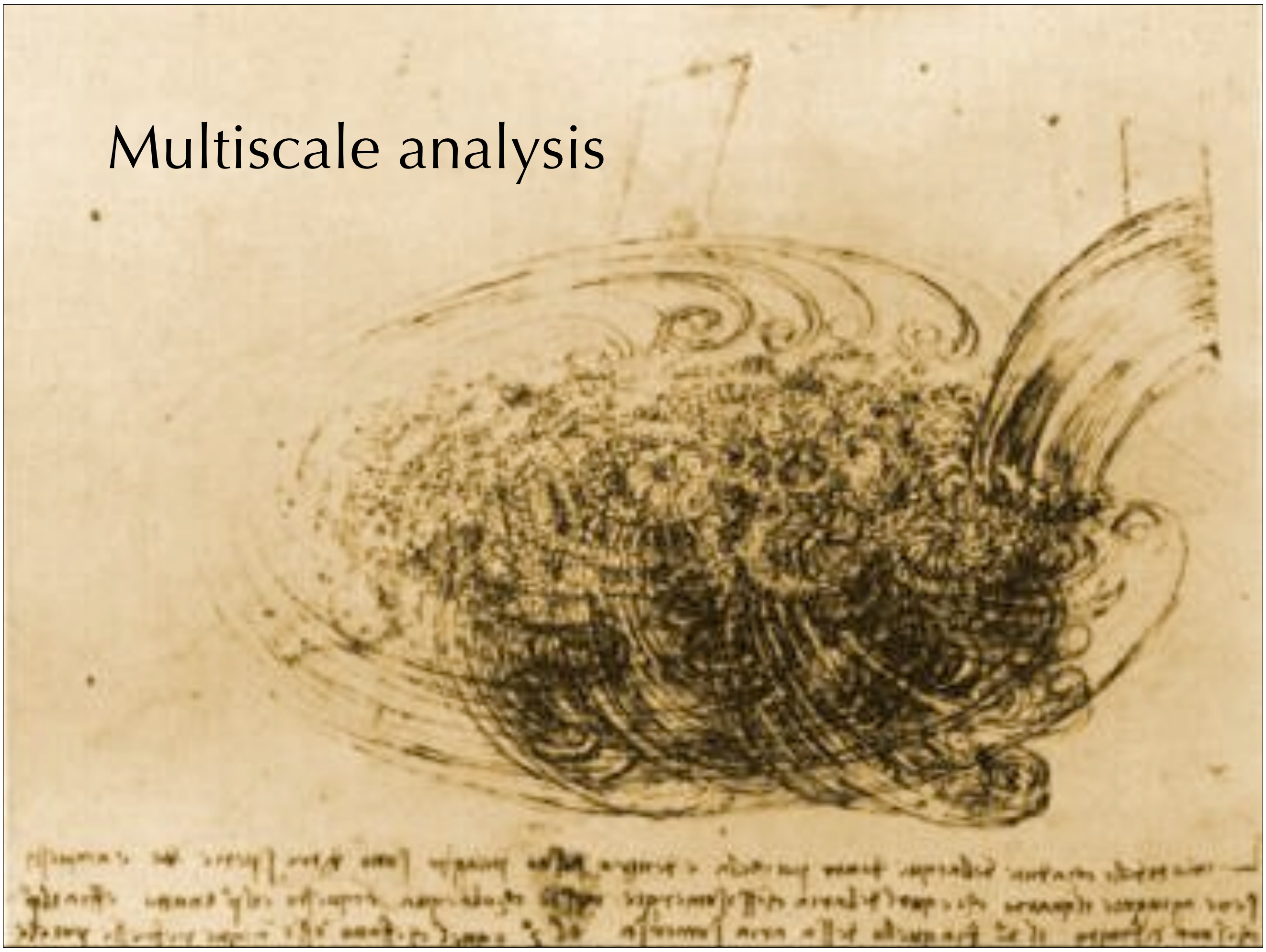
Multiscale analysis



- Most of our data analysis techniques involve strong assumptions on linearity, stationarity, etc.

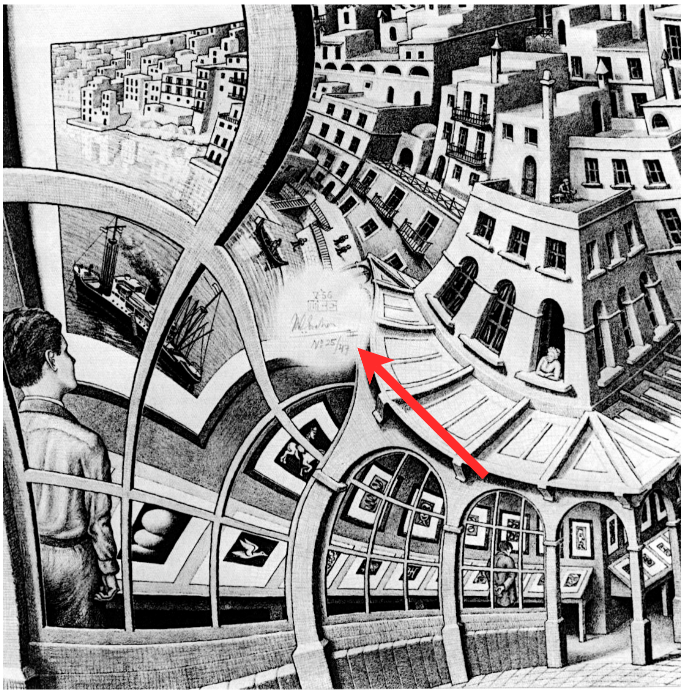
- Yet the world we live in is
  - non-linear
  - non-stationary
  - non-Gaussian
  - spatio-temporal
  - ...

- The proper understanding of these “***non-properties***” requires new (more advanced) tools.
- Many mistakes have been made by ignoring these non-properties
- These “non-properties” can however give deeper insight into the physics

---

An old mystery

- M. C. Escher didn’t know how to properly finish his drawing



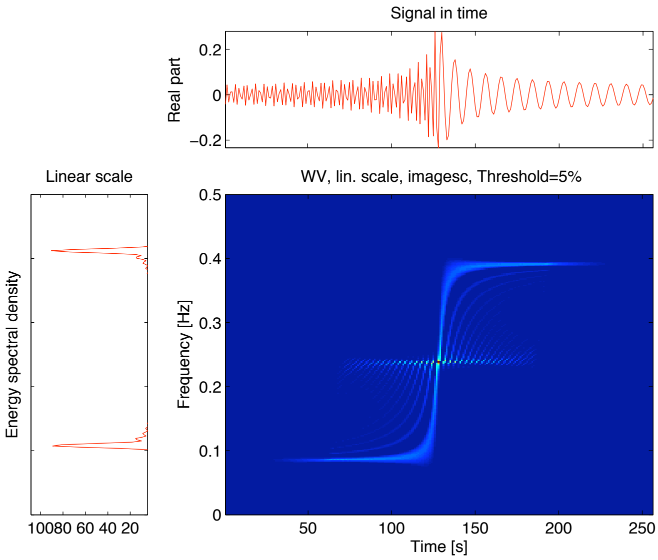
An old mystery

- The solution was found last year : self-similarity !

Wigner-Ville

- Unfortunately, ***interferences*** arise when signals are mixed

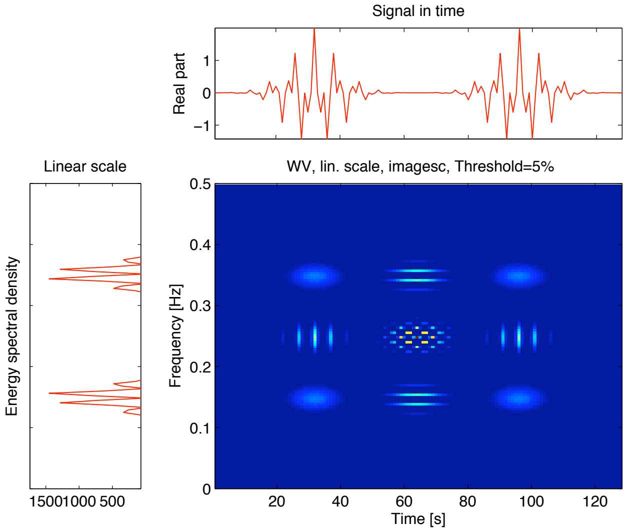
$$W_{x+y}(t,\nu) = W_x(t,\nu) + W_y(t,\nu) + 2\Re\{W_{x+y}(t,\nu)\}$$



Wigner-Ville

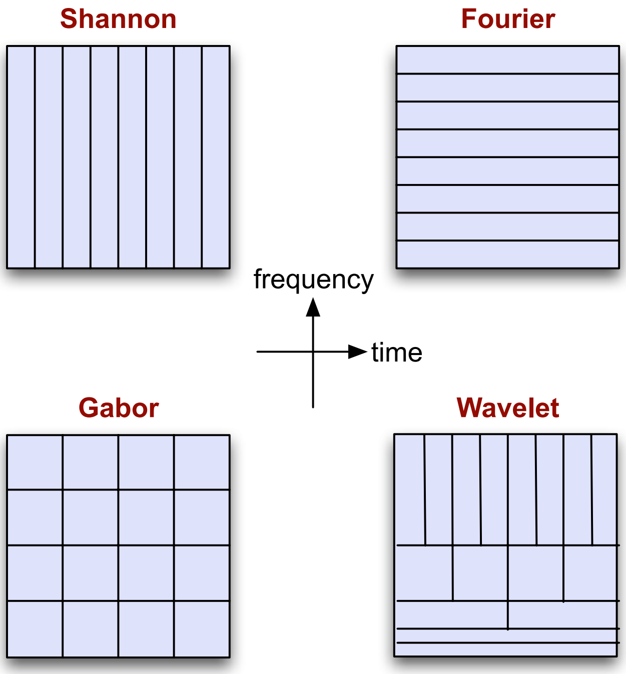
- Various methods have been developed for attenuating interferences in the Wigner-Ville transform

  but these cross-terms remain a severe problem



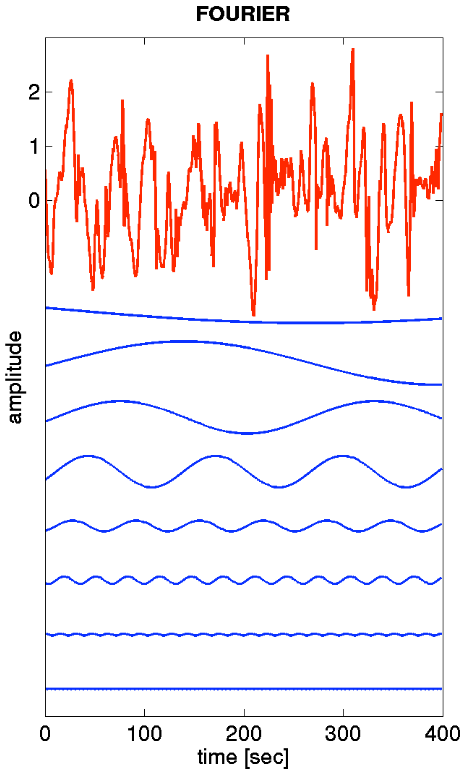
Tiling the plane

- The problem is : ***how to properly partition the time-frequency plane ?***



$$\Delta t \cdot \Delta \nu \geq \frac{1}{4\pi}$$

Fourier vs multiscale analysis



Wavelet transform : definition

Fourier transform

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

Continuous wavelet transform

$$X(\tau, a) = \langle x(t), \psi_{\tau,a}(t) \rangle = \int_{-\infty}^{+\infty} x(t)\, \psi^*_{\tau,a}(t)\, dt$$

mother wavelet

$$\psi_{\tau,a}(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t-\tau}{\underset{\nwarrow \text{scale}}{a}} \right)$$

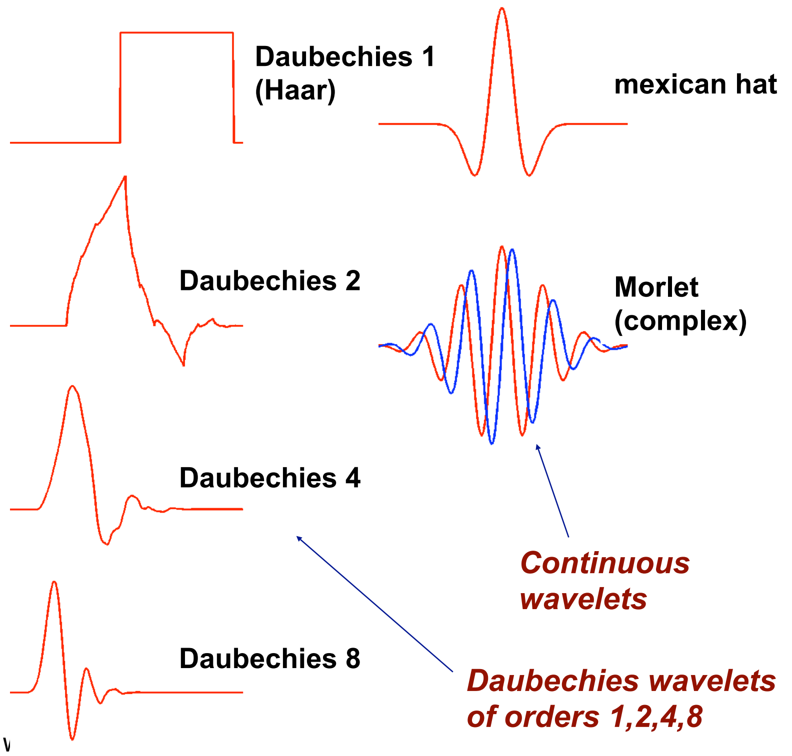
Discrete wavelet transform

$$X_{j,k} = \langle x(t), \psi_{j,k}(t) \rangle = \int_{-\infty}^{+\infty} x(t)\, \psi^*_{j,k}(t)\, dt$$

mother wavelet

$$\psi_{\tau,aj,k}(t) = \frac{1}{2^{j/2}} \psi \left(2^{-j} t - k \right)$$

Mother wavelets : examples



Daubechies 1 (Haar)

