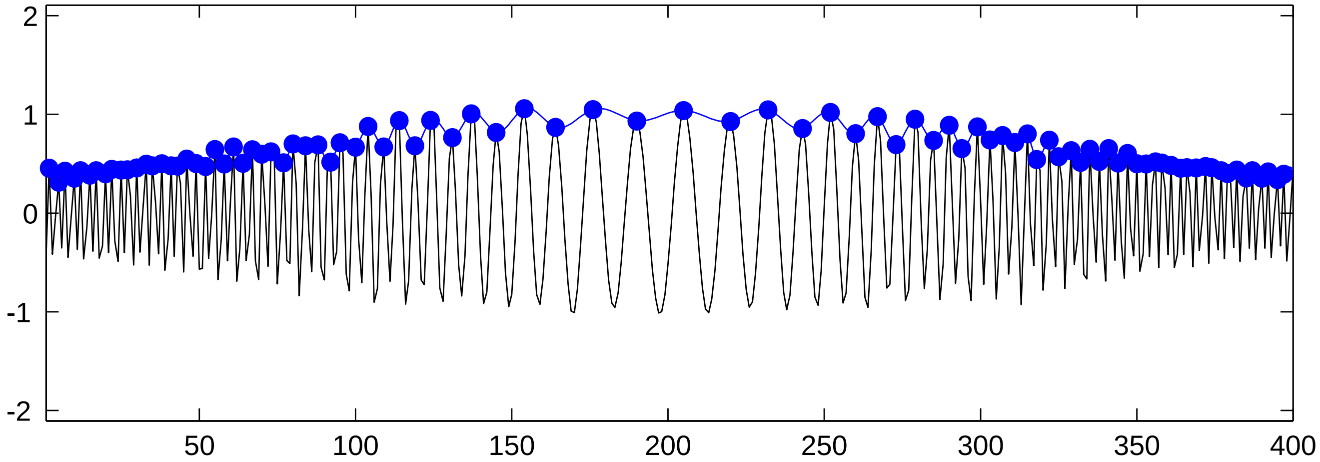
IMF 1;   iteration 2



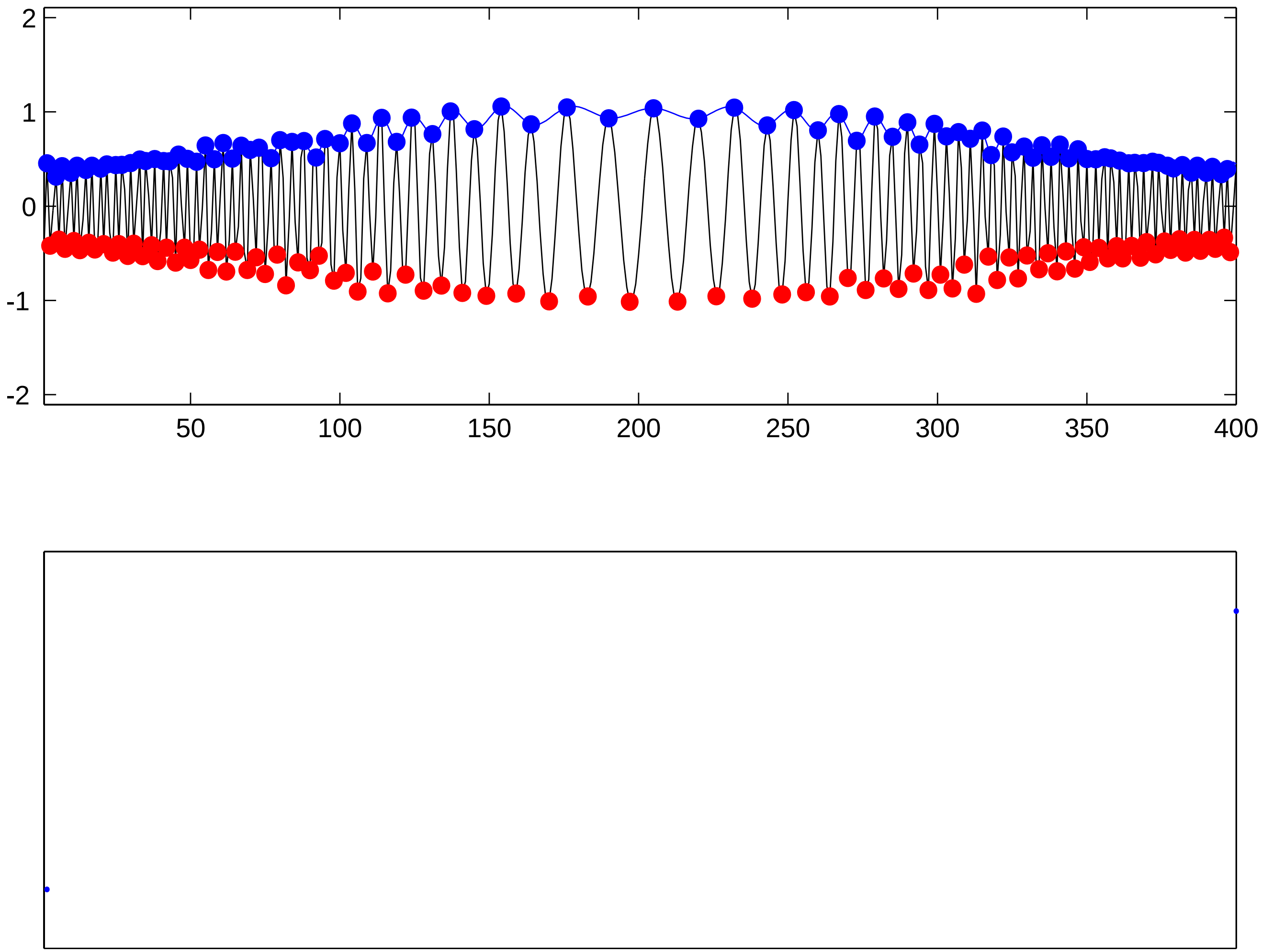



IMF 1; iteration 2
2
1
0
-1
-2
50
100
150
200
250
300
350
400

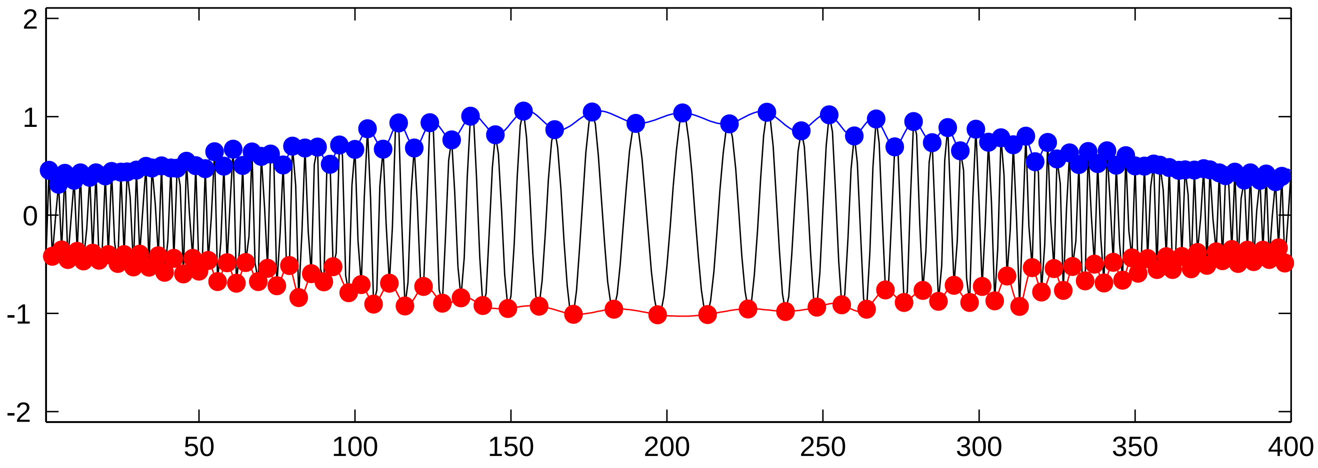
IMF 1; iteration 2



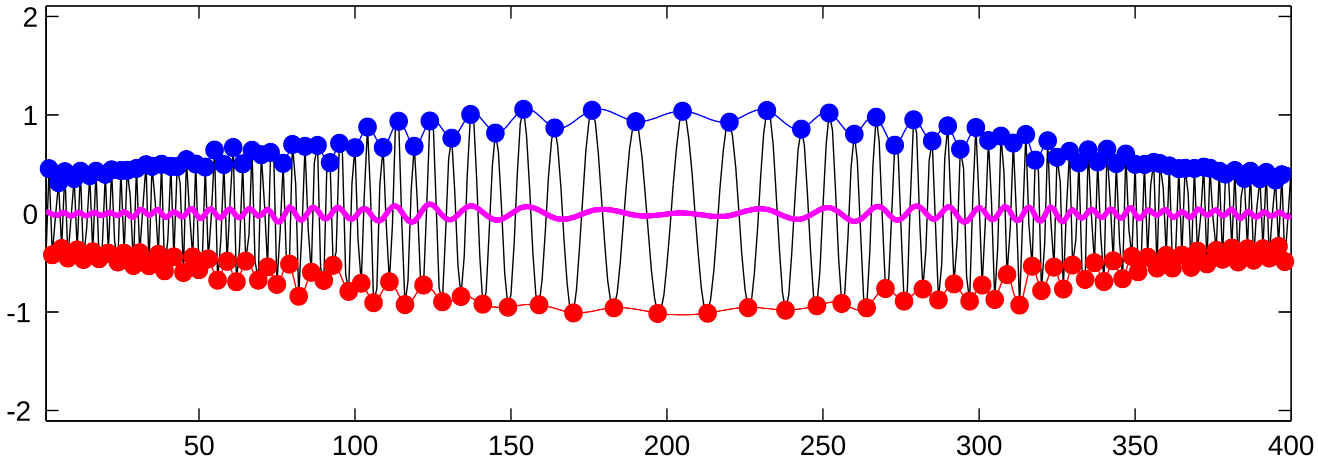
IMF 1;   iteration 2

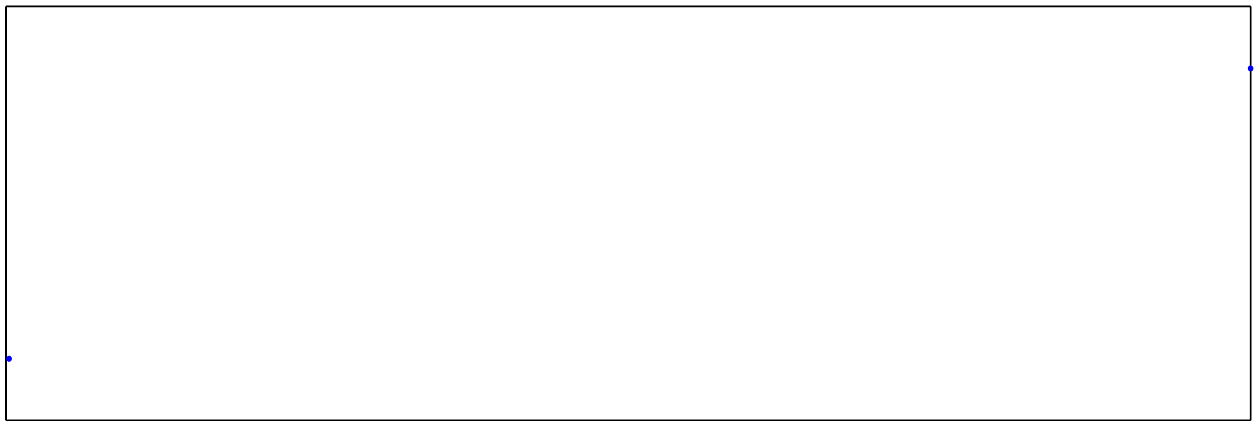


IMF 1; iteration 2

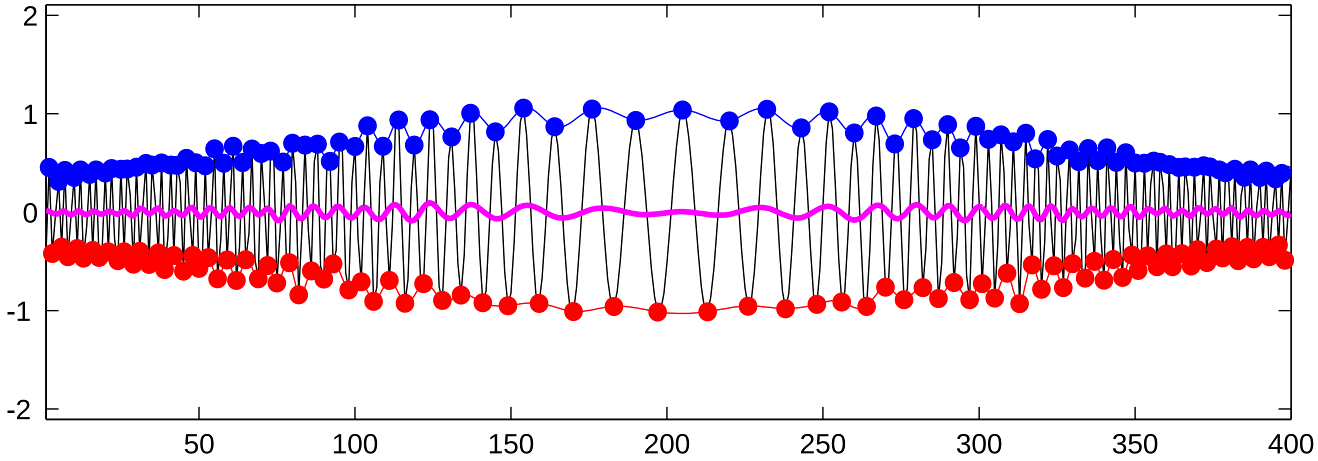


IMF 1;   iteration 2