Daubechies 2

Daubechies 4

Daubechies 8

mexican hat

$$\psi(t) = \frac{d^2}{dt^2}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-t^2/2\sigma^2}$$

Morlet (complex)

$$\psi(t) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-t^2/2\sigma^2}e^{-2\pi j\nu t}$$

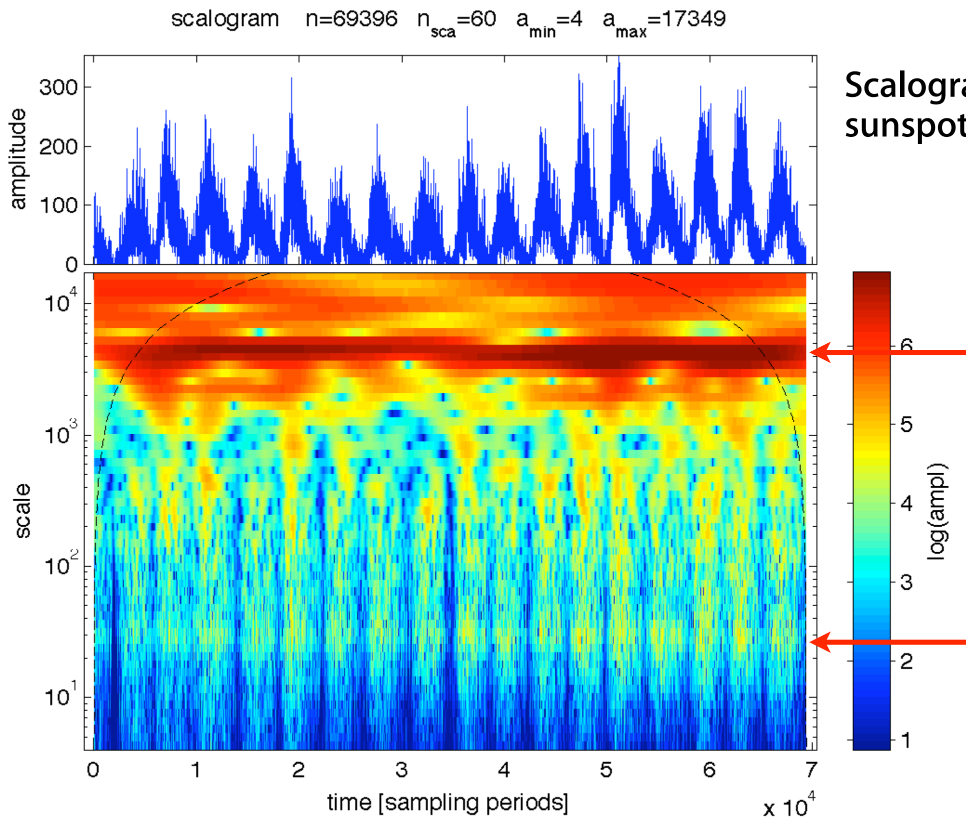
Continuous wavelets

Daubechies wavelets of orders 1,2,4,8

The wavelet shape should reflect the type of features that are present in the data

---

Scalogram



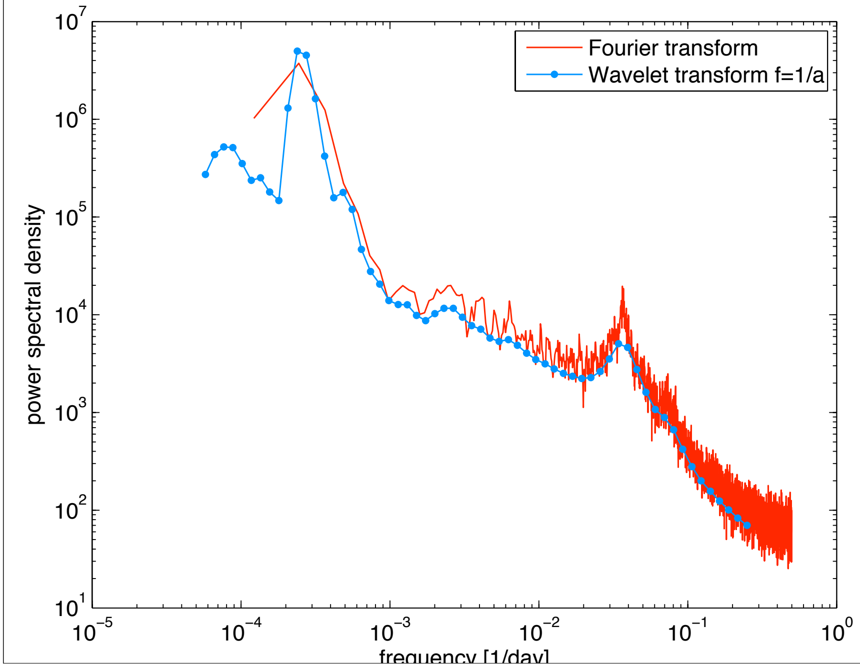
Scalogram of the daily sunspot number

a=11 years solar cycle

a=27 days solar rotation

Power spectral density

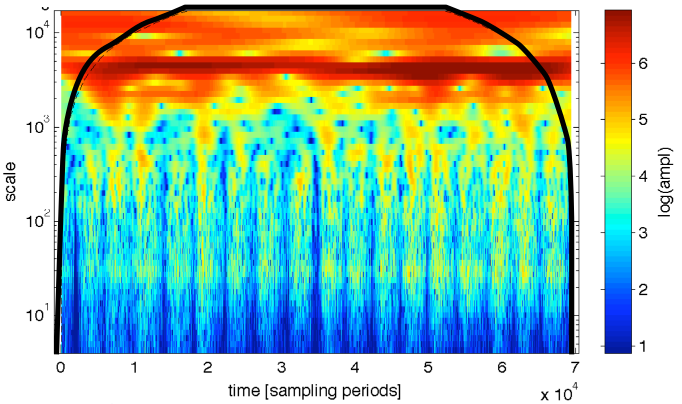
- By integrating $|\langle x(t), \psi_{a,\tau}(t)\rangle|^2$ over time, we recover a "power spectral density"



Morlet wavelet transform of the sunspot number, with
frequency = 1/scale

What scales ?

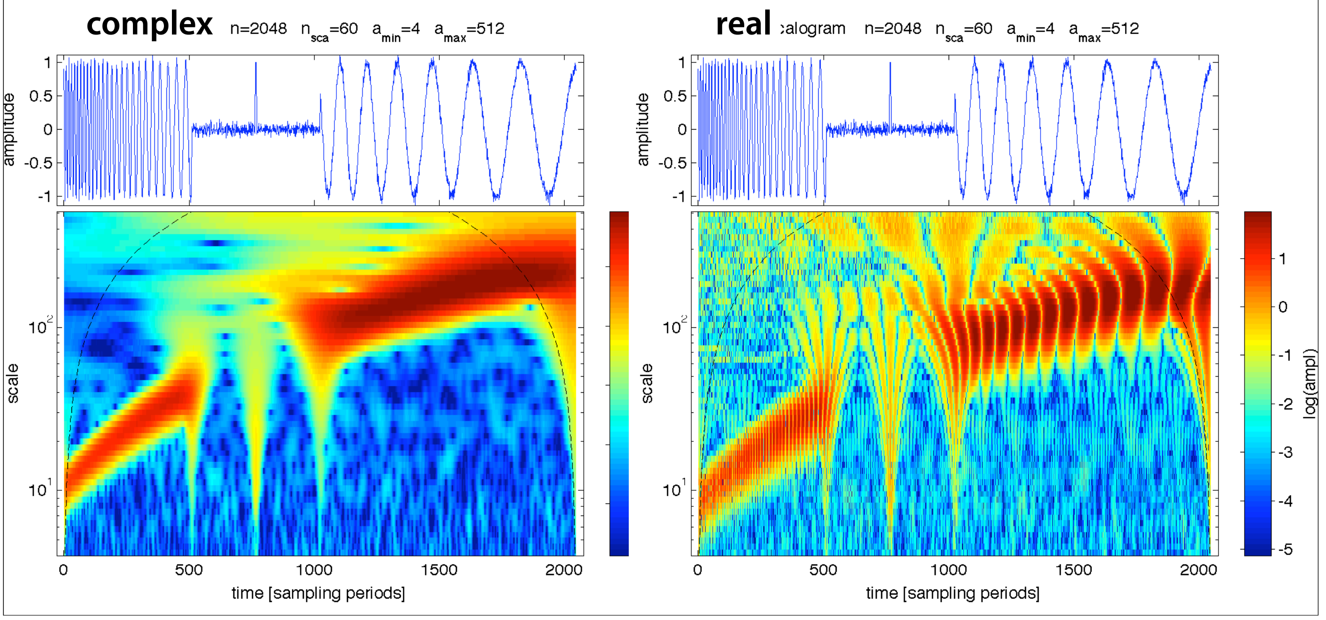
- As in Fourier analysis, the scales are bounded
  - The smallest scale should be > 4 sampling periods (2 for Fourier)
  - The largest scale should typically be < 0.25 x sequence length. This defines the ***cone of influence***



- Since the wavelets are self-similar, the scales should increase logarithmically (≠ linearly as in Fourier)
- For the discrete wavelet transform, a dyadic grid is imposed

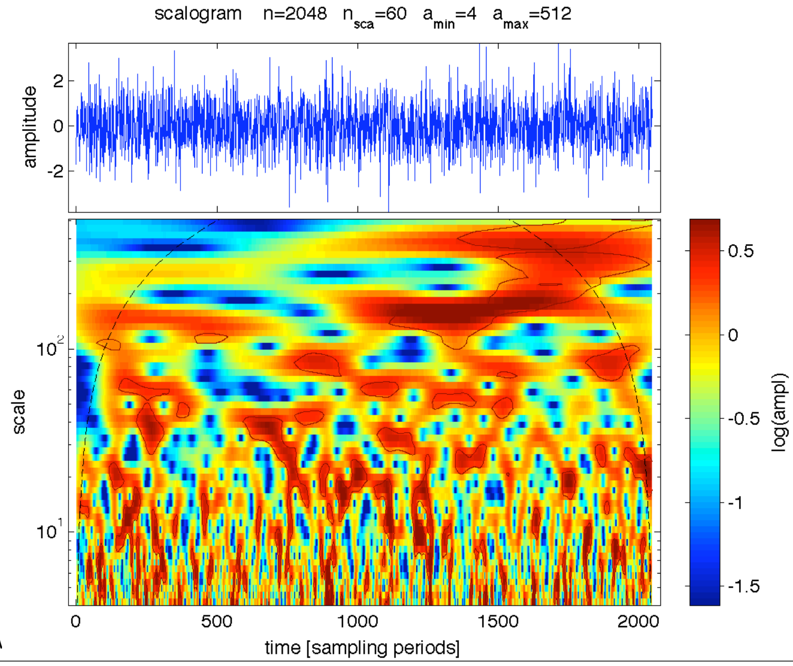
Complex or real ?

- ***Complex wavelets*** provide information about phase and amplitude. They are more appropriate for oscillatory behaviour.
- ***Real wavelets*** are used instead for probing peaks and discontinuities.


complex **real**

How significant are the wavelet coeffs ?

- Plotting confidence intervals is essential for determining the significance of peaks in the wavelet transform



Scalogram of a white noise process

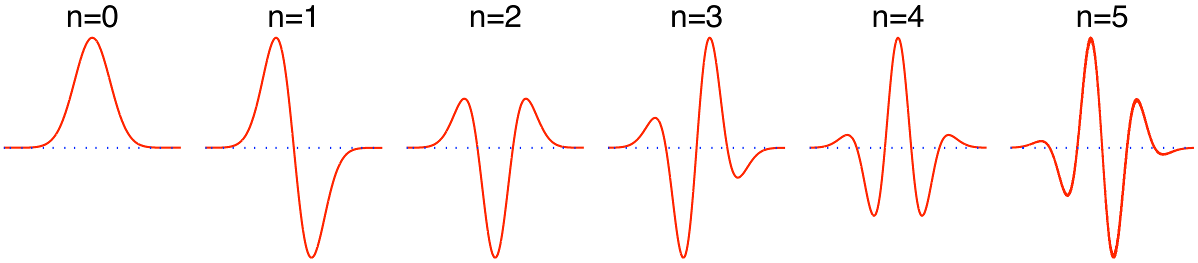
Are the peaks really that significant ?

Vanishing moments

A wavelet has ***n*** vanishing moments if

$$\int_{-\infty}^{+\infty} t^k \, \psi(t) \, dt = 0 \quad \text{for} \quad 0 \leq k < n$$

such a wavelet is orthogonal to polynomials of degree ***n-1***



Hölder (Lipschitz) regularity

A function ***x(t)*** has a local Hölder exponent **h** if

$$|x(t+\tau) - x(t)| \leq K|\tau|^{h(t)}, \quad \tau \to 0$$

Then the wavelet transform (with ***n***'th order wavelets) will satisfy

$$|X(\tau, a)| \leq \begin{cases} K\,|a|^{h(t)}, & a \to 0 \quad \text{if} \quad h(t) \leq n \\ K\,|a|^{n}, & a \to 0 \quad \text{if} \quad h(t) \geq n \end{cases}$$

If the wavelet order n is large enough, then the wavelet transform reveals local regularity

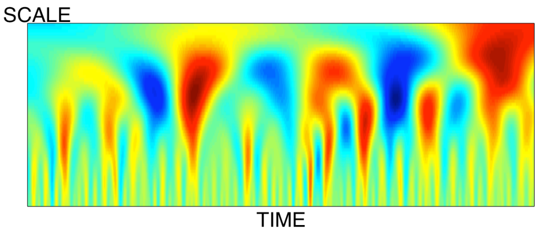
Local regularity

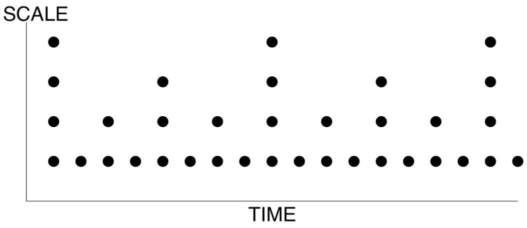
The consequences of the previous property are

- n’th order wavelets are insensitive to trends that can be described by polynomials of order < n
- low order wavelets thus should be avoided, unless the signal has a low regularity

Adapt you wavelet to the properties of your data set
• regularity
• symmetry
• time-frequency localisation
• ...

Discrete or continuous transform ?





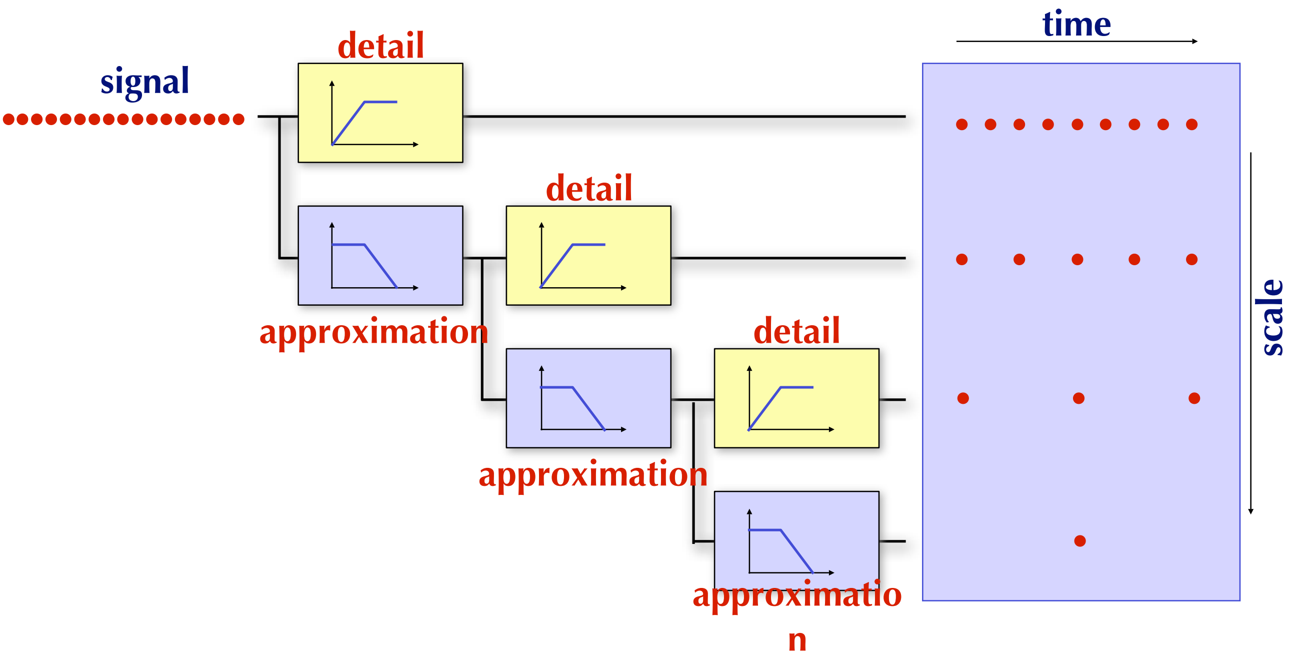
Continuous transform

- **highly redundant**
- **good for data analysis**
- **scales can be freely chosen**
- **computationally expensive**

Discrete transform

- non-redundant and uses orthogonal bases
- useful for multiresolution analysis (denoising) and compression
- scales are imposed
- very fast algorithms (faster than FFT)

Multiresolution analysis



The discrete wavelet transform can be implemented by means of a fast pyramidal recursive filter bank

---

Why use wavelets then ?

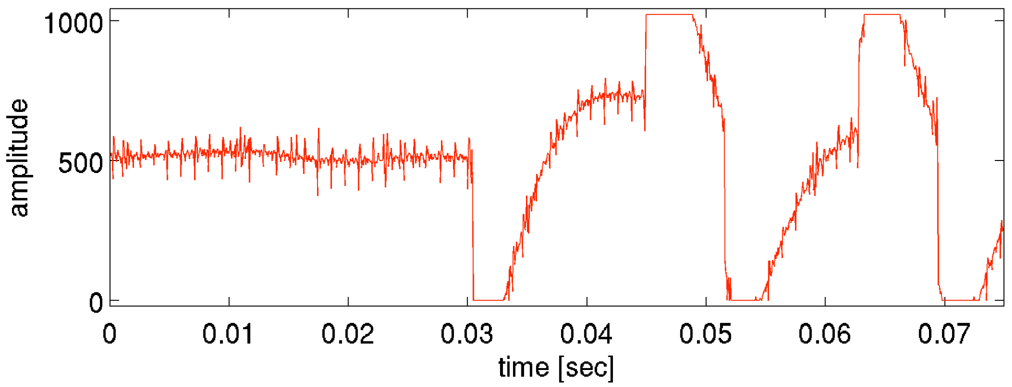
- Because wavelets are local in time AND frequency, they’re much more efficient than the Fourier transform for capturing transients.
- There have been attempts to use wavelet bases for simulating HD / MHD, with limited success so far.

Example

Wavelet denoising

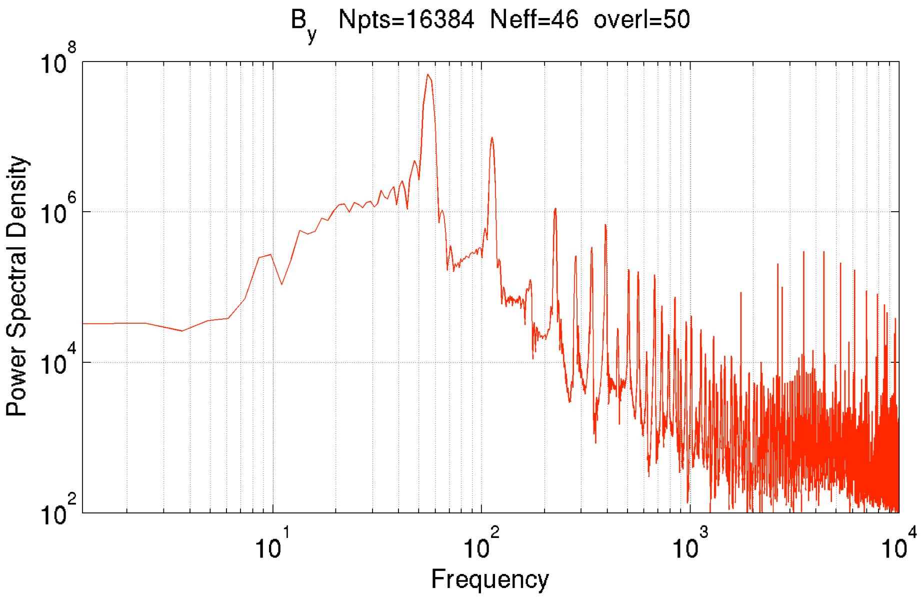
Example

AC magnetic field measurements made by the a sounding rocket in the terrestrial cusp region



The search coil data are heaviliy polluted by interference noise from nearby instruments. Is this hopeless ?

Spectral analysis



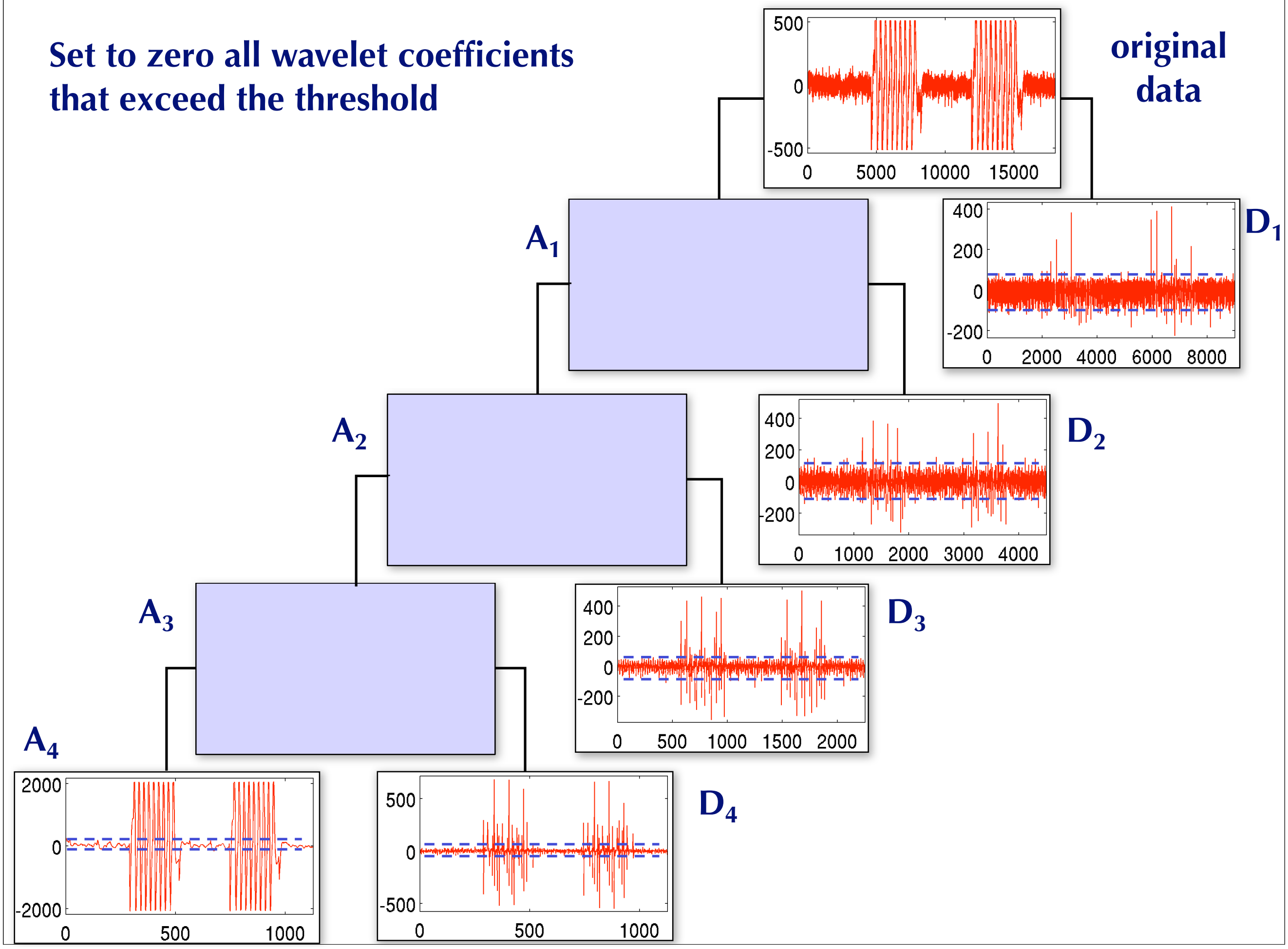
Interference noise occurs at all scales, and is non-stationary
➔ there is no way we can filter it out by Fourier analysis