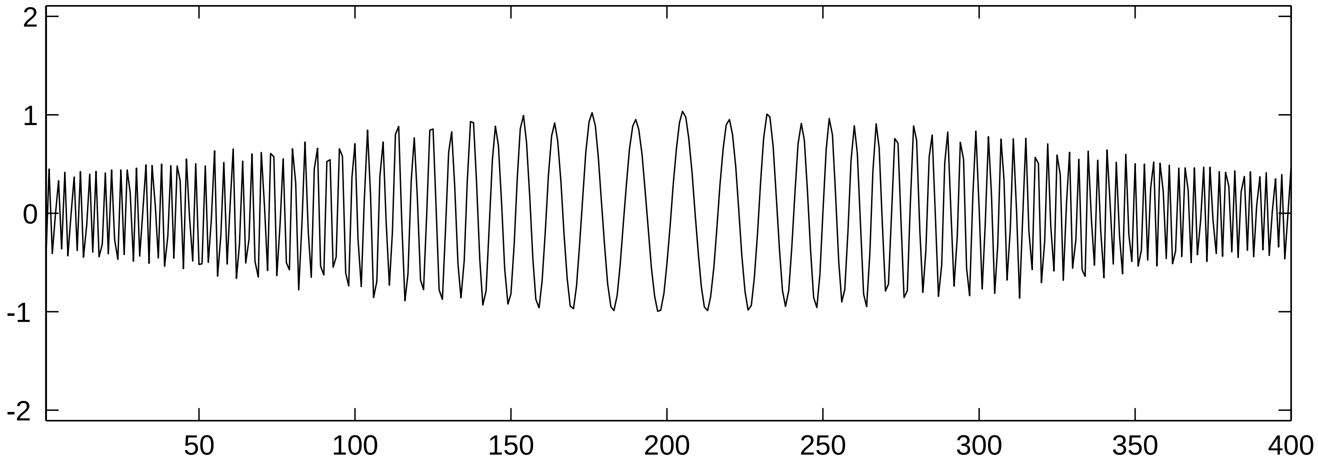




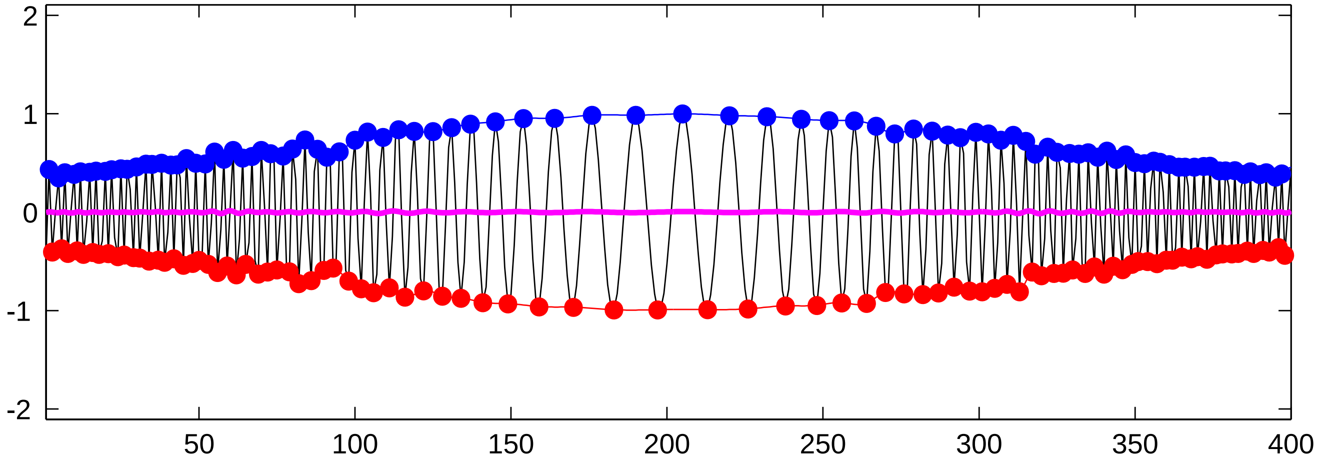
IMF 1;   iteration 2



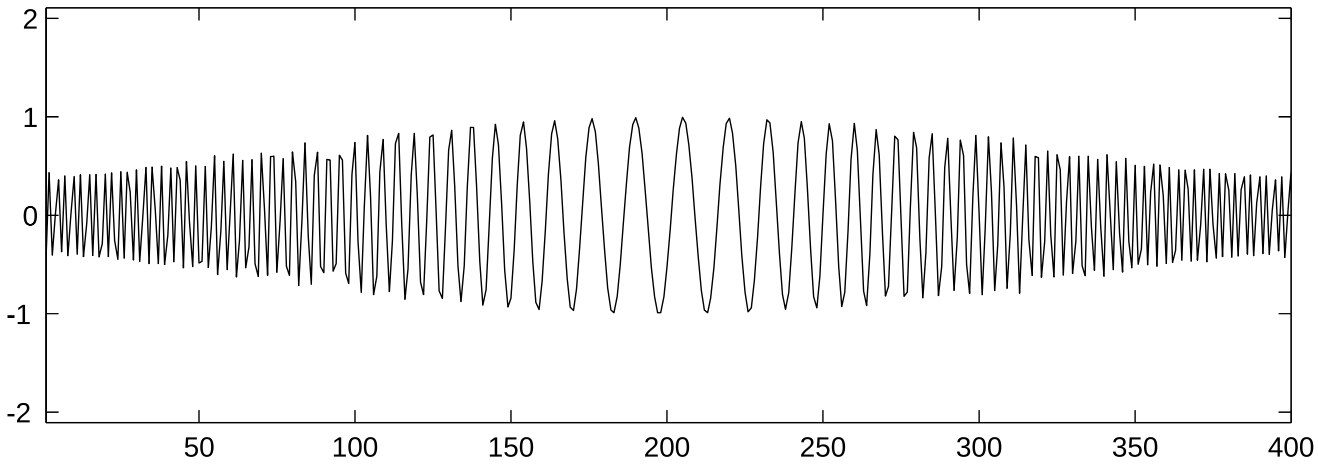
residue



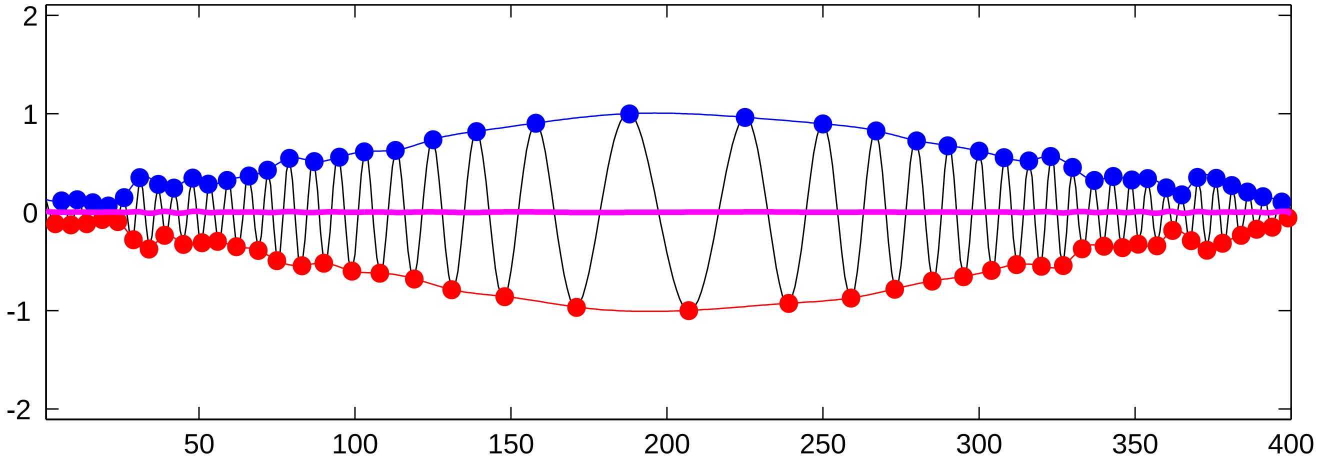
IMF 1;   iteration 5



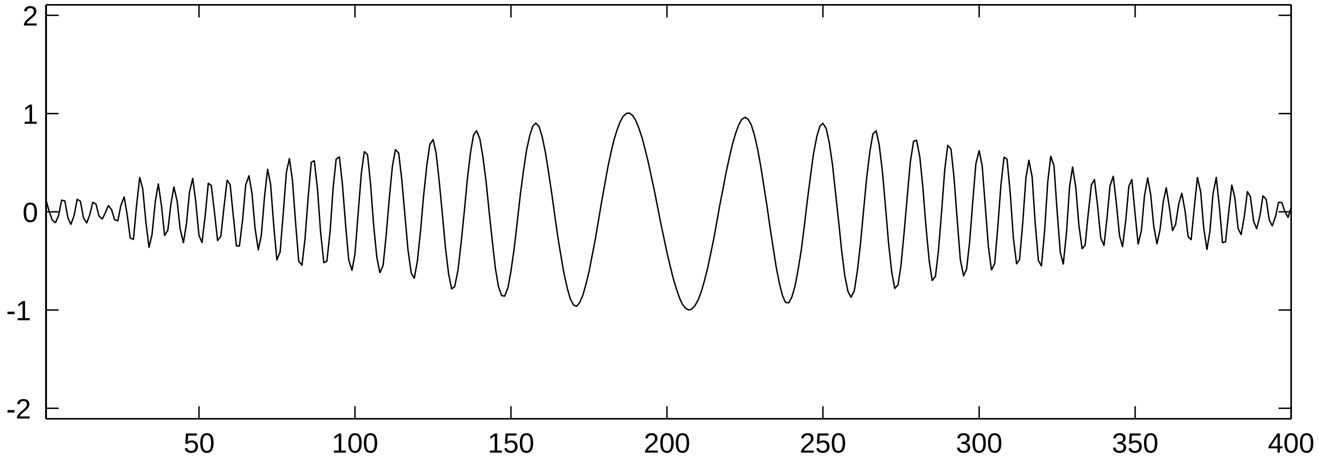
residue



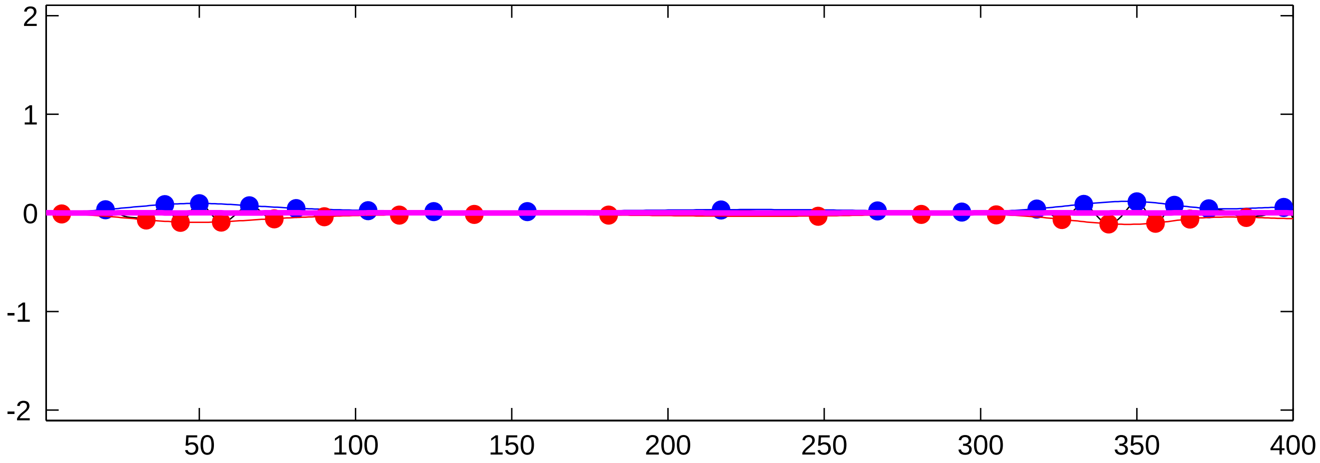
IMF 2;   iteration 2



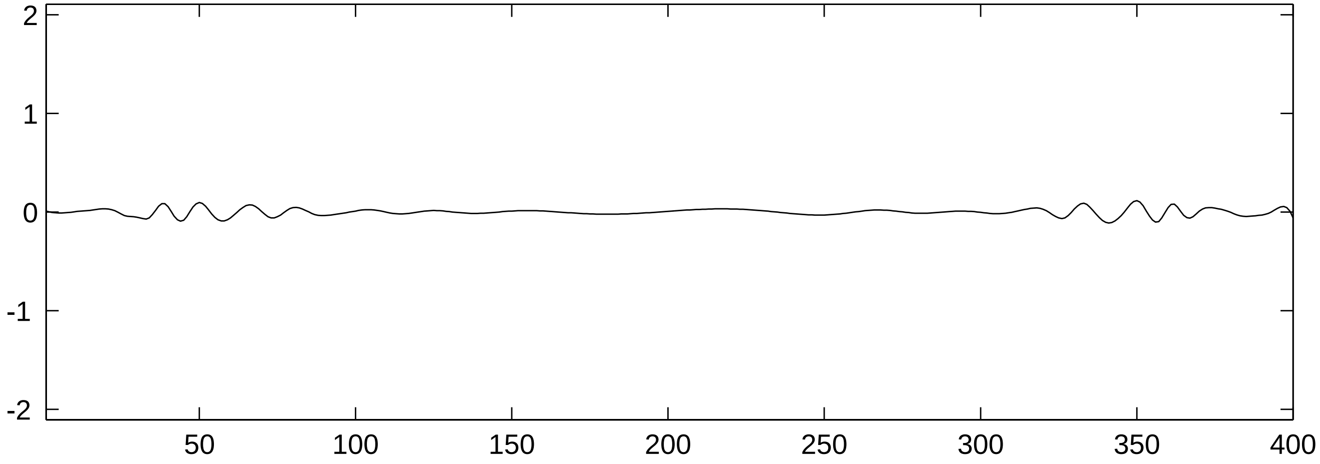
residue

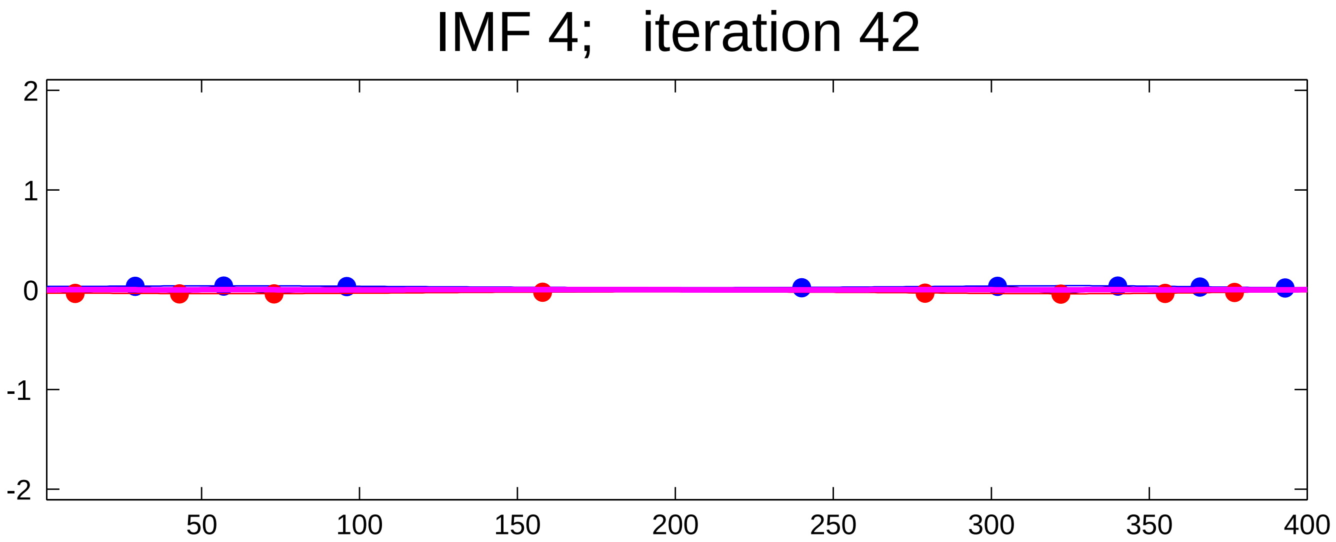