Discrete wavelet analysis

Recipe for denoising with the discrete wavelet transform

- Decompose the signal at different levels (**A**pproximations and **D**etails
- At each level, compare the wavelet coefficients for noisy and quiet periods
- Set to 0 all coefficients corresponding to “noise”
- Reconstruct the data

Set to zero all wavelet coefficients that exceed the threshold



Results

Time series

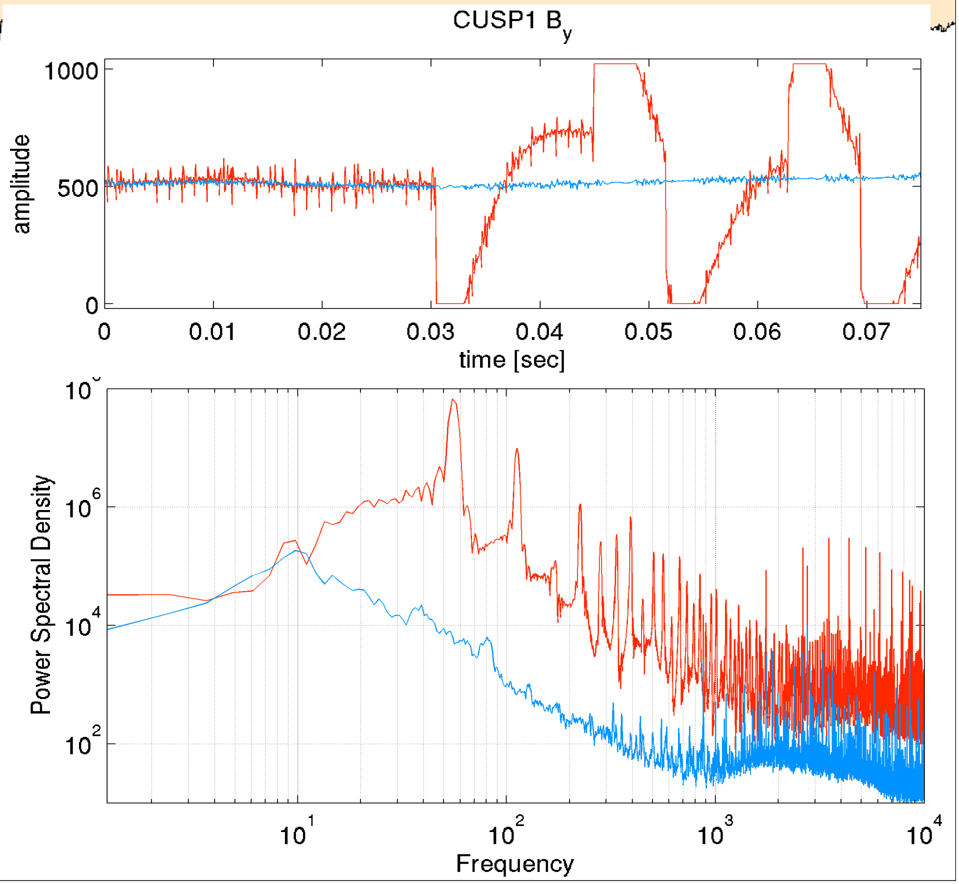
— raw signal

— denoised

Power spectral density

— raw signal

— denoised



Practical issues

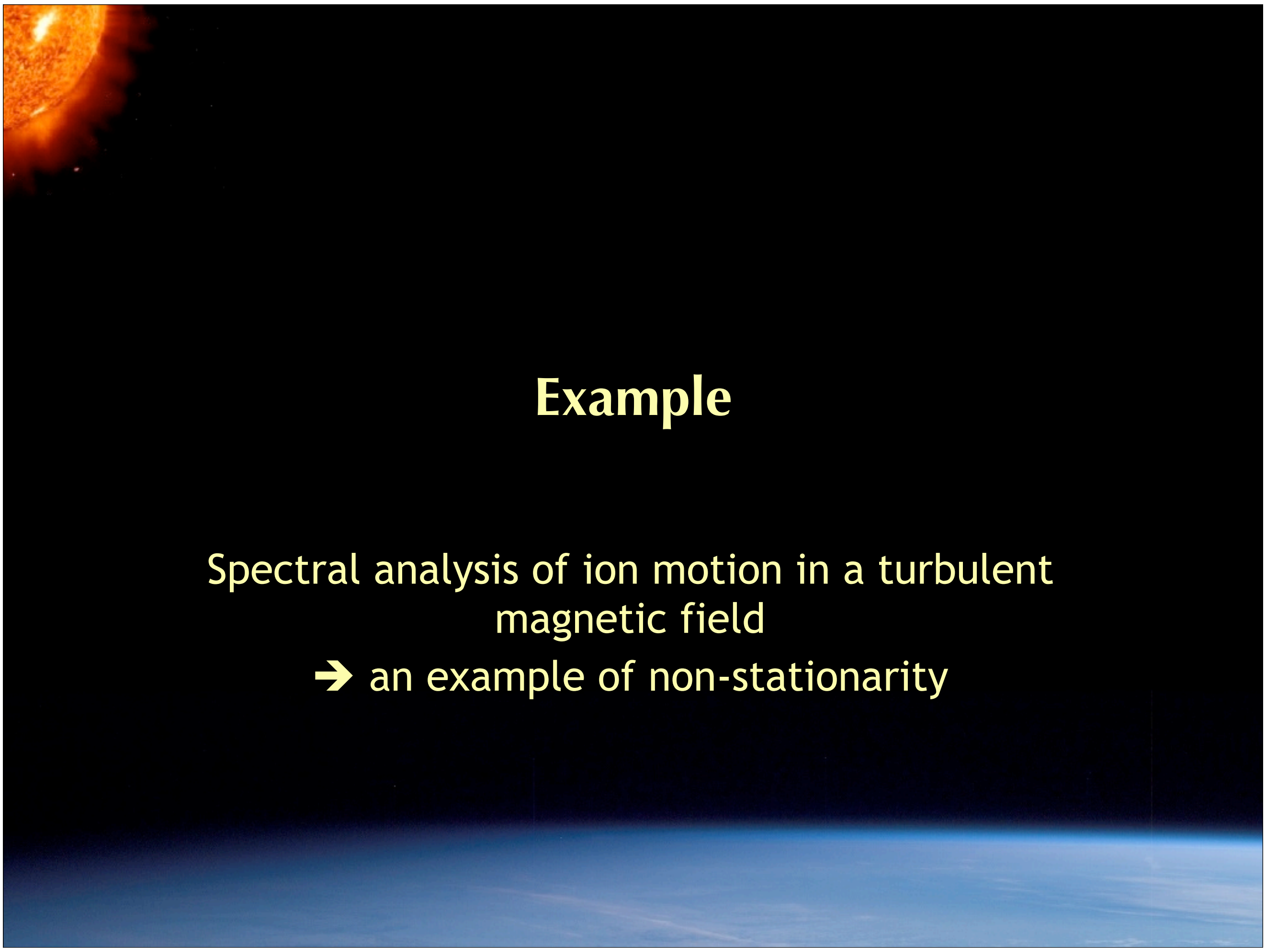
- Which wavelet should I use ?
  - Not so important, but symmetry considerations may play a role.
  - Much more important is the choice of the wavelet order :
    - high order = smoother + larger support
    - low order = more discontinuous + smaller support
- How many levels ?
  - More levels = large scales are investigated more in detail
- Where should I set the threshold ?
  - Depends a lot on the type of « noise »
  - Automatic thresholding criteria have been proposed, but selection must be driven by knowledge of the noise

---

Conclusion

- Discrete wavelet transforms are remarkably efficient for removing transient patterns
- Their lack of redundancy makes their interpretation difficult
  ➔ they're rarely used for analysis purposes

« Wavelet denoising is like removing weed while saving daisies »
(M. Wickerhauser)



Example

Spectral analysis of ion motion in a turbulent magnetic field

➔ an example of non-stationarity

Simulation model

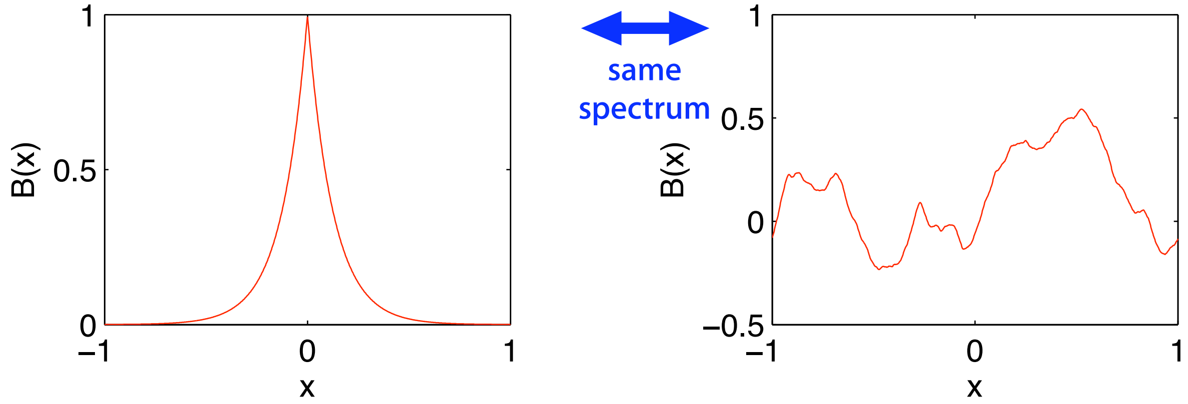
- Test particle simulations based on a 2-D model
  (Hada and Kuramitsu, 1999)

$$\begin{pmatrix} B_y \\ B_z \end{pmatrix} = \sum_{k,l} B_k^l \begin{pmatrix} \cos \\ \sin \end{pmatrix} (kx - \omega_k^l t + \phi_k^l)$$

- Phase correlation of adjacent modes can be adjusted by modifying the phases
- Wavefield is periodic in L (L=512), wavefield spectrum goes as

  $B_k \propto k^{-2}$
- amplitude and phase coherence are varied

Simulation model

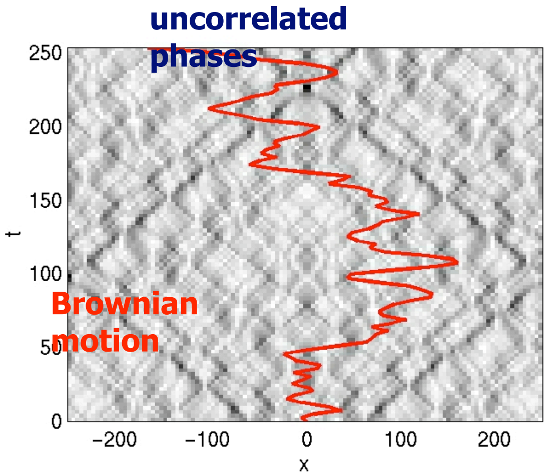


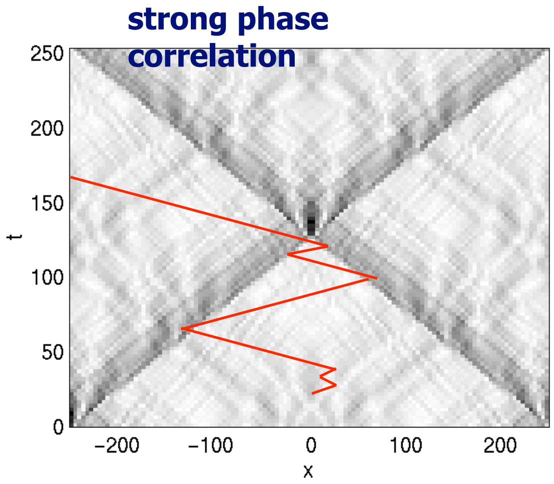
phase coherence strongly affects ion diffusion through nonlinear trapping by large-amplitude waves

---

Ion trajectories

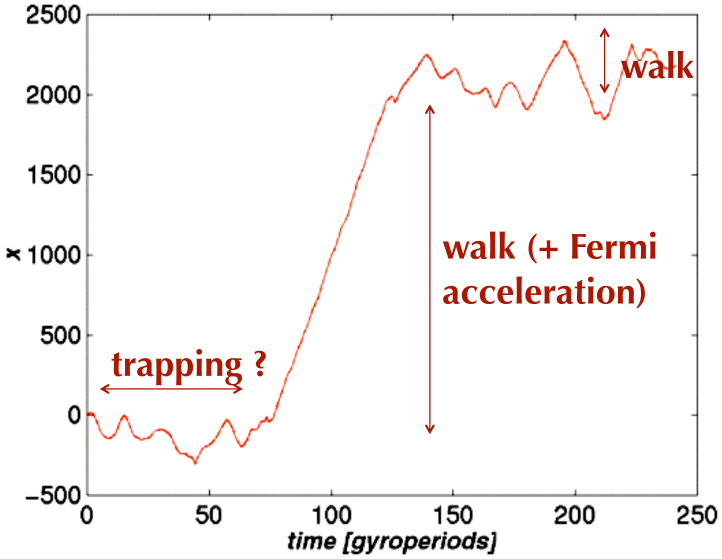
ion trajectories in the x–t plane (grey levels represent |B|)





Ion motion

Particle motion is a mixture of trappings (gyration) and walks (reflection by coherent structures)



walk

walk (+ Fermi acceleration)

trapping ?

Spectral analysis

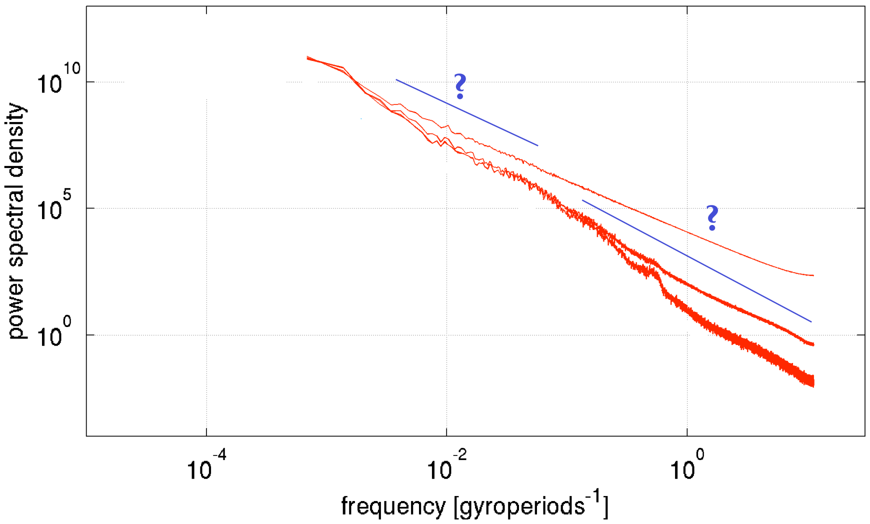
- We're dealing with a self-similar process

$$\Rightarrow u(x) = \lambda^{\alpha} u(\lambda x)$$

$$\Rightarrow u(k) = \mathcal{F}u(x) \sim k^{-\beta}$$

- Estimate the power spectral density of x(t) using windowed Fourier transform with various types of windows and detrendings

Spectral analysis

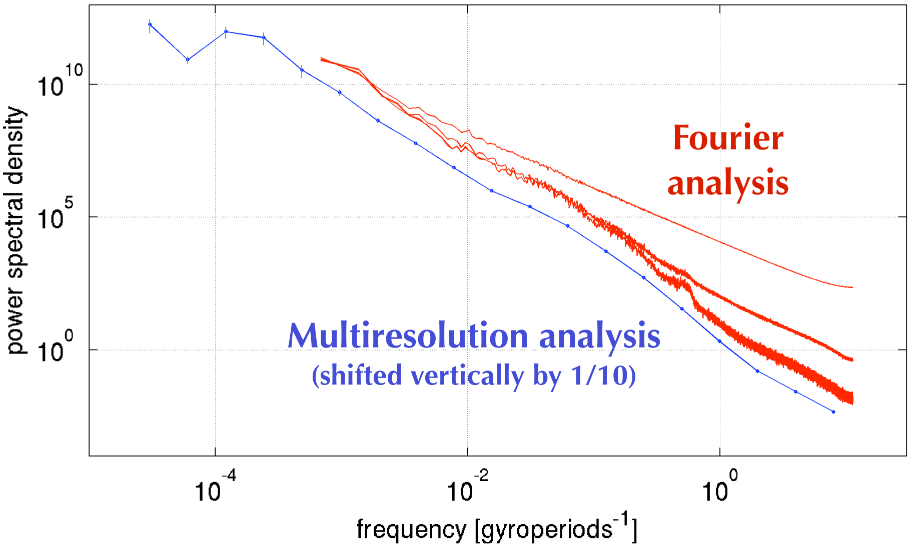


Self-similar ?

What is the correct spectral index ?

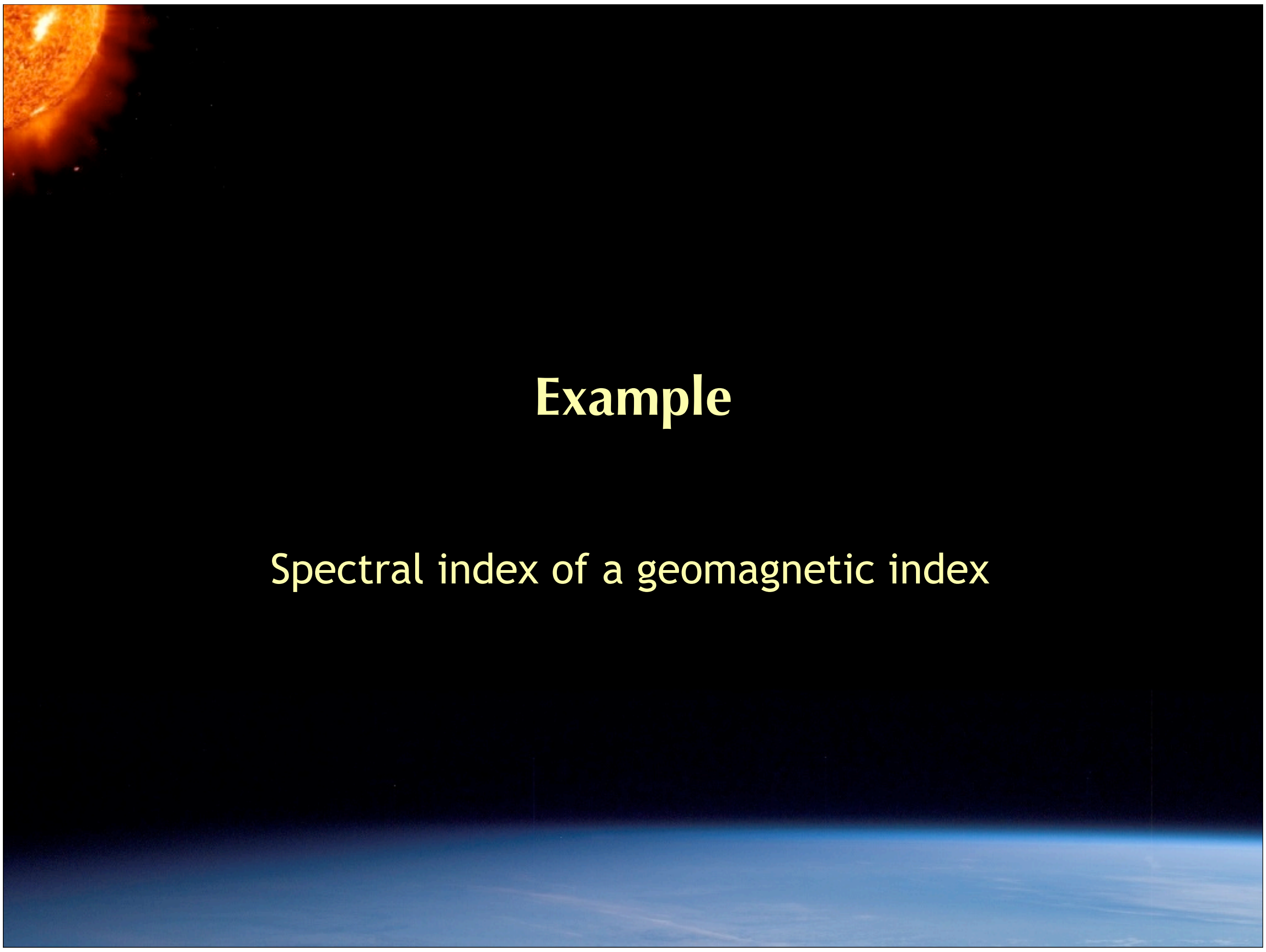
Spectral analysis

Multiresolution analysis provides a sound and unique estimate of the spectral index



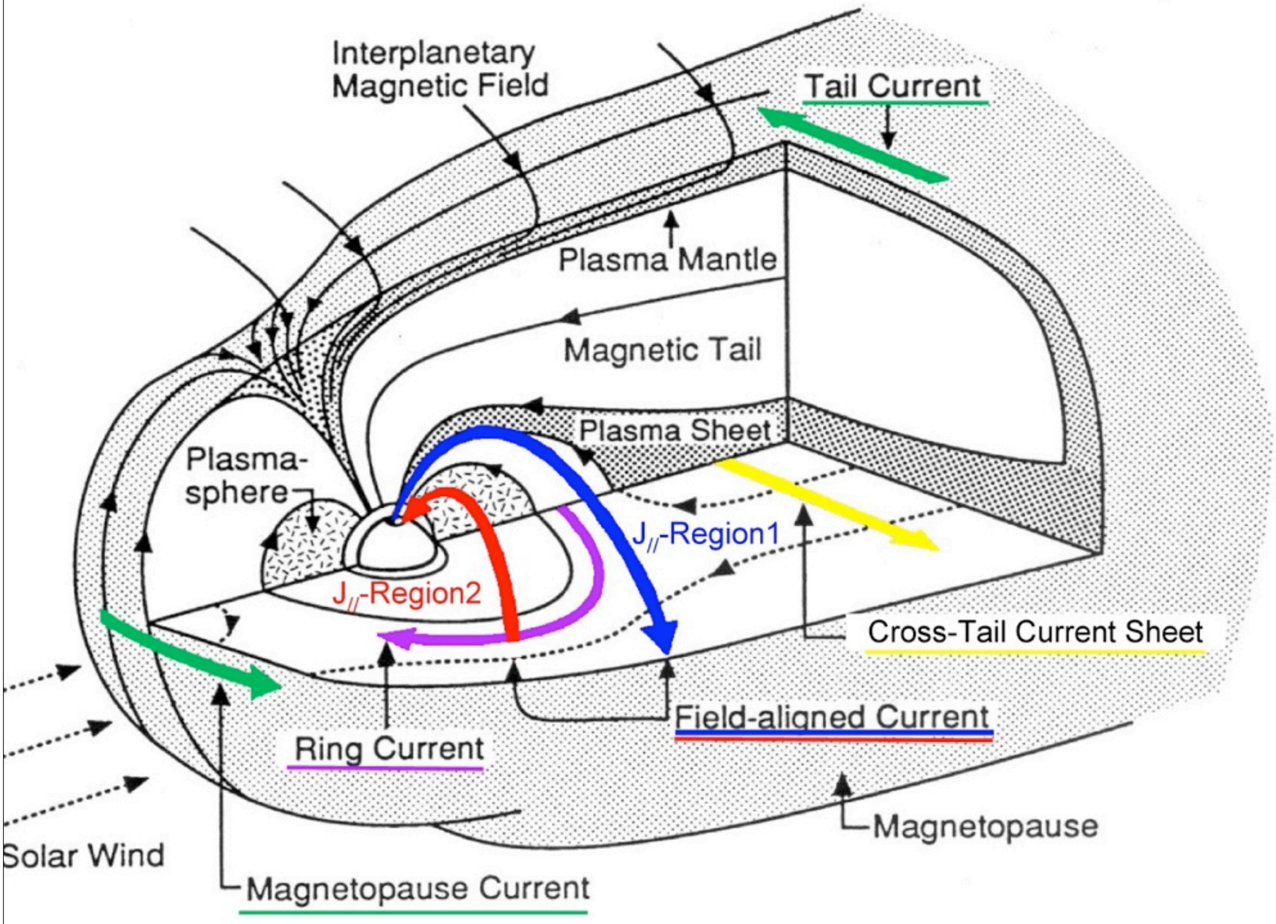
Conclusion

- Fourier analysis can be **wrong** when the time series is not sufficiently stationary
- Even for stationary data, Fourier analysis gives a biased estimate of the spectral index
  ➔ multiresolution analysis provides an **unbiased** and **more robust** estimate of the spectral index (Abry et al., 2000)
- But be careful to take wavelets of sufficiently high order