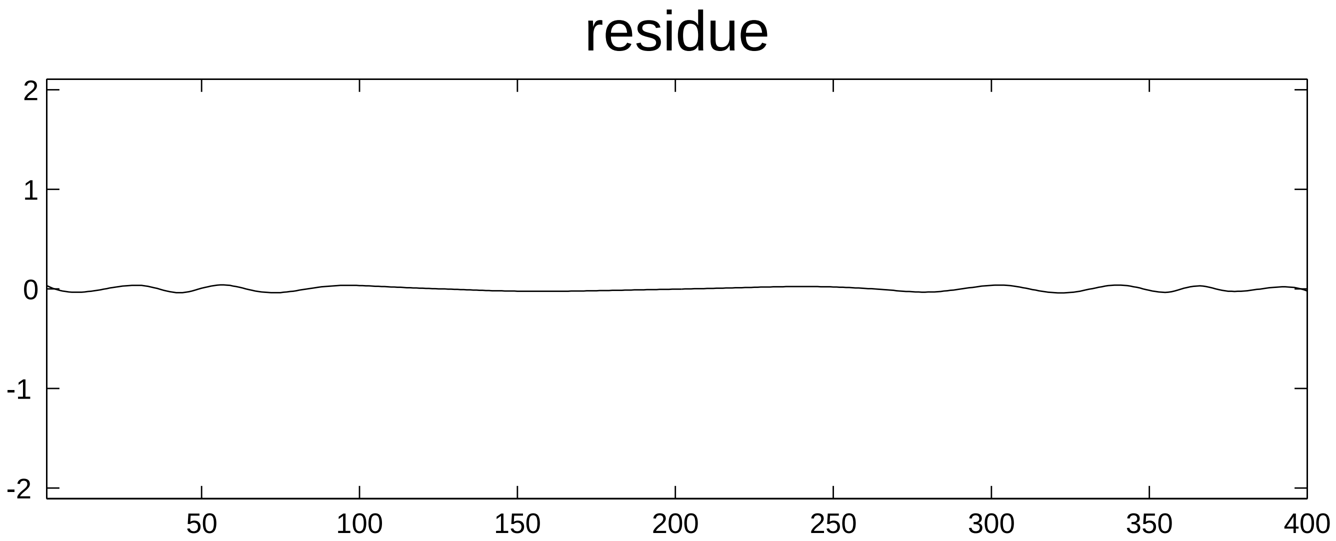
IMF 3; iteration 14

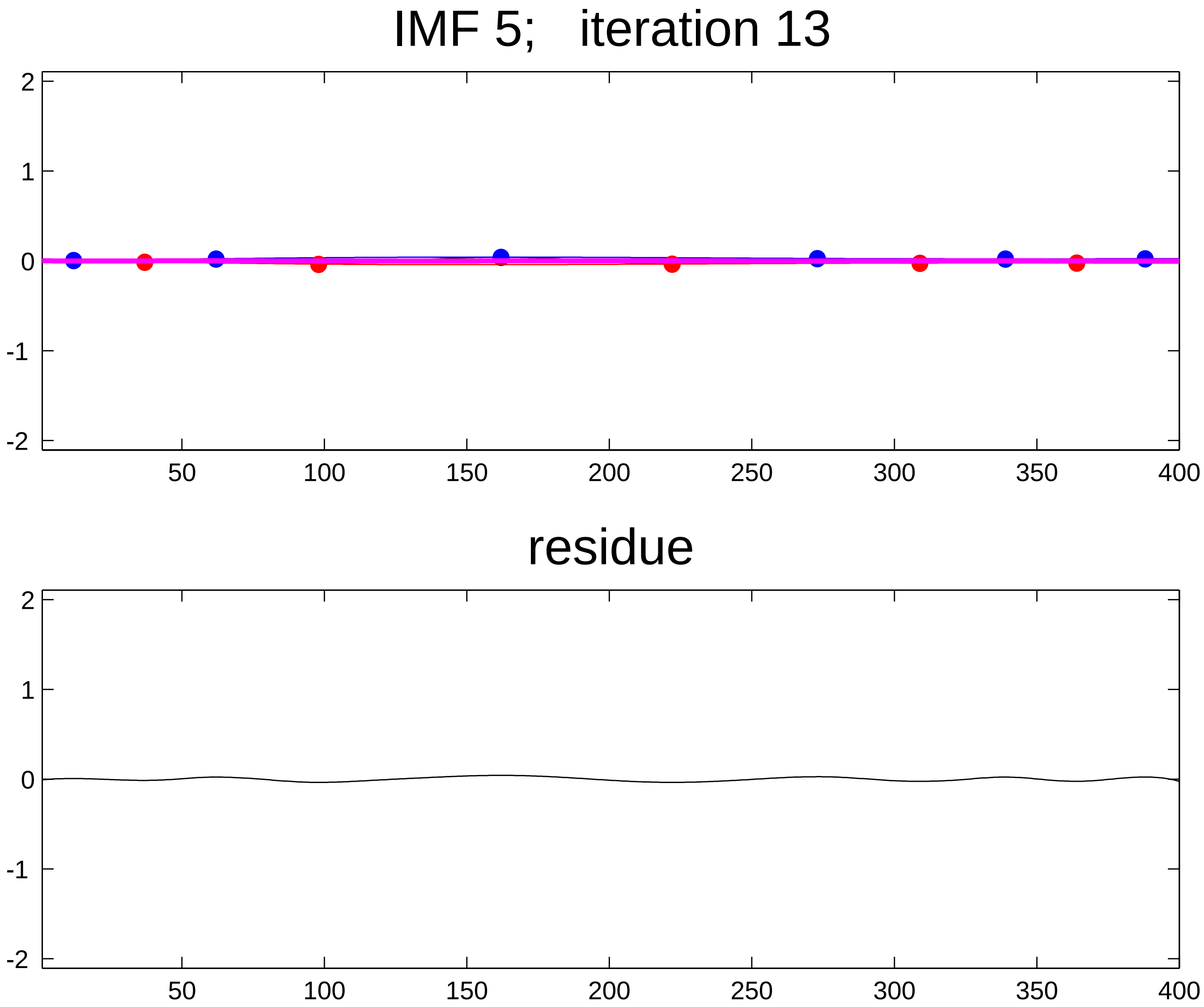


residue

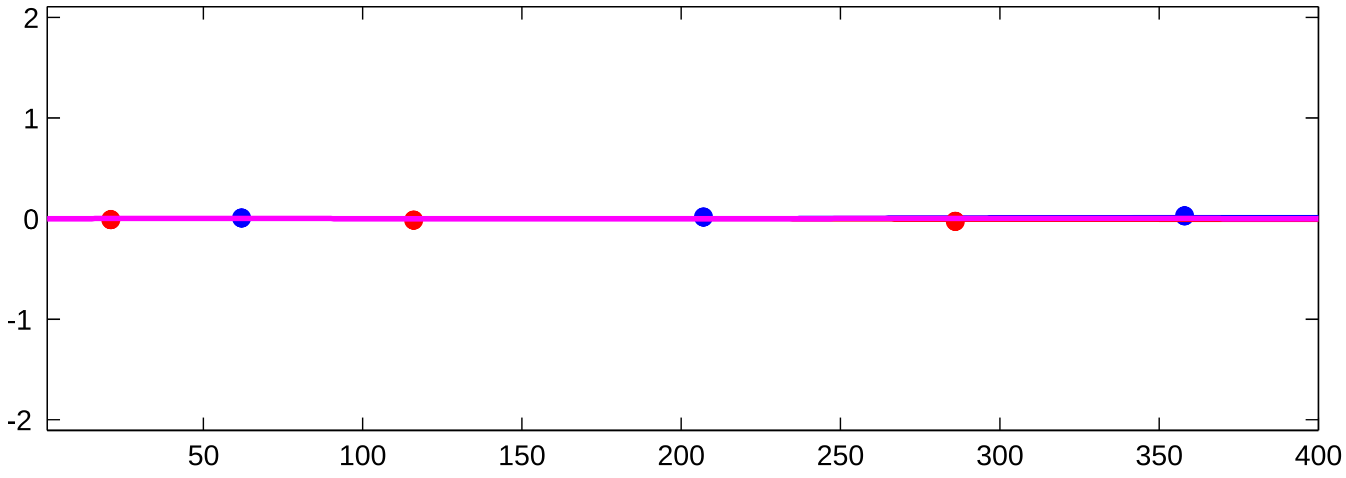



IMF 4; iteration 42
2
1
0
-1
-2
50
100
150
200
250
300
350
400

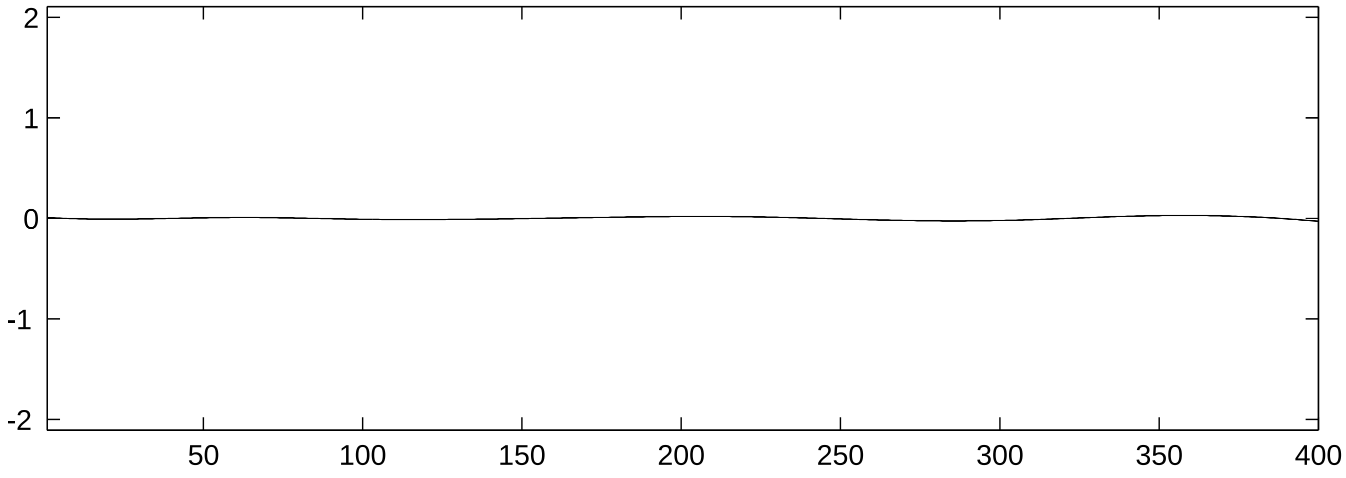

residue
2
1
0
-1
-2
50
100
150
200
250
300
350
400


IMF 5; iteration 13
2
1
0
-1
-2
50
100
150
200
250
300
350
400
residue
2
1
0
-1
-2
50
100
150
200
250
300
350
400