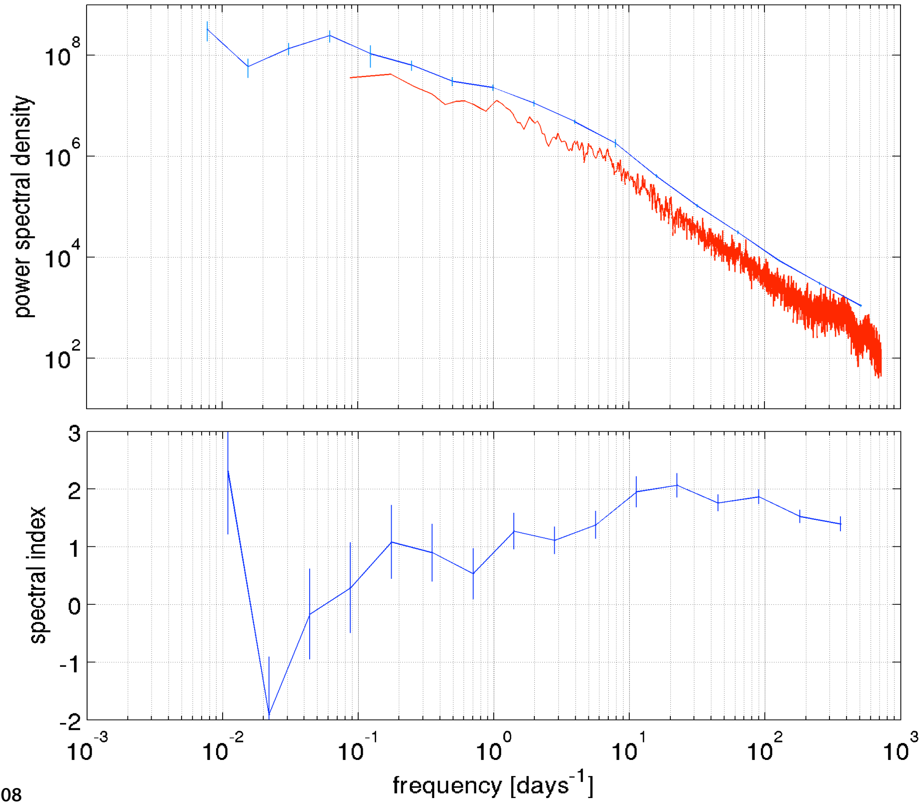


Example

Spectral index of a geomagnetic index



Spectral index for AE



Characterizing scale invariance

- Scale invariance is a key concept in the statistical analysis of turbulence
- The property of interest is the **interplay** between scales, rather than the role played by each individual scale
  ➔ spectral indices reveal one aspect only of scale invariance
- To get a more complete picture, other estimators are needed
  - structure functions
  - waiting times
  - singularity spectra
  - ...
- For the particular case of our ion motion, it is interesting to look at the distance travelled by the ion

---

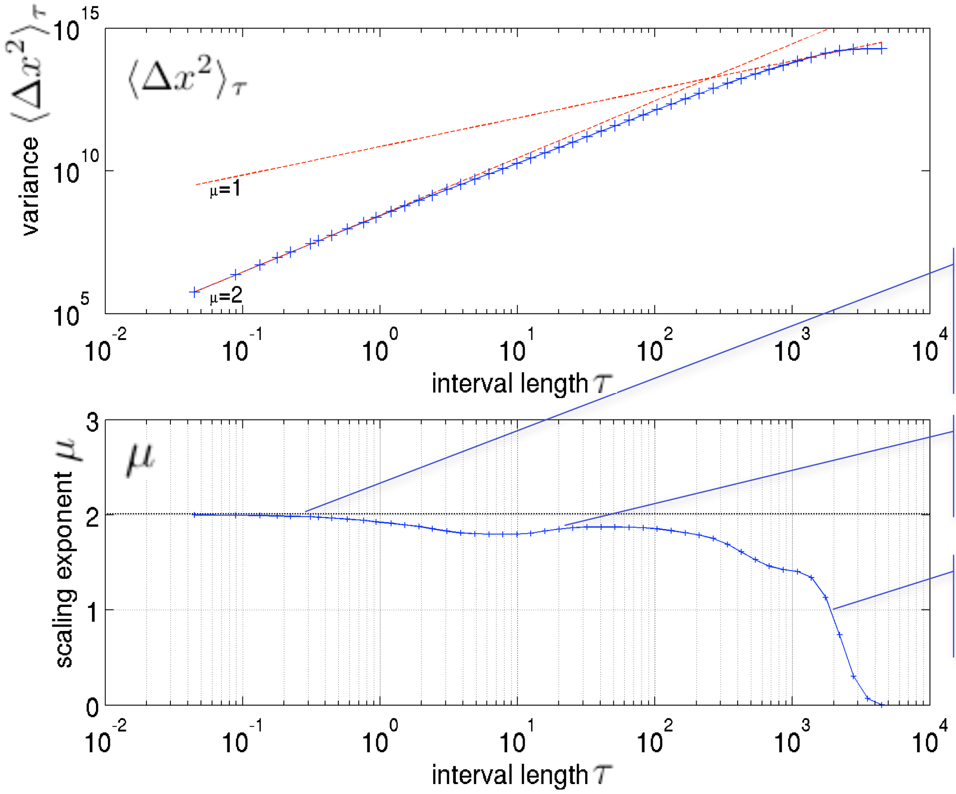
Ion motion

Let $\langle \Delta x^2 \rangle_\tau$ be the mean squared distance travelled by an ion after a time τ

- For diffusion (Brownian motion) $$\langle \Delta x^2 \rangle_\tau = 2\ D\ \tau$$
- For convection (ballistic motion) $$\langle \Delta x^2 \rangle_\tau = v^2 \tau^2$$
- More generally, we have $$\langle \Delta x^2 \rangle_\tau = \alpha\ \tau^\mu$$

Regimes with $1 < \mu < 2$ are particularly interesting, for the transport may then be neither ballistic, nor diffusive (anomalous transport)

Ion motion

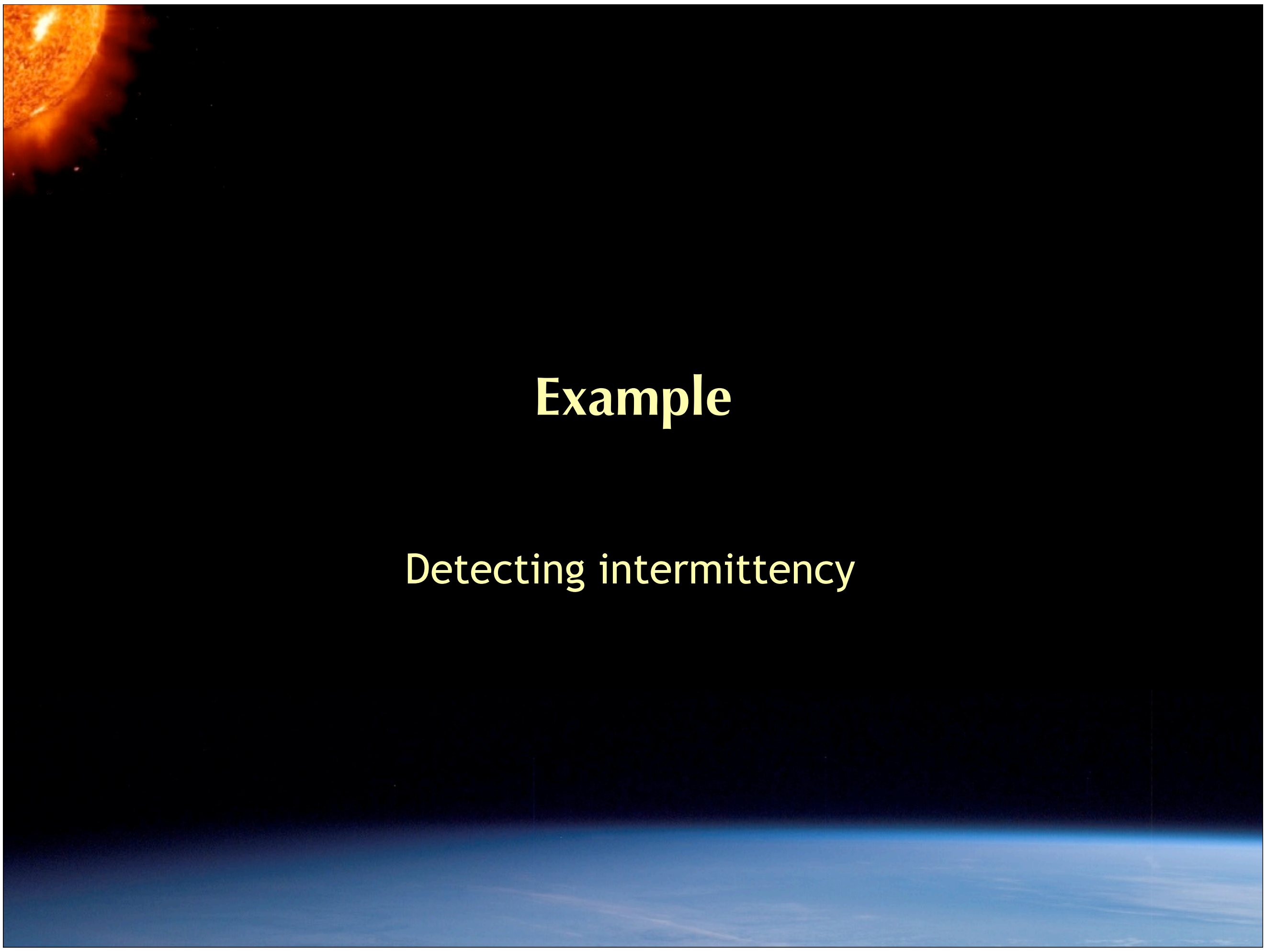


Application to ion motion in turbulent wavefield

motion below gyrofrequency is ballistic

motion is superdiffusive

finite sample size effects !



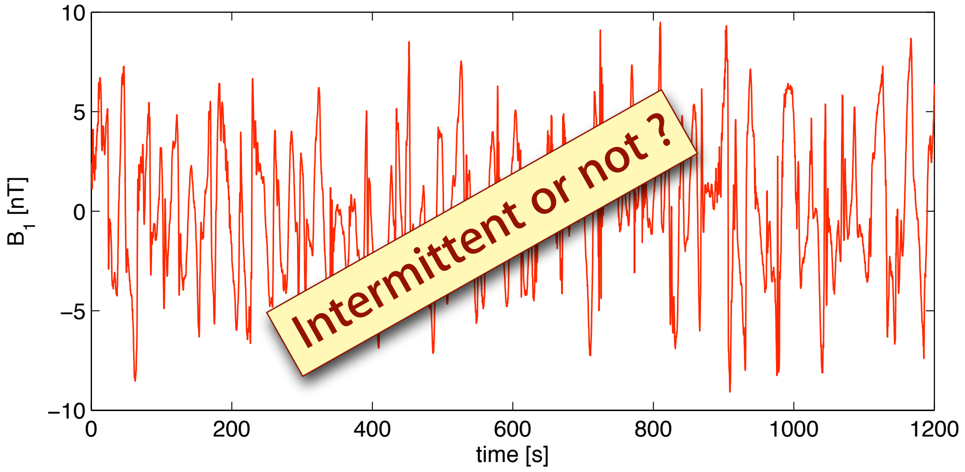
Example

Detecting intermittency

Detecting intermittency in flows

- Here ***intermittency*** = aternation of regimes with normal spectral content, and regimes with significant ***excess of energy*** in a given range of scales

- ***Example*** : magnetic field measurements upstream the Earth's bow shock



Intermittent or not ?

Intermittency

- To quantify the anomalous energy content, we define the ***local intermittency measure*** (LIM)

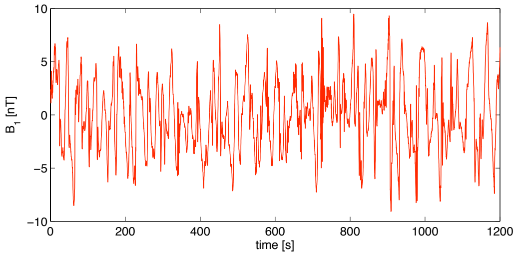
$$\gamma(t,\tau) = \frac{|X(t,\tau)|^2}{\int_{t_{min}}^{t_{max}} |X(t',\tau)|^2 \, dt'}$$

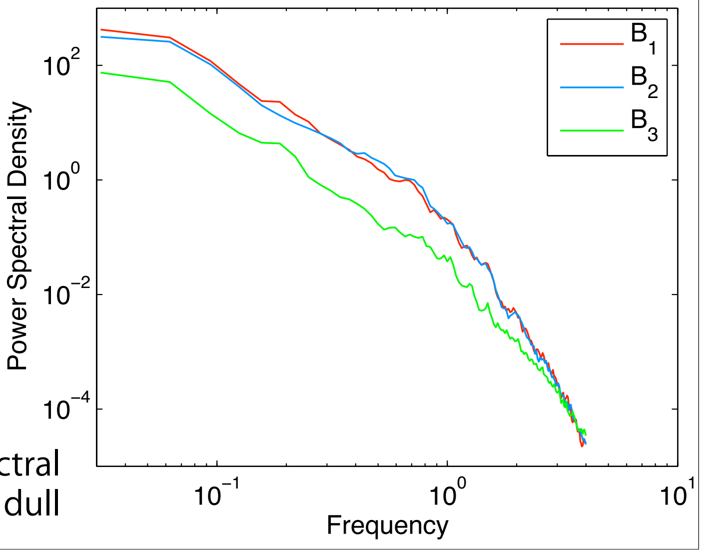
Any value of LIM that significantly exceeds 1 implies a local excess of energy

We then need to do a statistical test (chi-square)