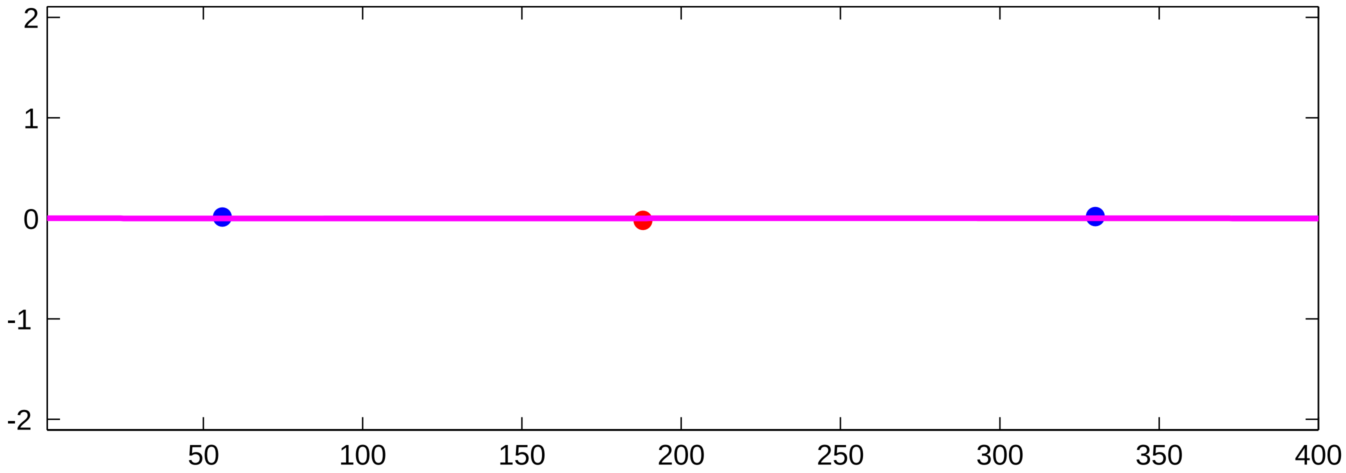
IMF 6;   iteration 8



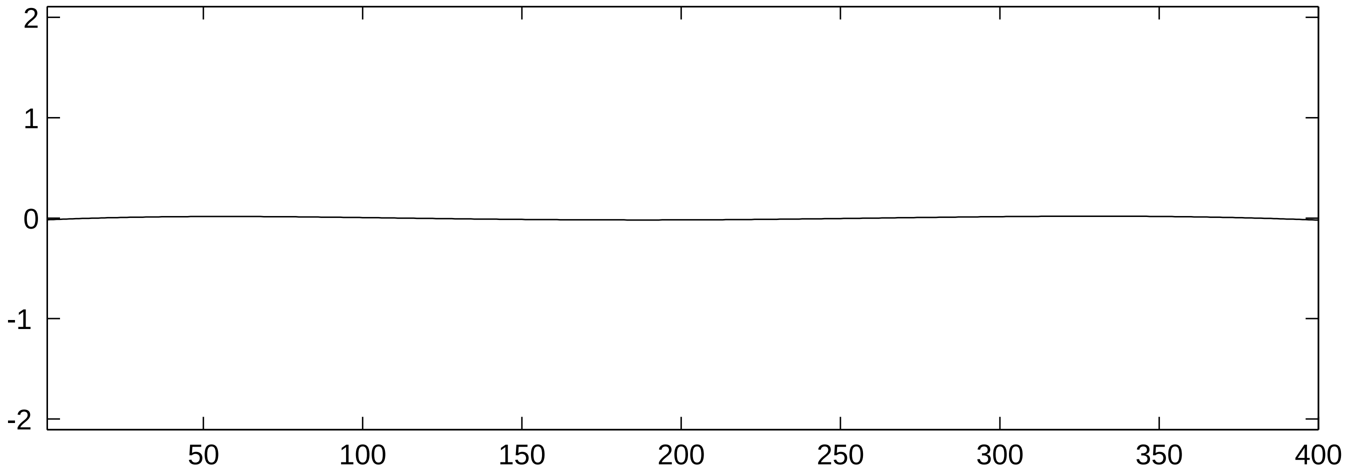
residue

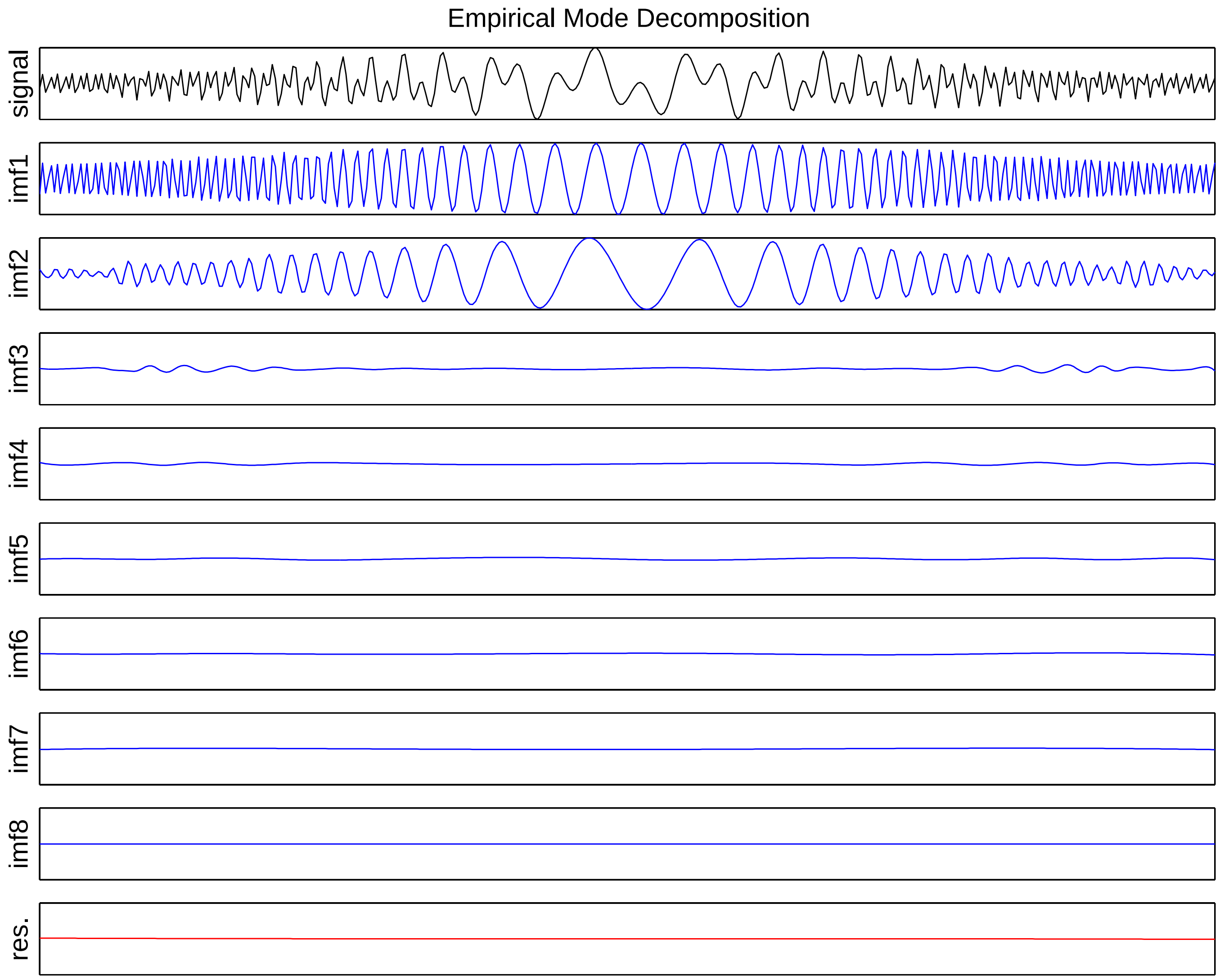


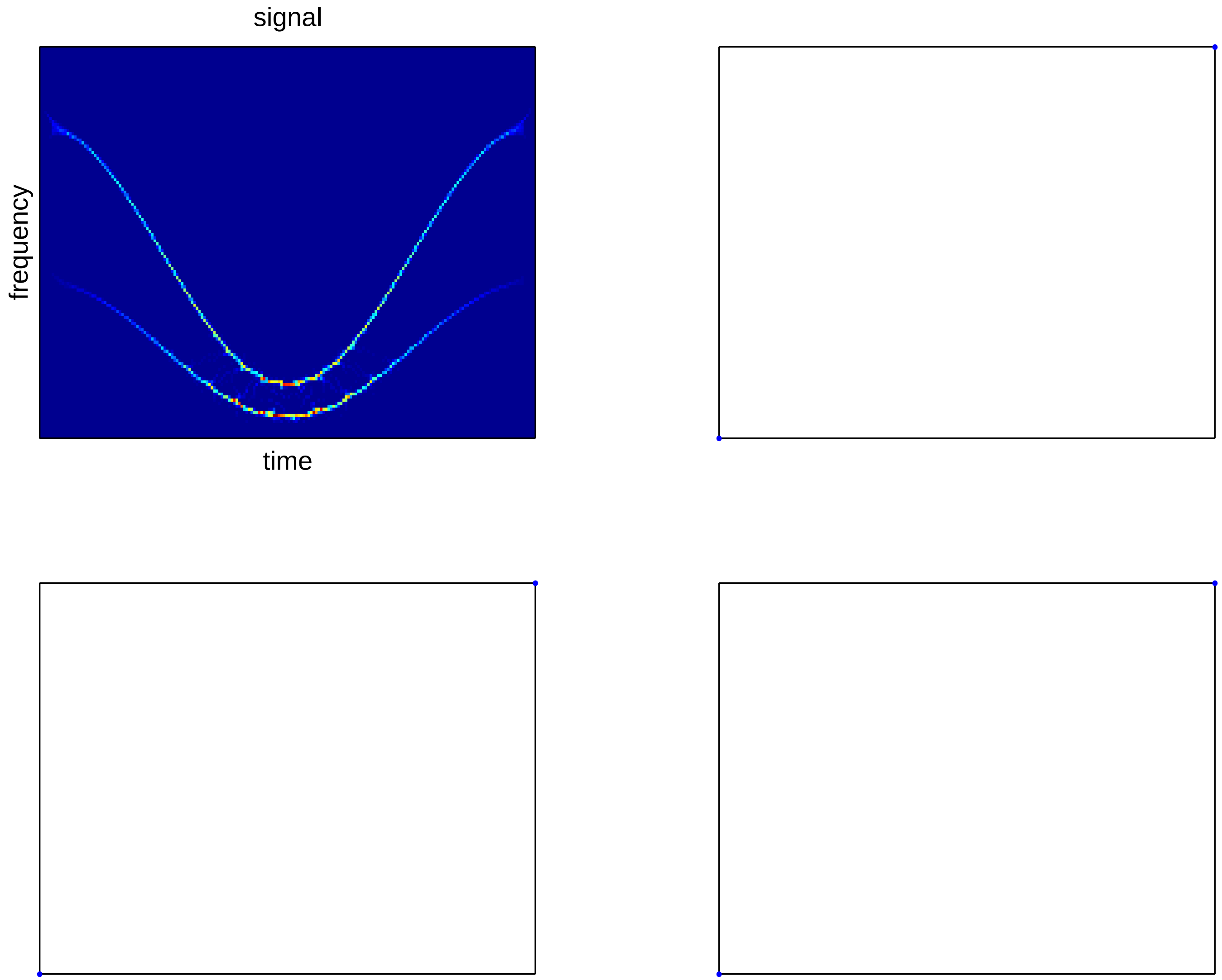
IMF 7;   iteration 21

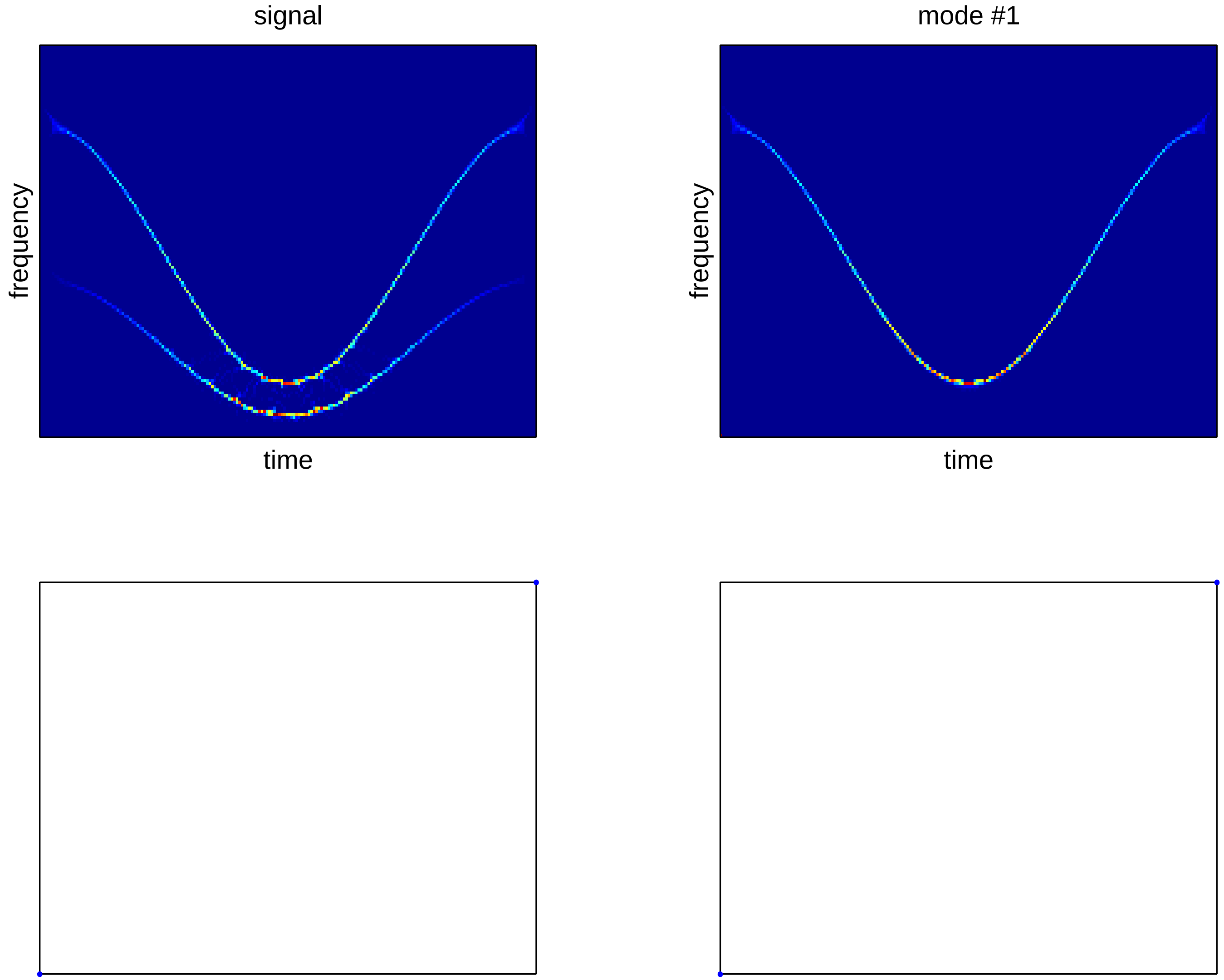


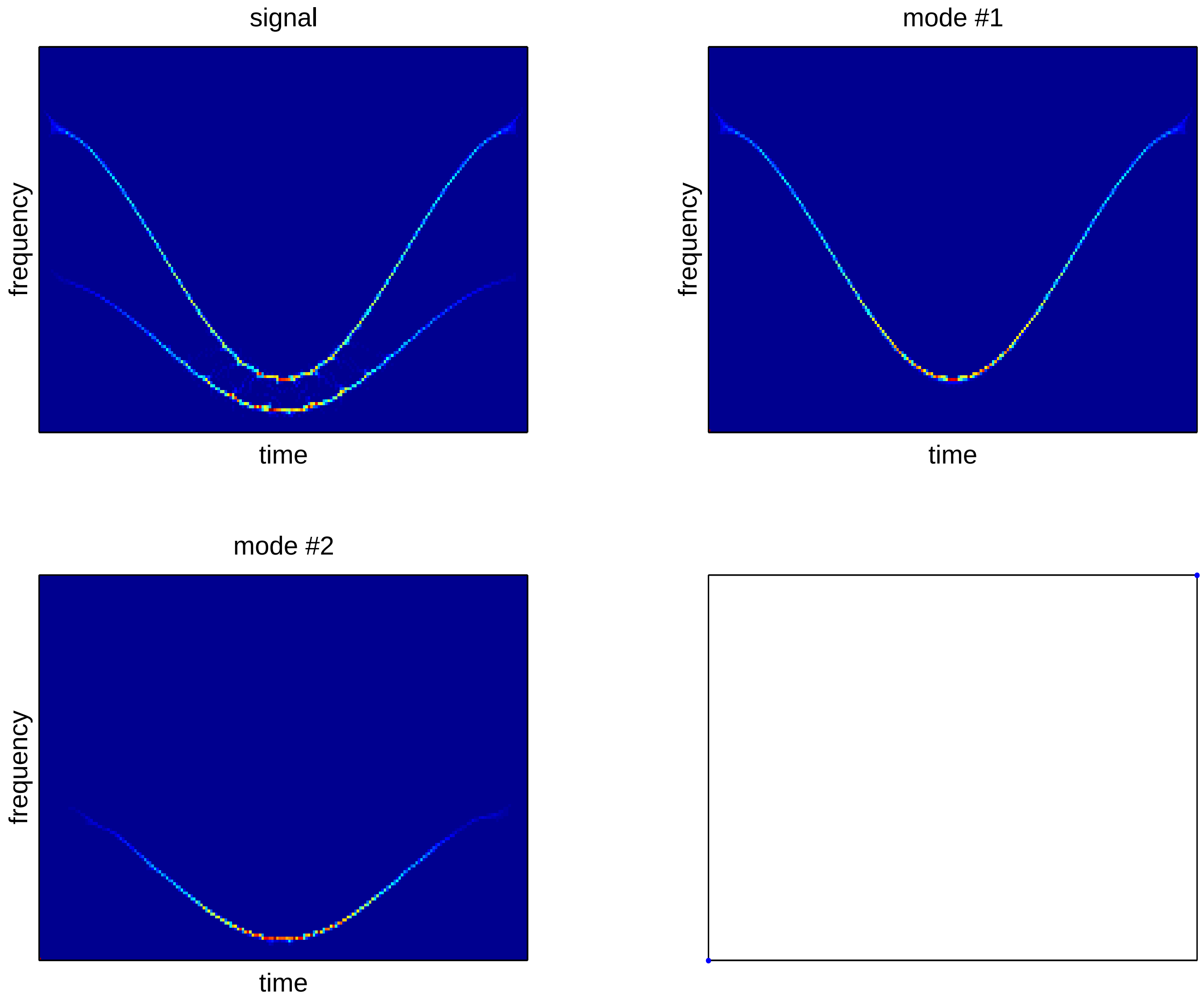
residue

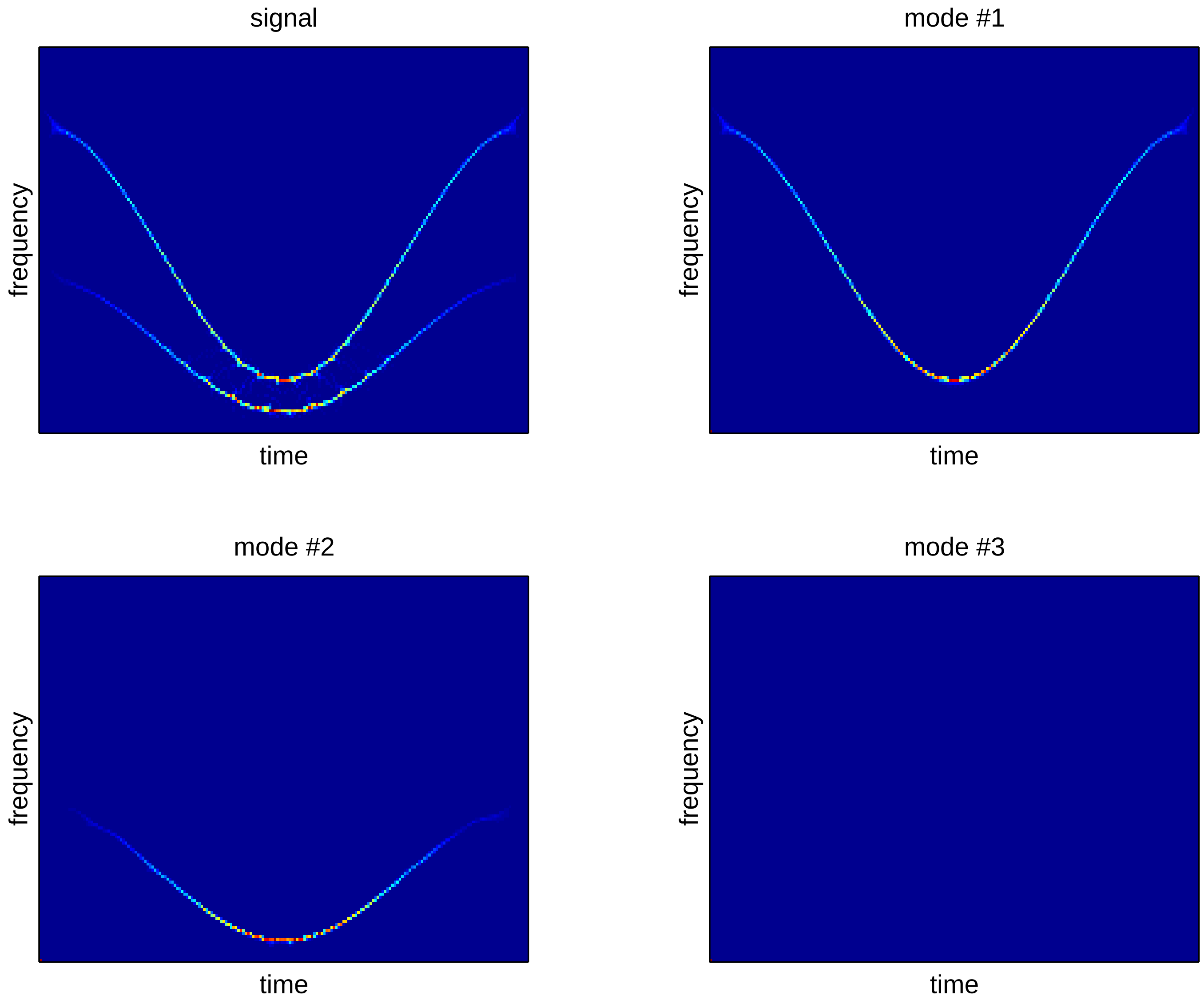



Empirical Mode Decomposition
signal
imf1
imf2
imf3
imf4
imf5
imf6
imf7
imf8
res.


signal
frequency
time


signal
frequency
time
mode #1
frequency
time


signal
mode #1
mode #2
frequency
time


signal
frequency
time
mode #1
frequency
time
mode #2
frequency
time
mode #3
frequency
time

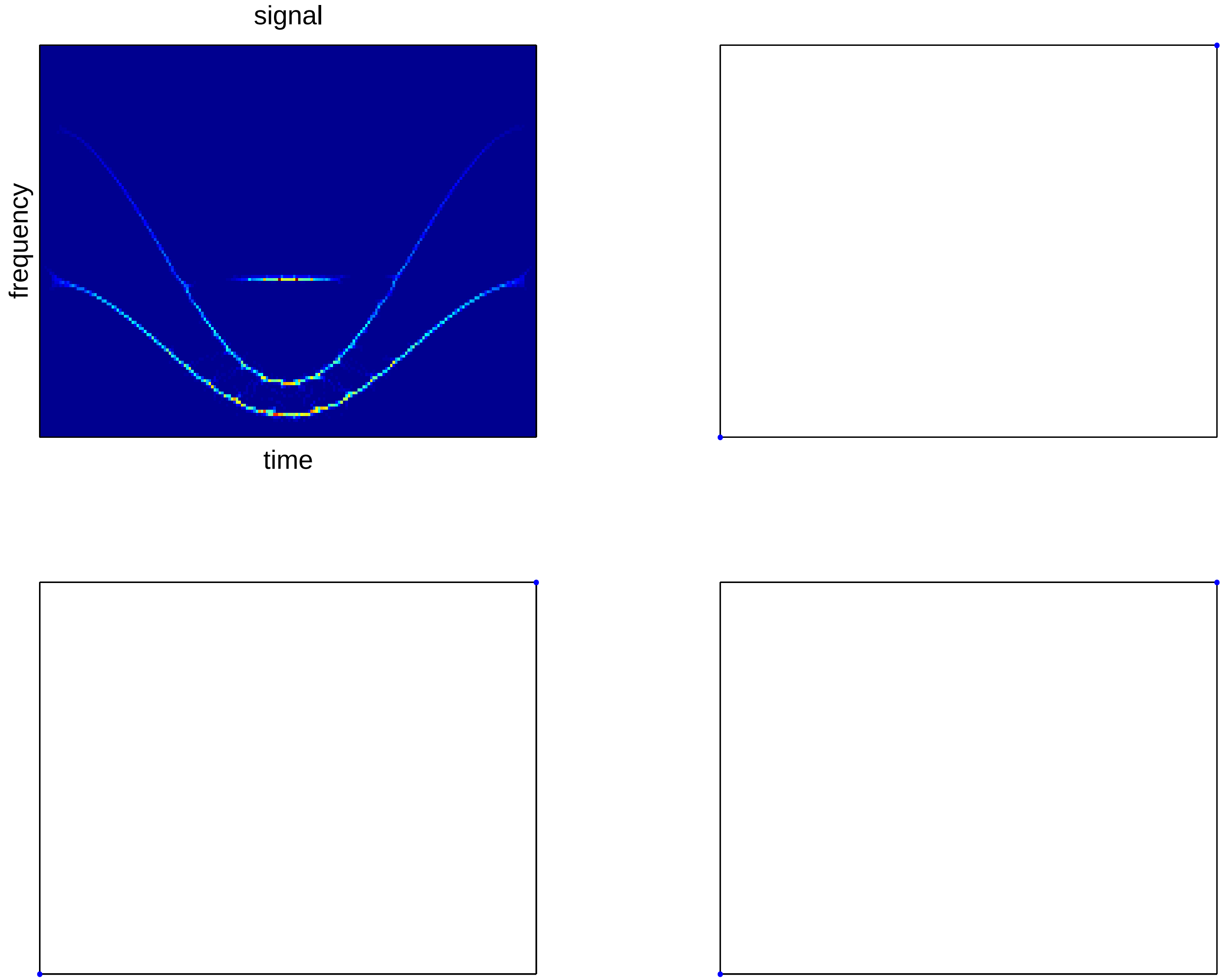
Decomposition

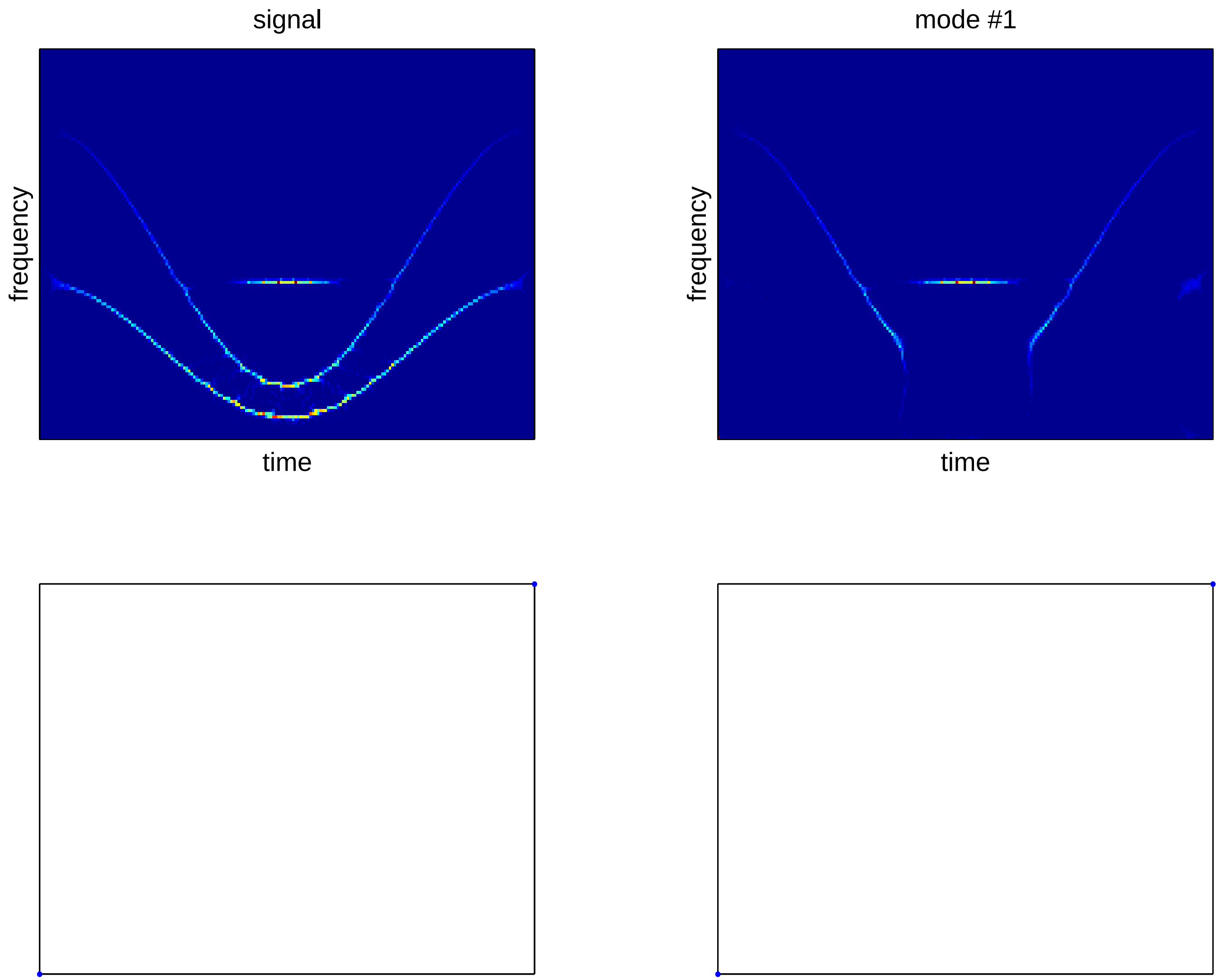
Iterative process — *Sequential* extraction, locally "fine to coarse" :

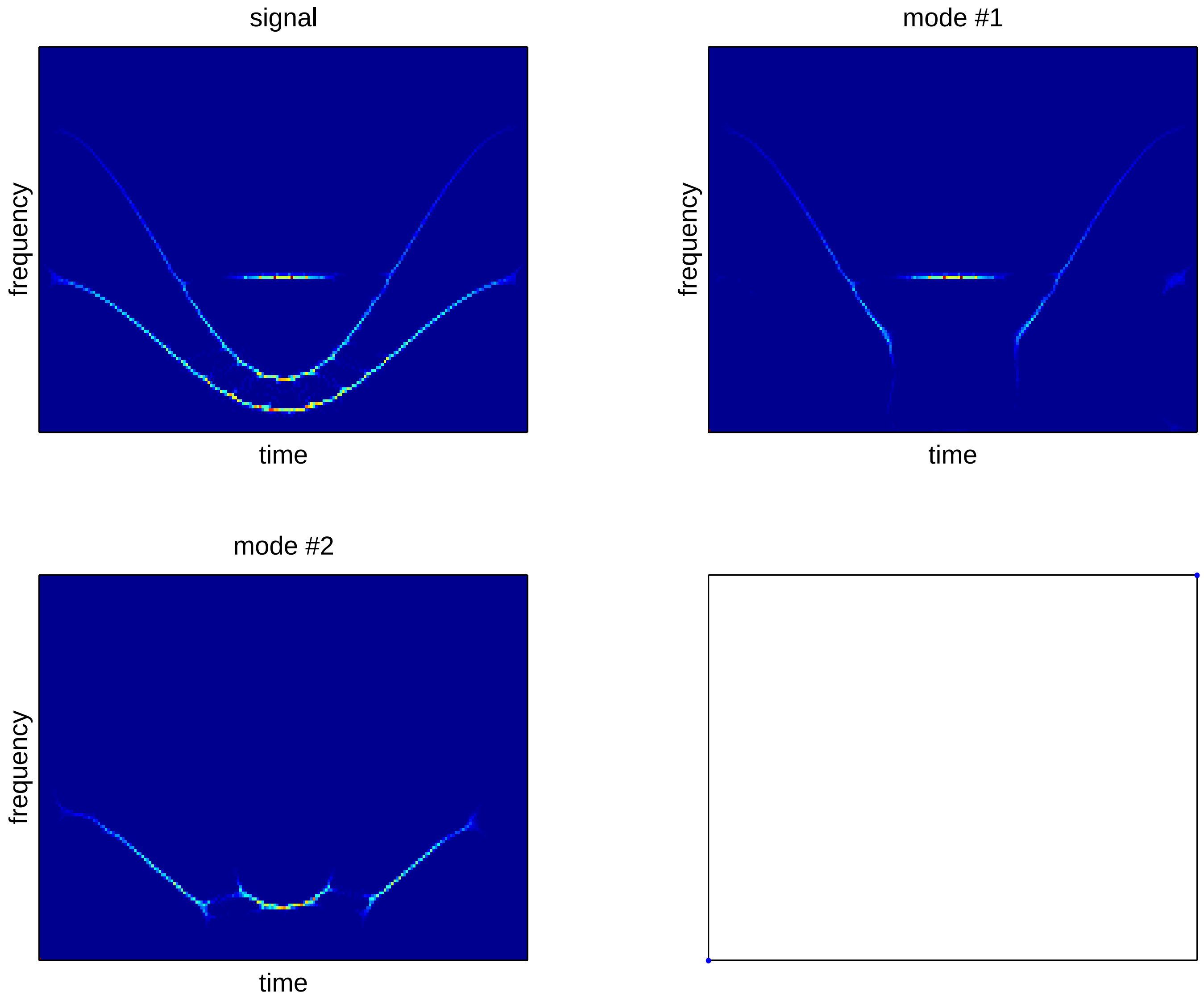
$$\begin{aligned} x(t) &= d_1(t) + m_1(t) \\ &= d_1(t) + d_2(t) + m_2(t) \\ &= \ldots \\ &= \sum_{k=1}^{K} d_k(t) + m_K(t). \end{aligned}$$

Modes — Automatic selection of *modes*—referred to as "Intrinsic Mode Functions" (IMF)—which are *zero-mean* and (wide sense) *AM-FM*.

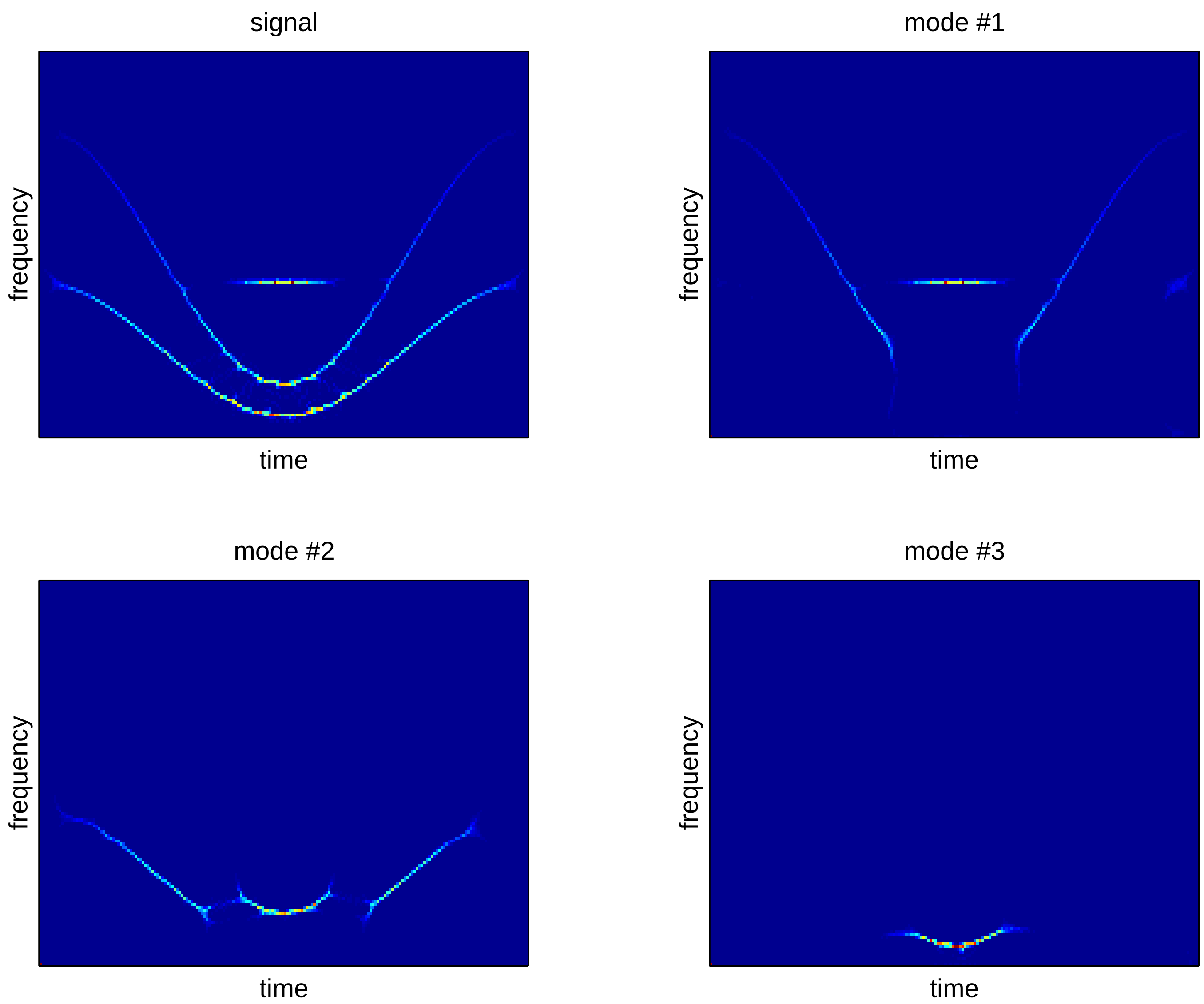
Implicit assumption — *Non-vanishing* modes: $a_k(t) > 0$.


signal
frequency
time


signal
mode #1
frequency
time
frequency
time



signal
mode #1
frequency
time
frequency
time
mode #2
frequency
time


signal
frequency
time
mode #1
frequency
time
mode #2
frequency
time
mode #3
frequency
time