Intermittency

- ***Example from magnetic field at bow shock***: Fourier spectral analysis doesn't reveal anything unusual

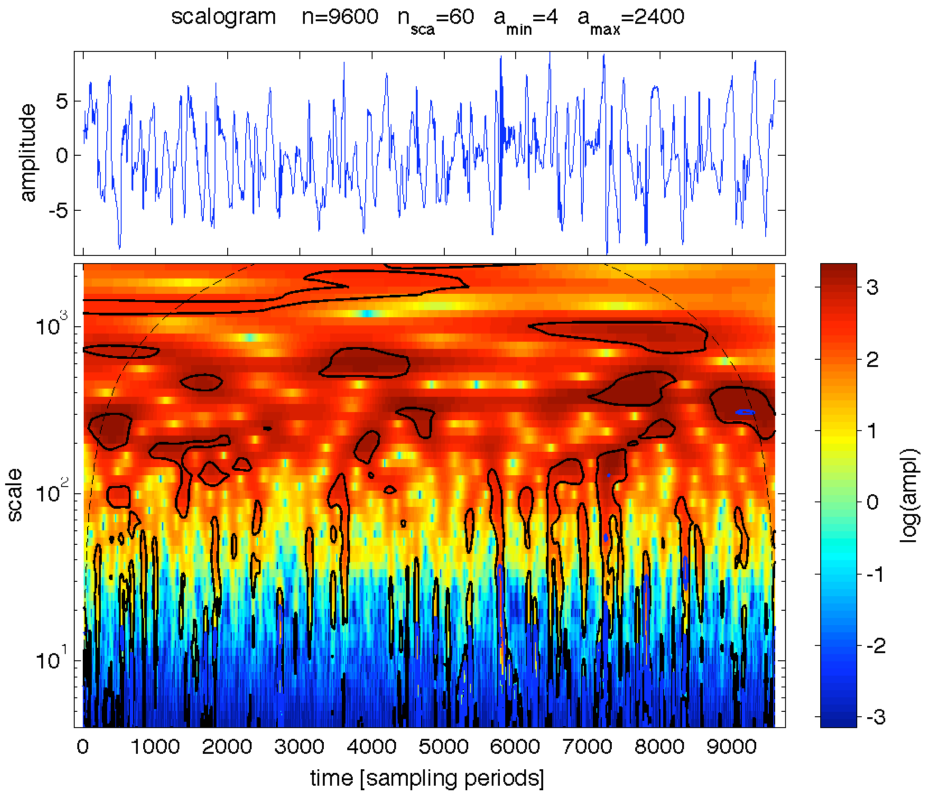




The power spectral density look rather dull

Intermittency

A wavelet scalogram reveals the presence of structures at various scales, but is it really that intermittent ?

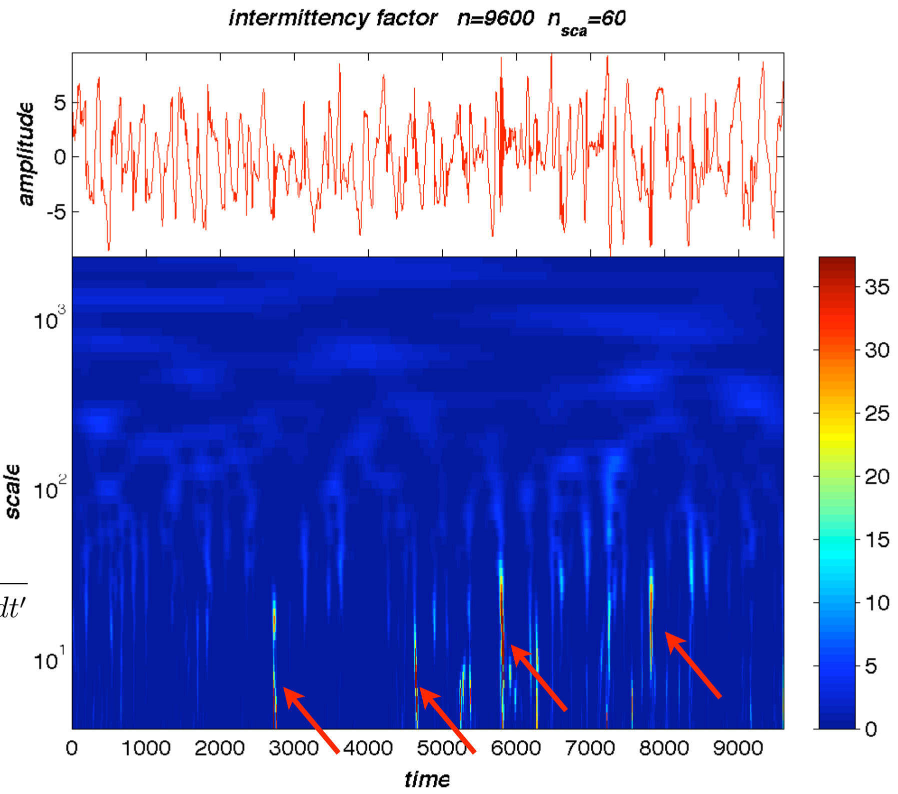


Intermittency

The local intermittency measure reveals clear bursts

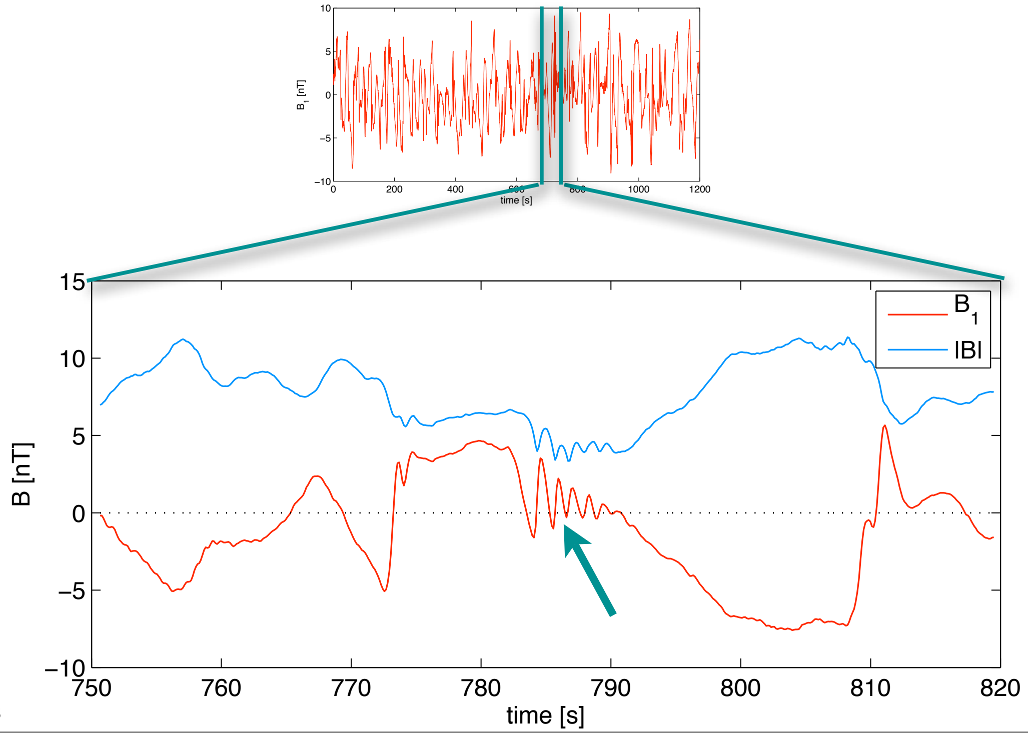
What are these ?

$$\gamma(t,\tau) = \frac{|X(t,\tau)|^2}{\int_{t_{min}}^{t_{max}} |X(t',\tau)|^2 \, dt'}$$



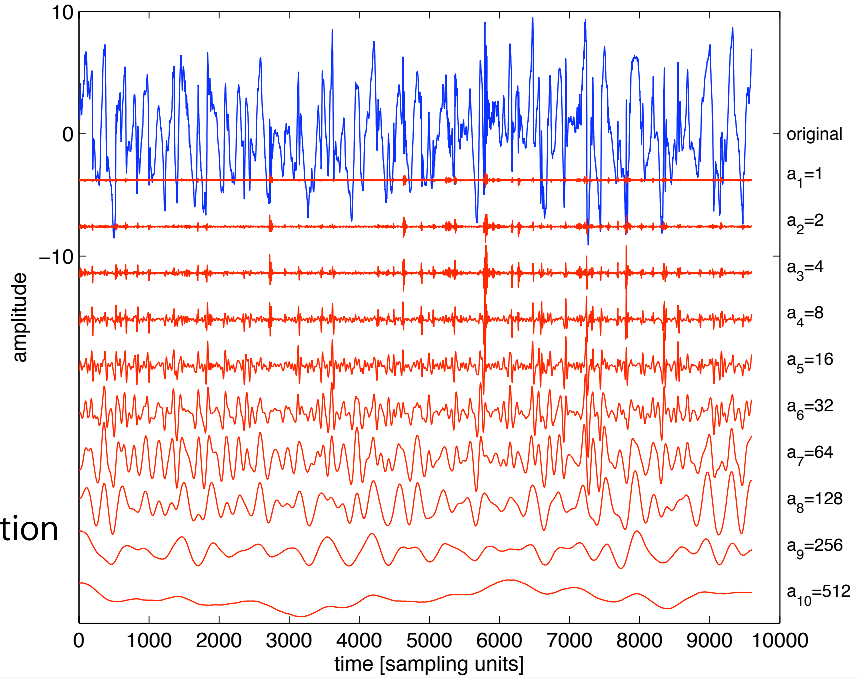
Intermittency

- The scalogram reveals the existence of occasional bursts of energy



Intermittency

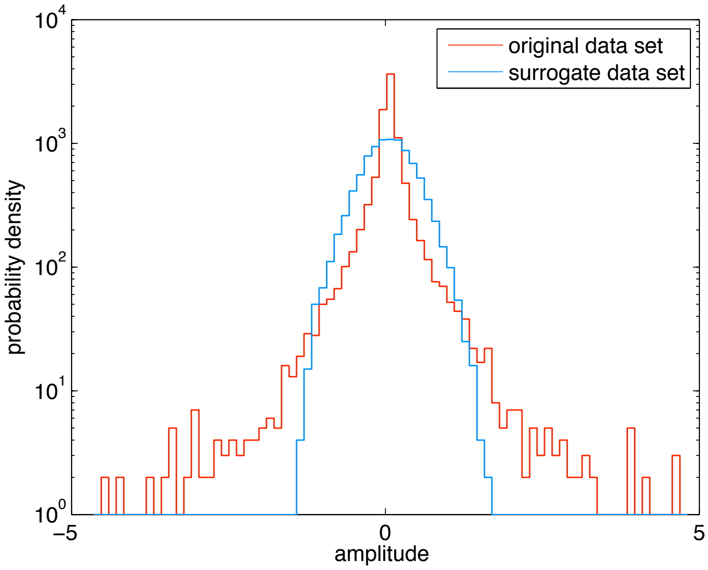
- This bursty behaviour can be observed both in the continuous and in the discrete wavelet transform



Multiscale decomposition of the magnetic field

Intermittency

- Since the wavelet transform is linear, many interesting mathematical property remain valid

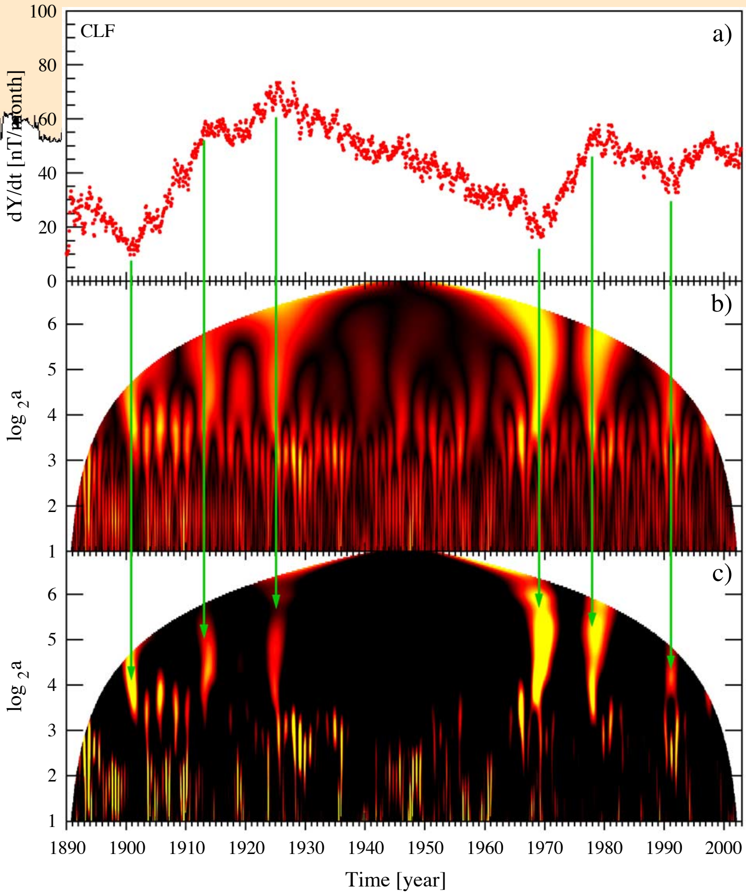


probability density function of wavelet coefficients (DOG wavelet) at scale=8

Example

The same concept can be used to detect changes :

monthly mean of eastward component of geomagnetic field
(de Michelis et al., 2003)