Some features

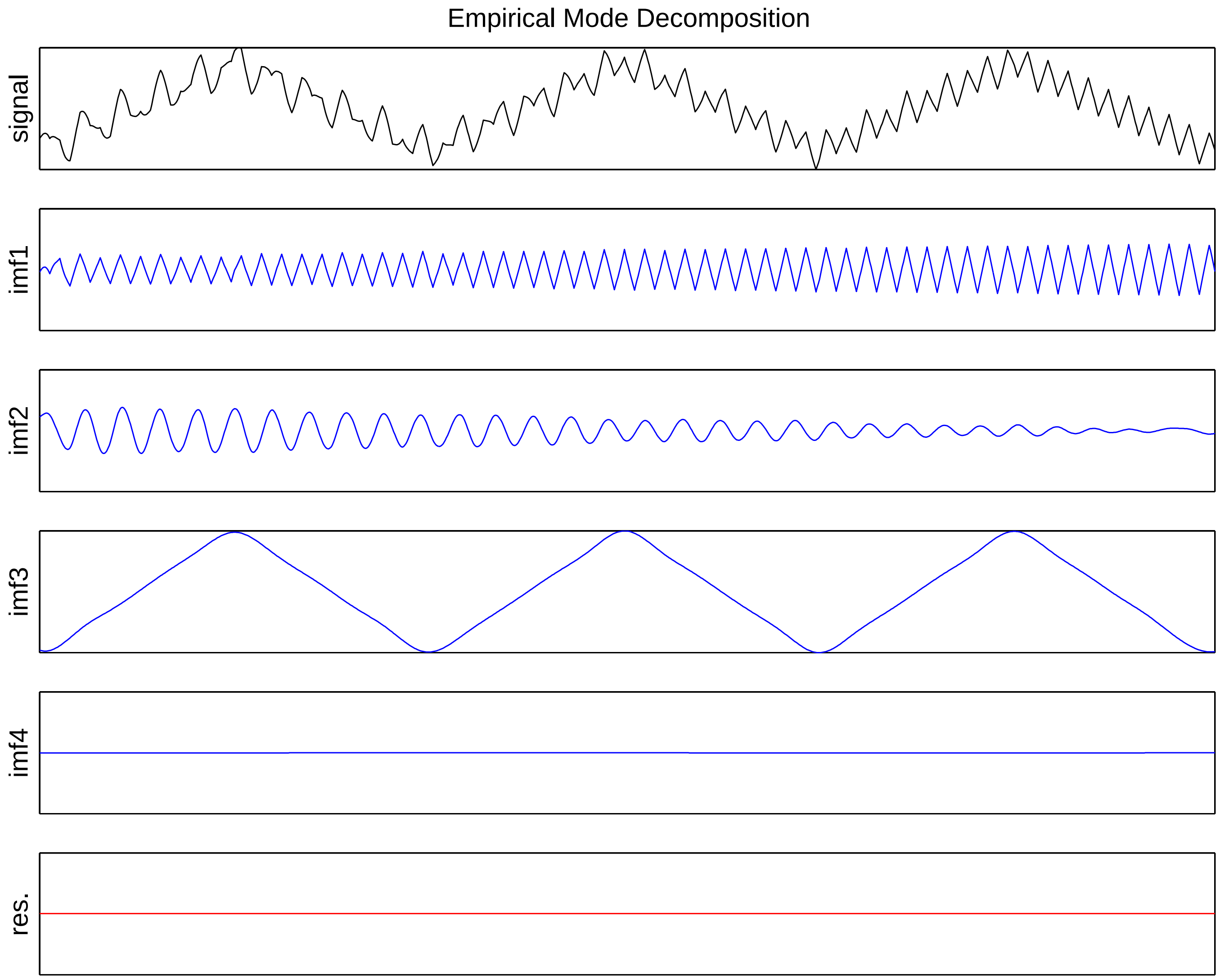
No analytic definition — The decomposition is only defined as the output of an algorithm $\Rightarrow$ performance evaluation via extensive *numerical simulations* in well-controlled situations.

Locality — The method operates at the scale of *one* oscillation.

Adaptivity — The decomposition is fully *data-driven*.

Multiresolution — The iterative process explores *sequentially* the "natural" constitutive scales of a signal.

Oscillations of any type — No assumption on the *harmonic* nature of oscillations $\Rightarrow$ 1 *nonlinear* oscillation $=$ 1 mode.


Empirical Mode Decomposition
signal
imf1
imf2
imf3
imf4
res.

EMD as a filter — 1.

A stochastic approach in the frequency domain — Decomposition and spectrum analysis, mode by mode, of a wideband noise.

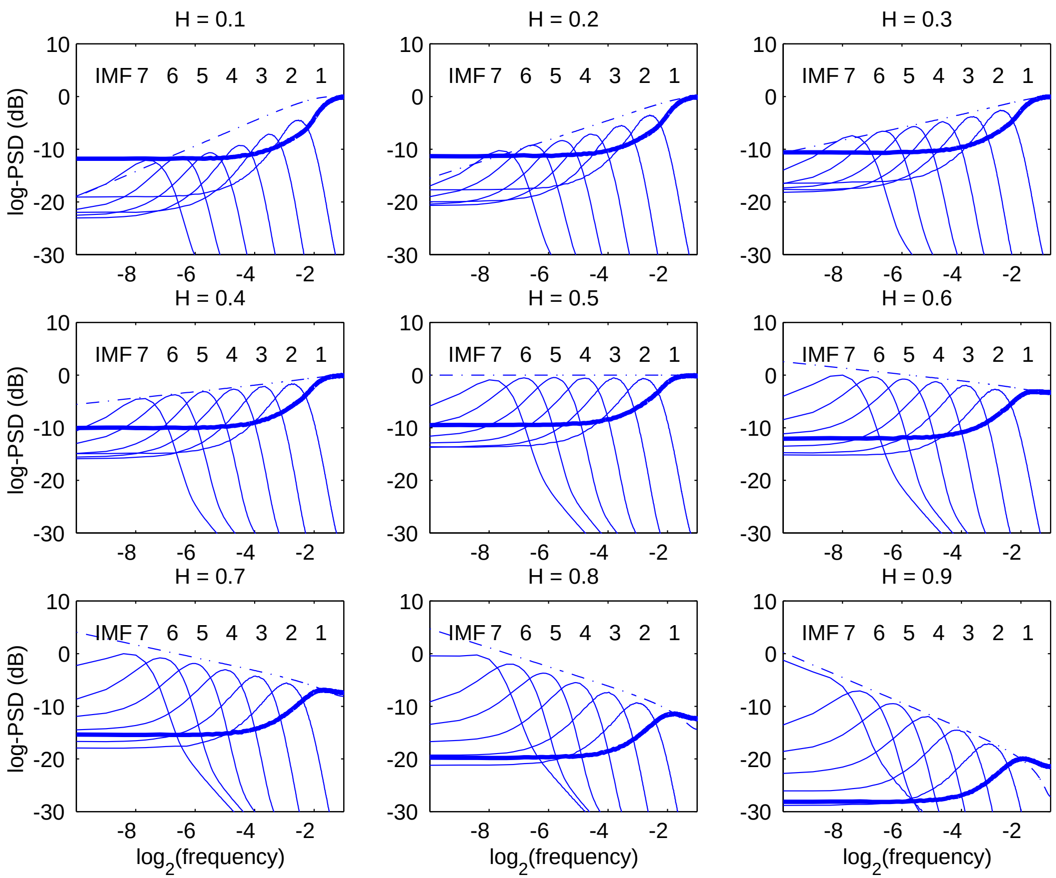
Model — Fractional Gaussian noise (fGn), of spectrum density $\mathcal{S}(f) \sim |f|^{1-2H}$, with $0 < H < 1$ (Hurst exponent).

Results — Ensemble average $\Rightarrow$ "spontaneous" emergence of a filter bank structure, almost dyadic and self-similar (F., Gonçalvès & Rilling, '03):

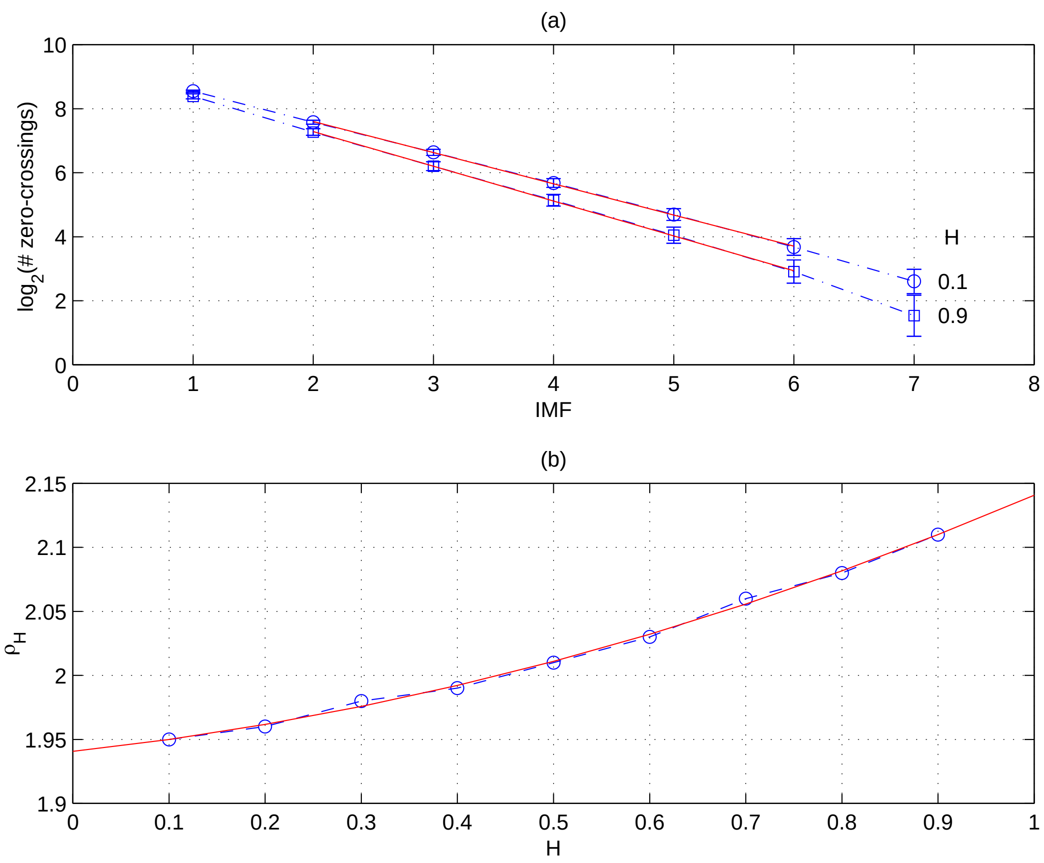
$$\mathcal{S}_{k',H}(f) = \rho_H^{\alpha(k'-k)} \mathcal{S}_{k,H}(\rho_H^{k'-k} f)$$

for any $k' > k \geq 2$, with $\alpha = 2H - 1$ and $\rho_H \approx 2$.

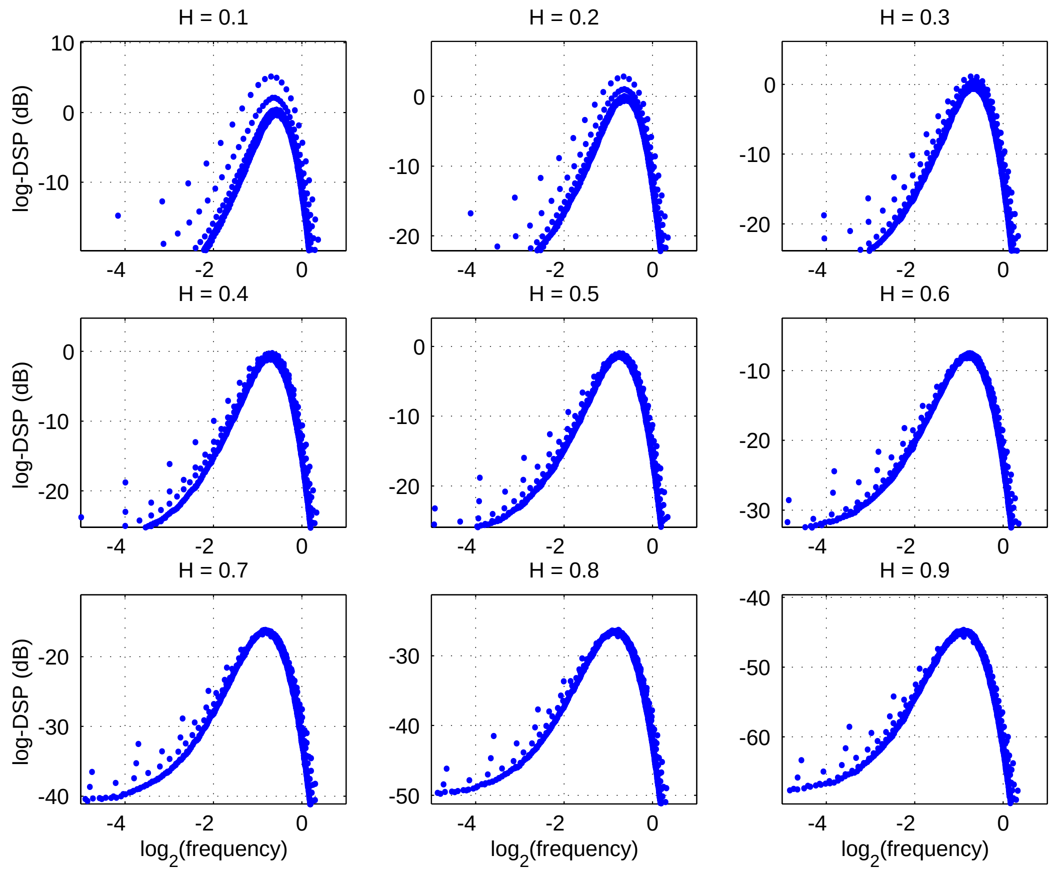
EMD equivalent filter banks



IMFs zero-crossings



Renormalized EMD spectra



EMD as a filter — 2.

A deterministic approach in the time domain — Decomposition, mode by mode, of a Dirac impulse.

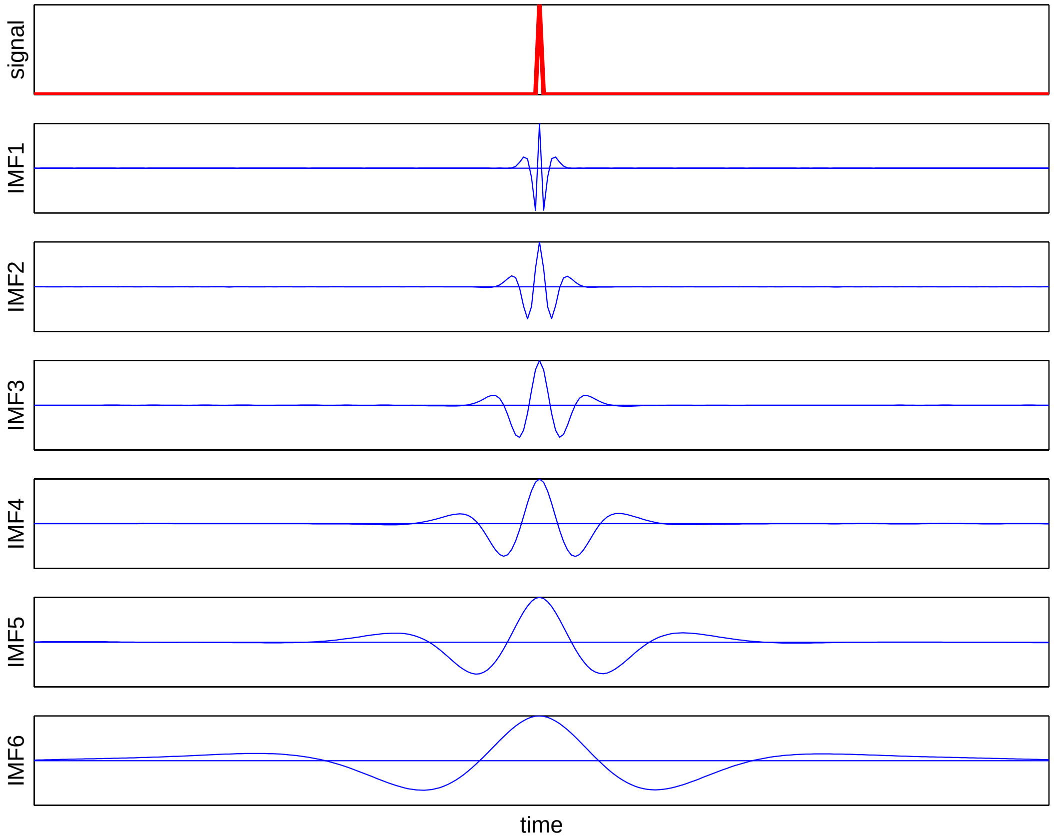
Model — Multiple extrema required $\Rightarrow$ small amount of *additive noise*.

Results — Ensemble average $\Rightarrow$ "spontaneous" emergence of a filter bank structure, almost dyadic and self-similar (F. & Gonçalvès, '03) :

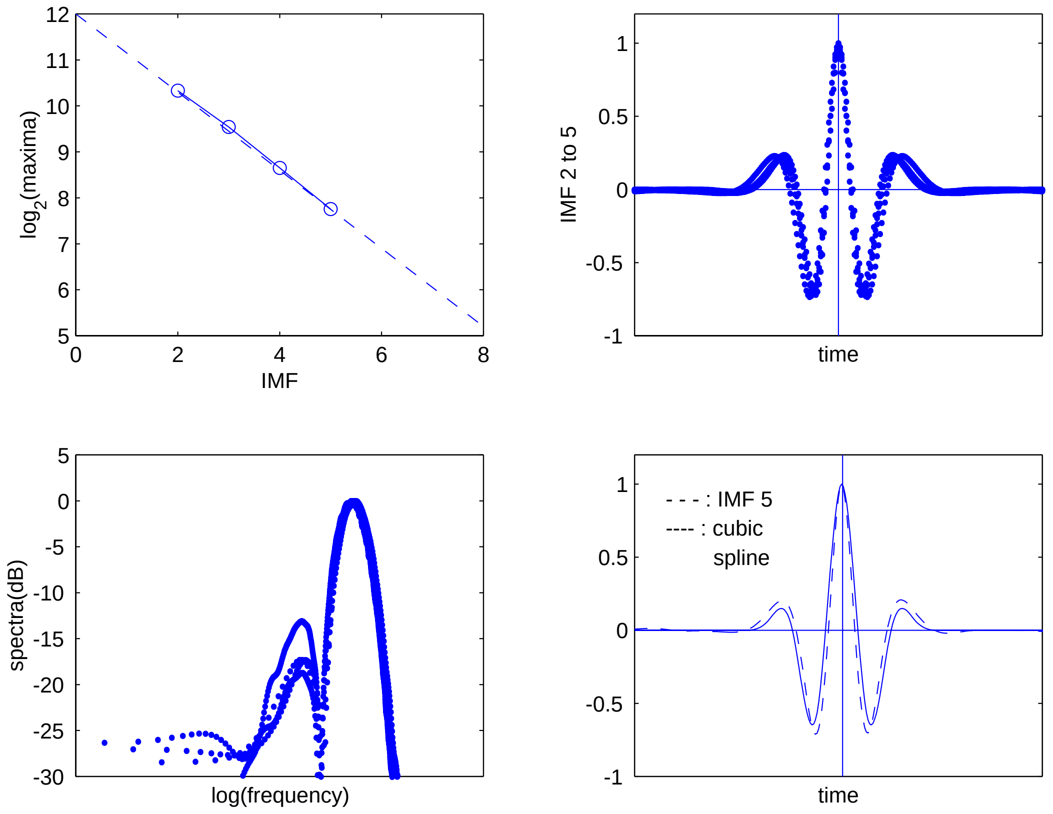
$$d_k[n] = \frac{1}{\alpha^k} \psi \left(\left\lfloor \frac{n}{\alpha^k} \right\rfloor \right),$$

for any $k' > k \geq 2$ and where $\alpha \approx 2^p$, with p s.t. $d_k[0] = 2^{C-pk}$.

EMD equivalent impulse response



EMD of a pulse and renormalization



Back to fGn

Modes variance — The self-similarity of spectra leads, for the variance $V_H[k] := \mathrm{var}\, d_{k,H}[n]$, to:

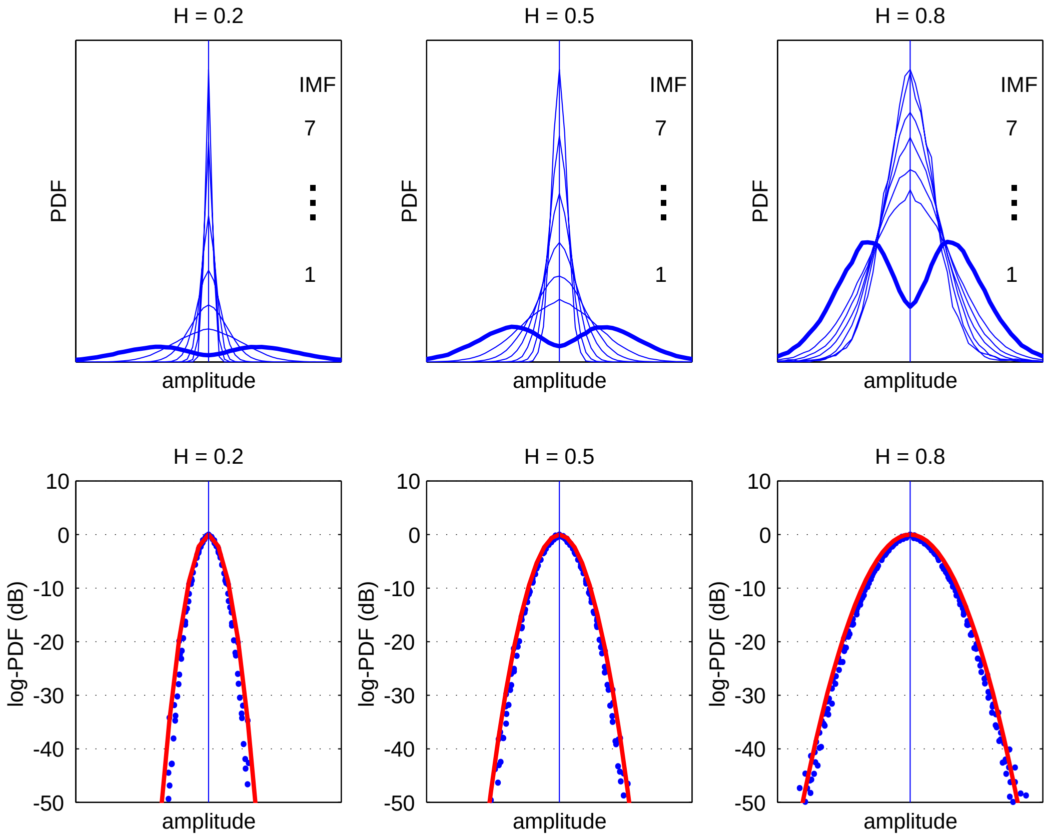
$$\begin{aligned} V_H[k'] &= \int_{-1/2}^{1/2} \mathcal{S}_{k',H}(f)\, df \\ &= \rho_H^{\alpha(k'-k)} \int_{-1/2}^{1/2} \mathcal{S}_{k,H}(\rho_H^{k'-k} f)\, df \\ &= \rho_H^{(\alpha-1)(k'-k)}\, V_H[k], \end{aligned}$$

and, hence, to: $V_H[k] = C\, \rho_H^{2(H-1)k}$.

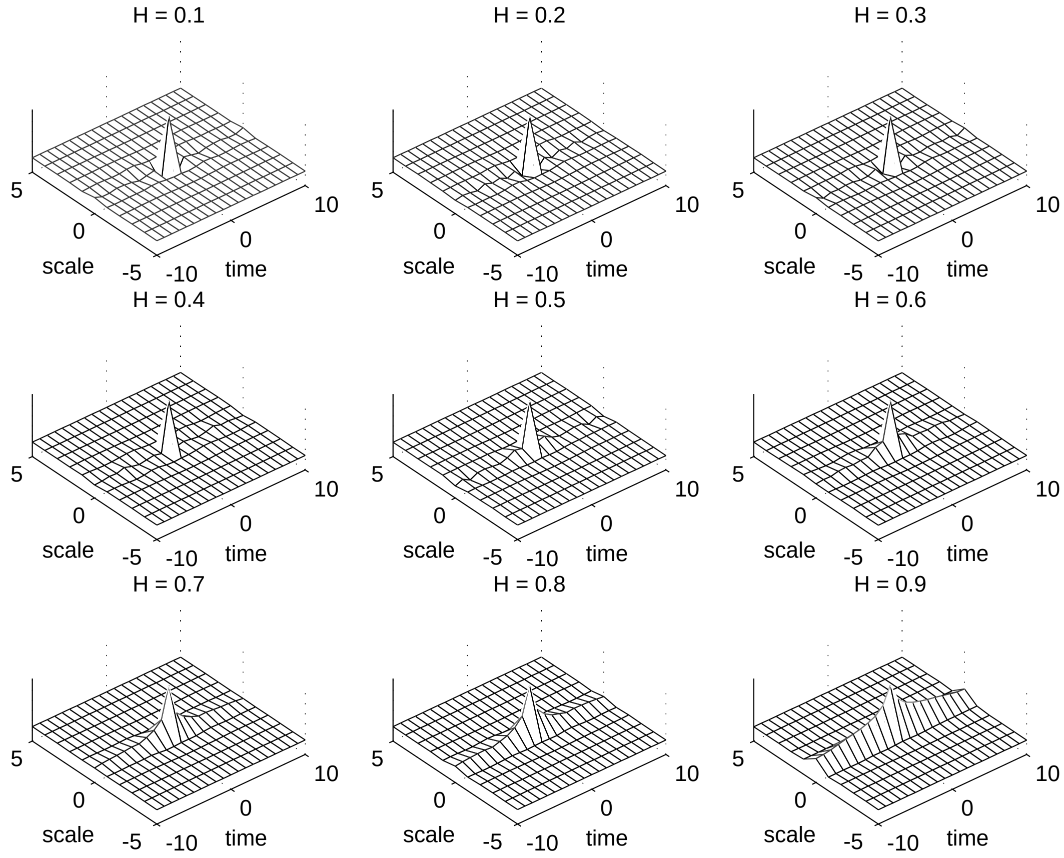
Marginal statistics — Gaussiannity for $k \geq 2$.

Correlations — Reduced, except intra-mode for large H's.

IMFs marginal statistics



IMFs correlation



Estimation of the Hurst exponent

Empirical variance — Evaluation of energy, mode by mode:

$$\widehat{V}_H[k] = \frac{1}{N}\sum_{n=1}^{N} d_{k,H}^2[n].$$

Slope — The variance model being *linear* in a semi-log diagram:

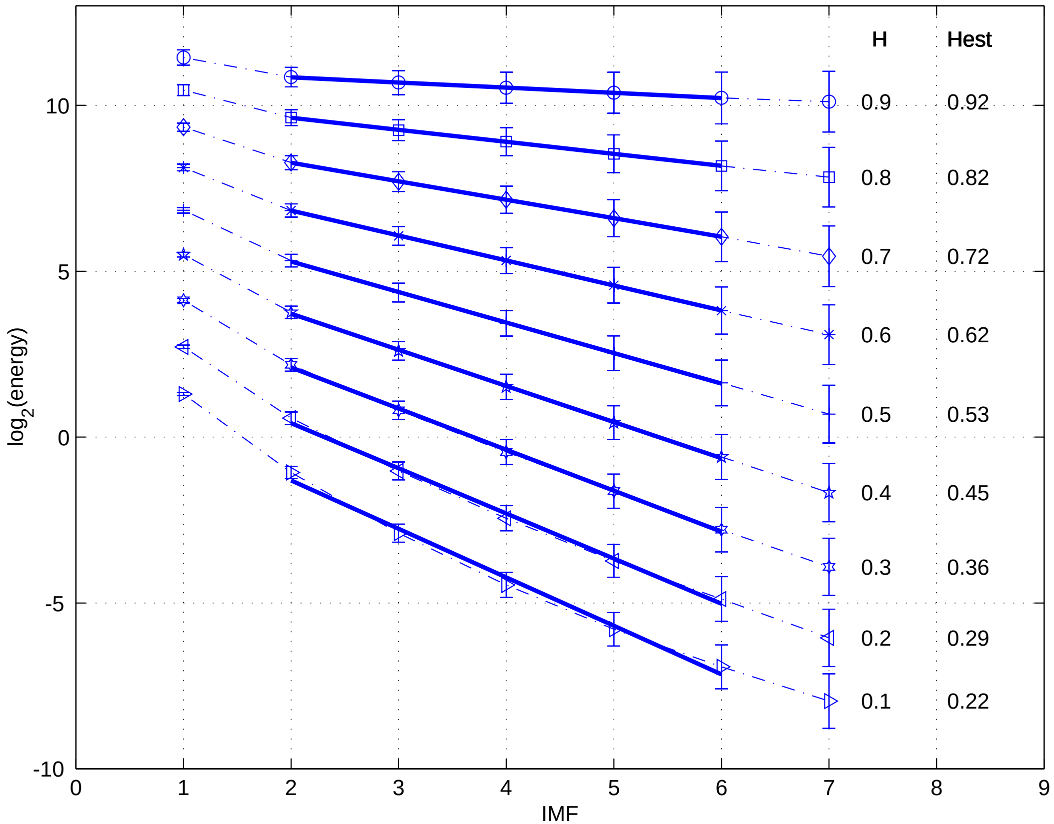
$$\log V_H[k] = [2(H-1)\log \rho_H]\, k + \mathrm{Cst},$$

an *estimate* $\widehat{H}$ of the Hurst exponent H reads:

$$\widehat{H} = 1 + \frac{\kappa_H}{2},$$

where κ_H is the measured slope.

EMD-based estimation of the Hurst exponent



Modes manipulation

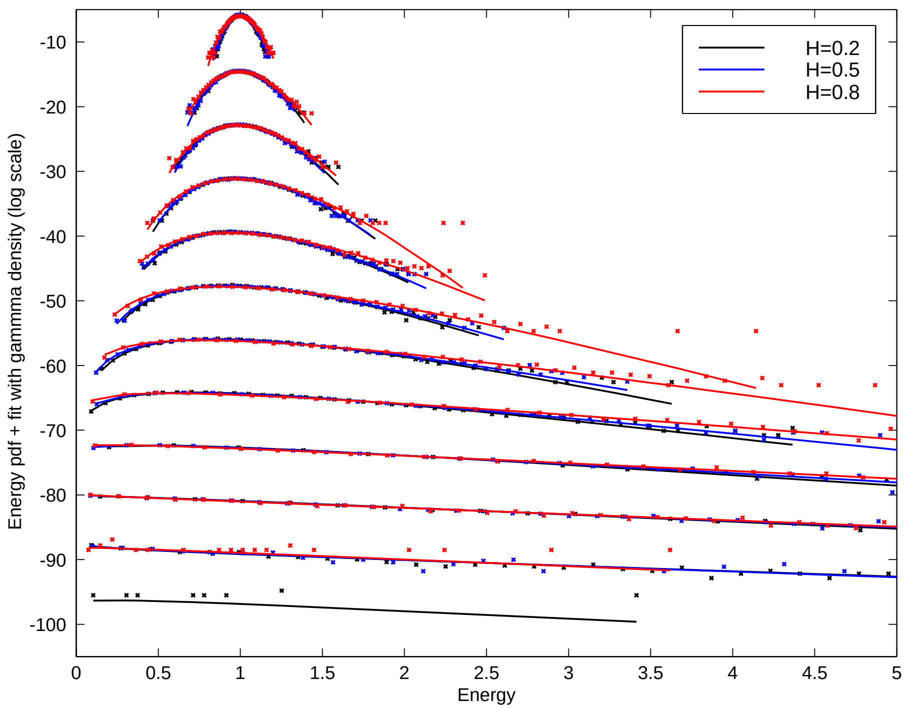
Reconstruction — The modes extraction process is *nonlinear*, but their linear recombination is *exact*. Modes selection $\Rightarrow$ possibilities of partial reconstructions.

Significant modes — For a given noise model, the (empirical) evaluation of dispersion allows, mode by mode, to reject or not the noise only hypothesis.

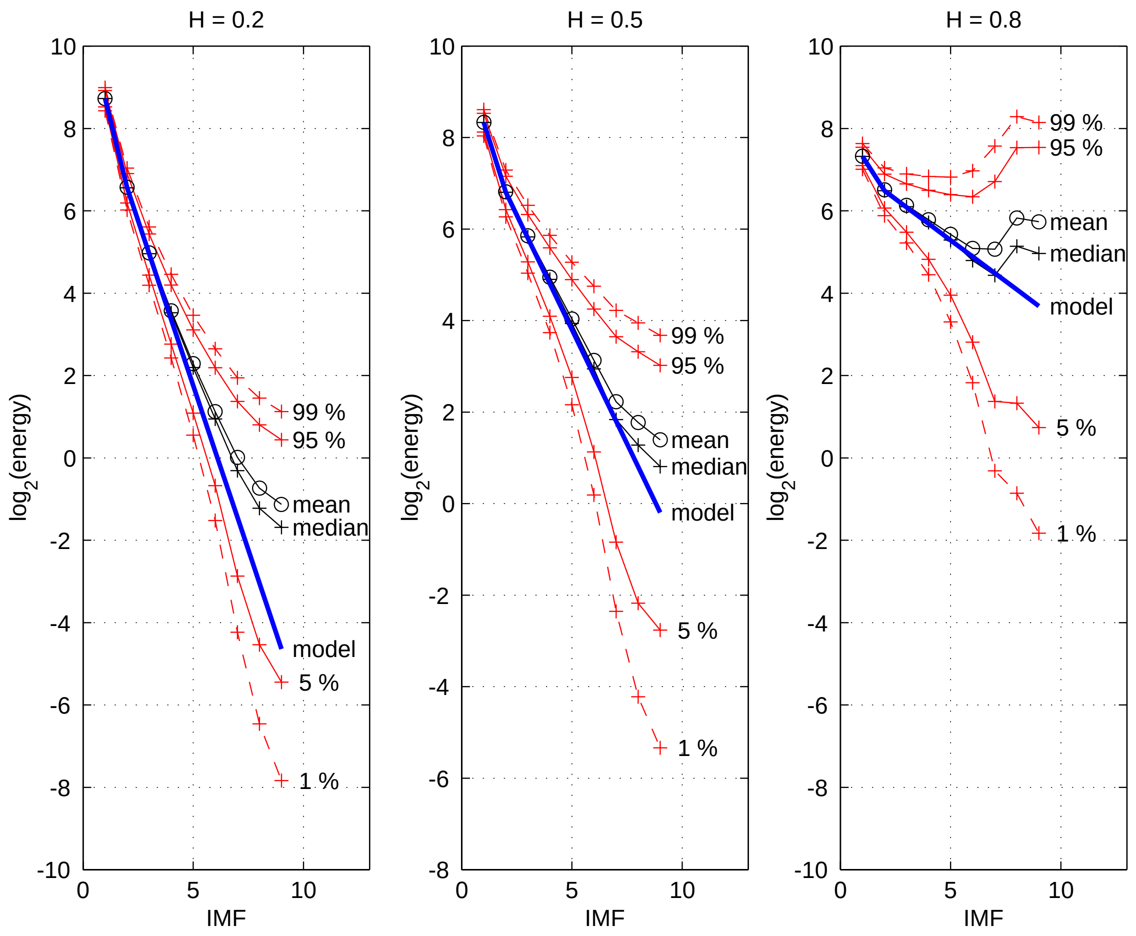
Applications — *Denoising* and *detrending*, depending on whether one keeps or rejects those modes whose energy is greater than a threshold controlled by the model and some acceptation level.

Modes dispersion

IMFs energy — Following Wu '01, empirical histograms show that the IMFs energy is *Gamma distributed* (χ^2 when normalized), with # degrees of freedom ≈ # extrema.



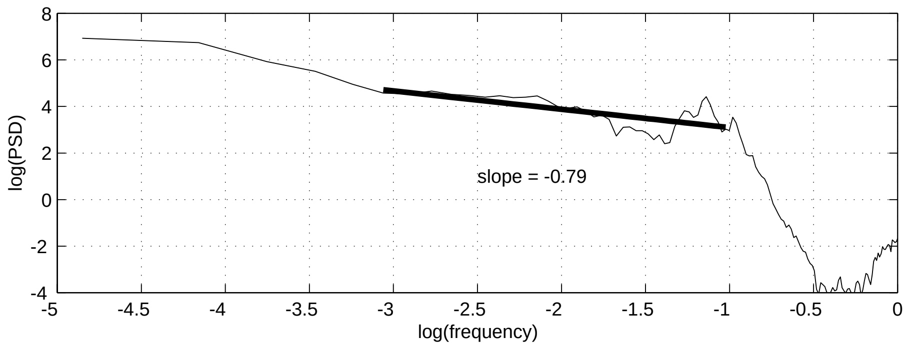
Mean, median and confidence intervals



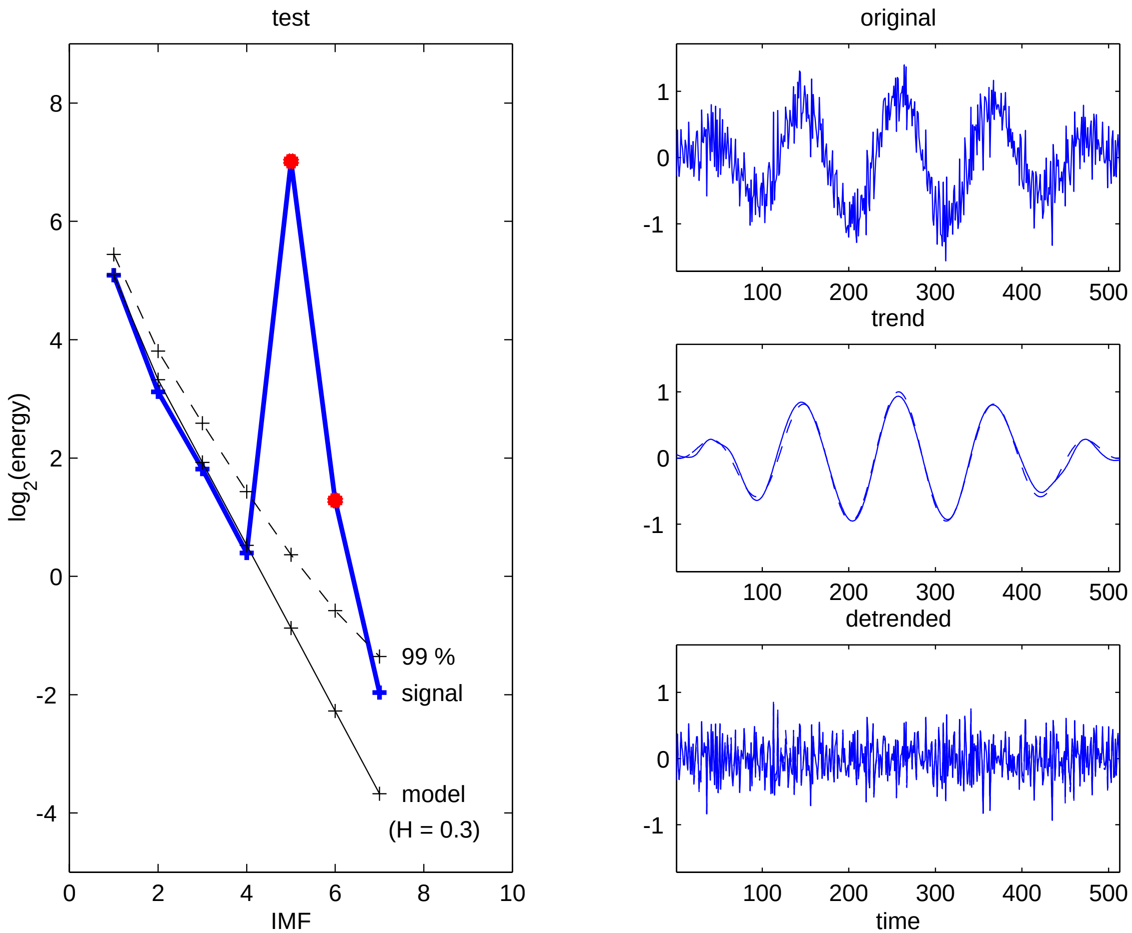
Two examples

Toy example — Tapered low frequency sine wave embedded in fGn with known Hurst exponent $H = 0.3$.

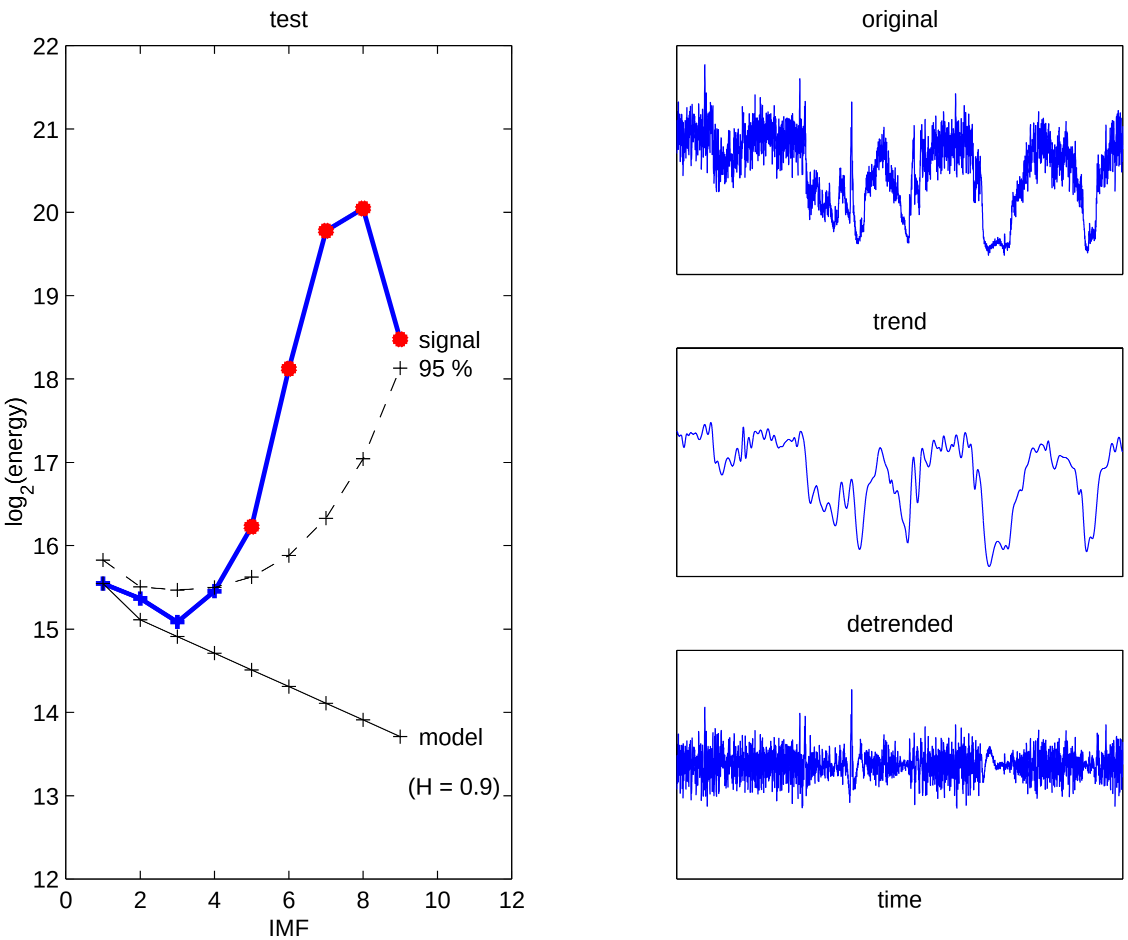
Real data — Heart-Rate Variability (HRV) time series, with H estimated from the observed spectrum.



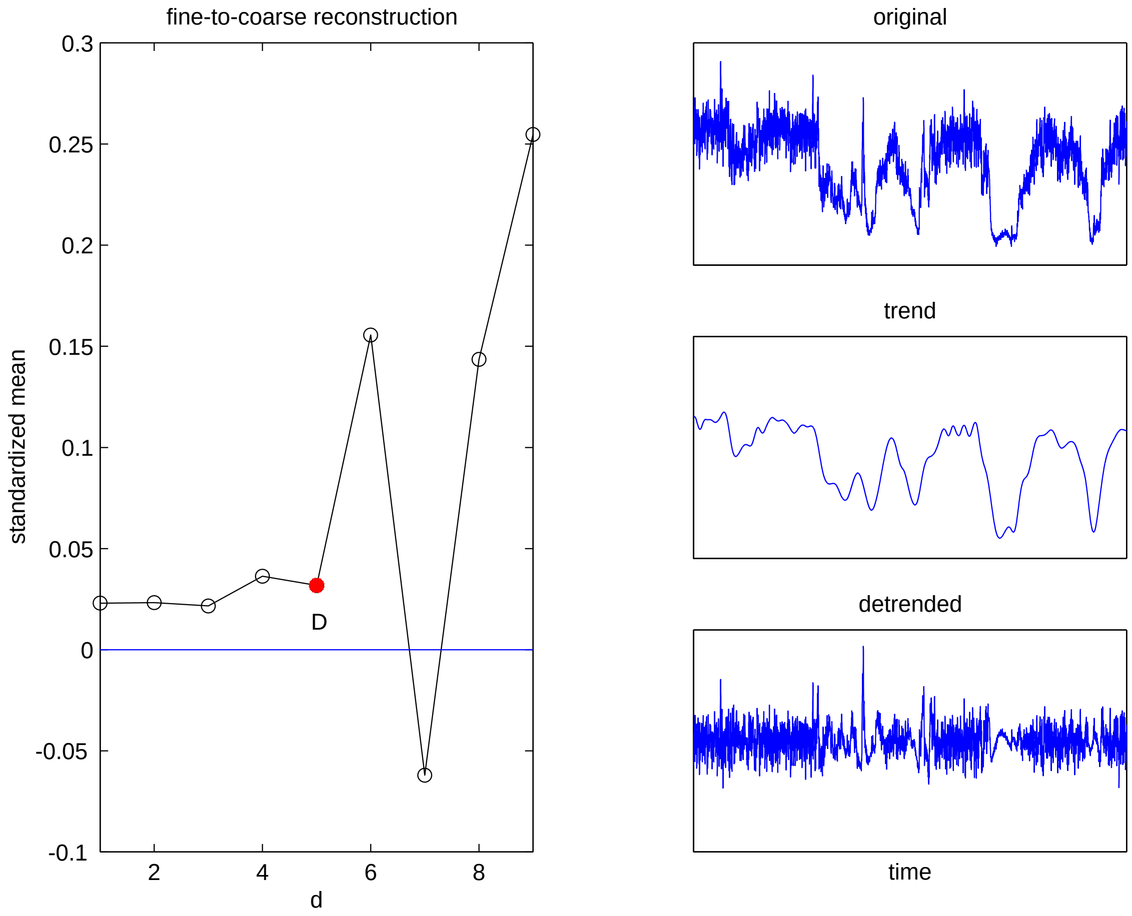
Toy example


test
log2(energy)
IMF
99 %
signal
model
(H = 0.3)
original
trend
detrended
time

HRV data


test
22
21
20
19
18
17
16
15
14
13
12
0
2
4
6
8
10
12
IMF
log2(energy)
signal
95 %
model
(H = 0.9)
original
trend
detrended
time

Another approach



Concluding remarks on Huang's EMD

Rationale — Intuitive, simple, local and fully data-driven.

Benefits — Spontaneous organization as a time-varying wavelet-like filter bank, new approach to decomposition, denoising and detrending, adapted to nonlinear oscillations.

Issues — *Ad hoc* and user-controlled tunings, sensitivity to small perturbations.

still lacks from solid theoretical grounds references, preprints, software, demos

`http://perso.ens-lyon.fr/patrick.flandrin/`