Bx

wavelet transform

LIM

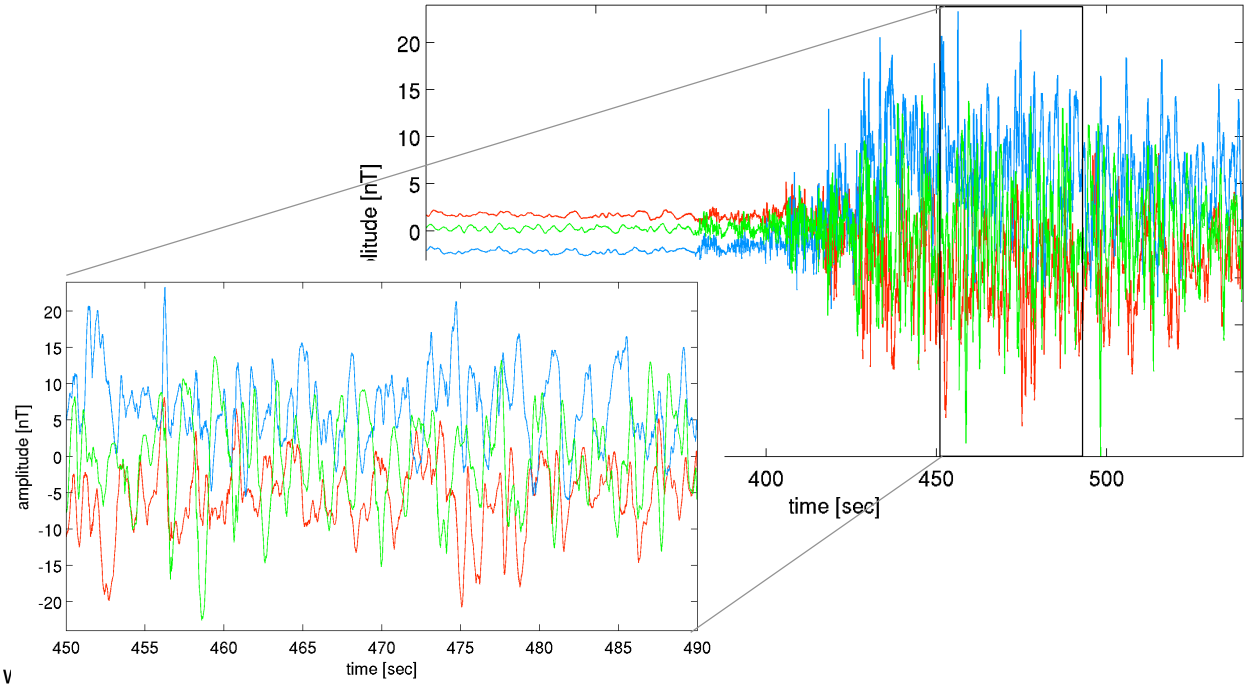
Plate I. Monthly mean values of the geomagnetic secular variation for the eastward component (*Y*) measured at Chambon la Forêt (CLF) observatory from 1890 to 2002 (panel a). Wavelet transform absolute value obtained using Alexandrescu et al. [3] methodology (panel b). Results from LIM analysis using the condition LIM$_{a,b}^2$>3 (panel c).

Example

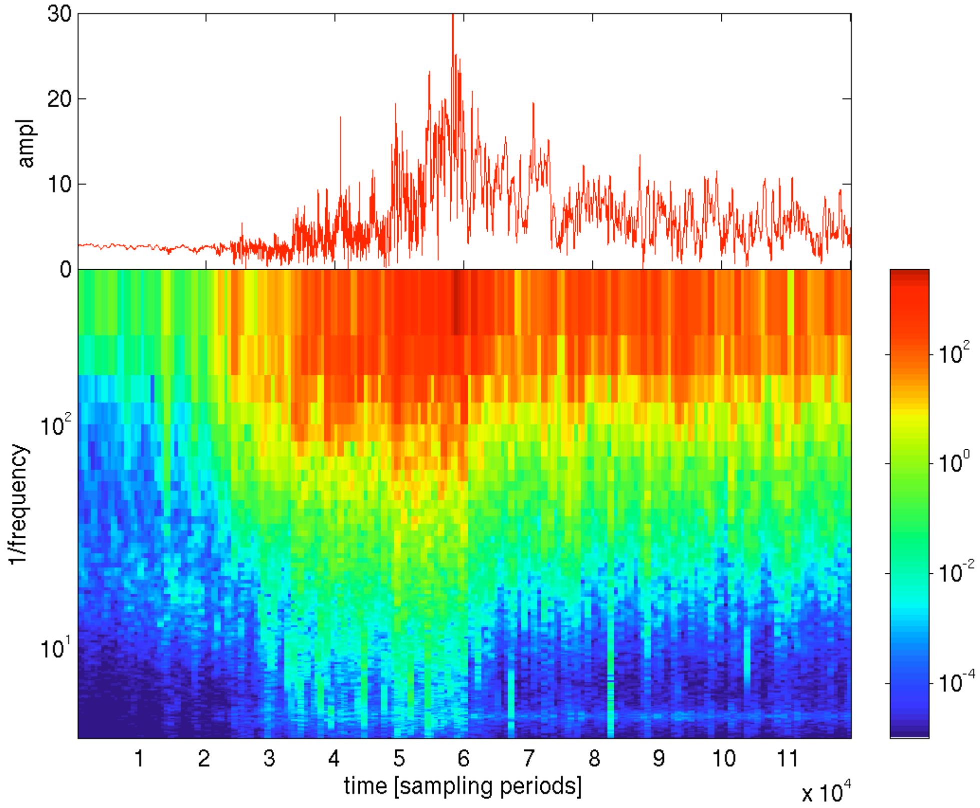
Identifying polarised wave packets
in AC magnetic field measurements

Magnetic field fluctuations

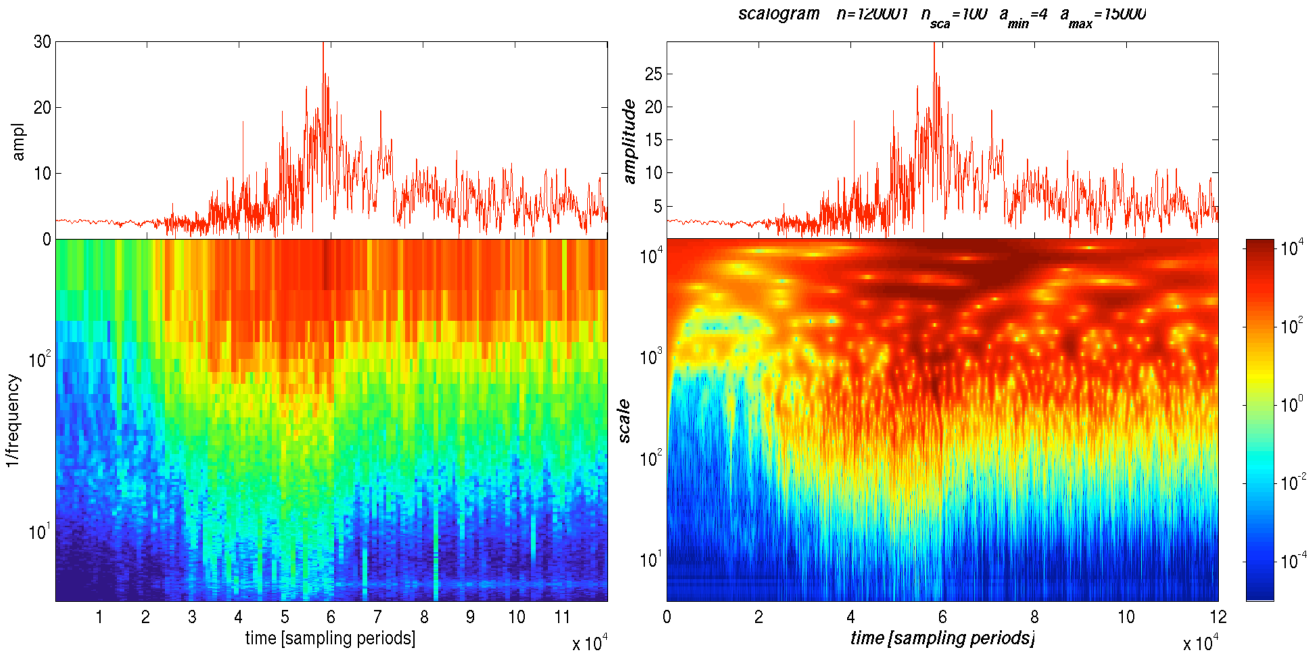
- AC magnetic field measurements at a quasiperpendicular bow shock (STAFF instrument)



STAFF spectrogram



Spectrogram vs scalogram



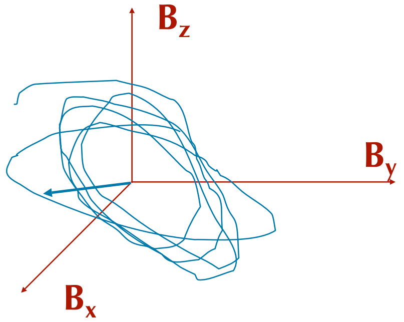
spectrogram (Fourier)

scalogram (Morlet wavelets)

there is nothing really special about the

Polarisation

- Since spectral characteristics alone are not very appealing here, we look for the polarisation vs time and frequency



- Compute

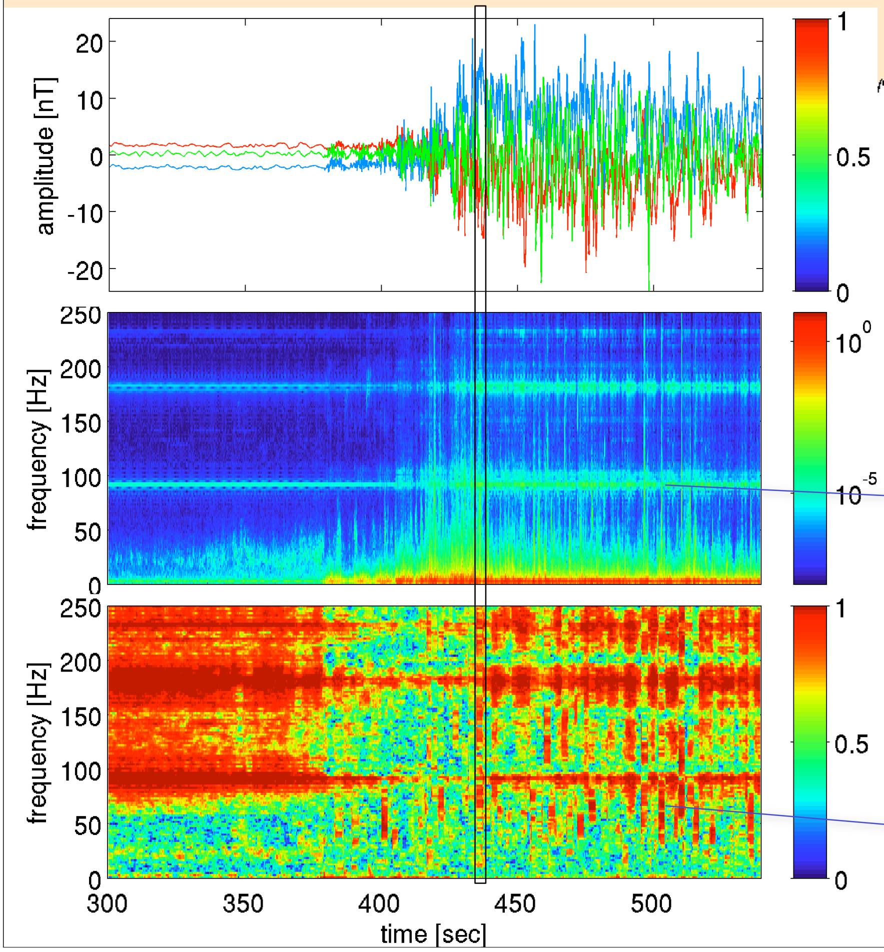
$$S_{i,j} = \langle B_i(\omega)^* B_j(\omega) \rangle \quad \text{for} \quad i,j = x,y,z$$

- Diagonalise

$$\mathsf{S}(\omega) = \begin{pmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{pmatrix}$$

- The polarisation is

$$0 \leq p \leq 1 \qquad p = 1 - \sqrt{\frac{\lambda_{min}}{\lambda_{max}}}$$



Time series

Spectrogram of B_y

interference noise

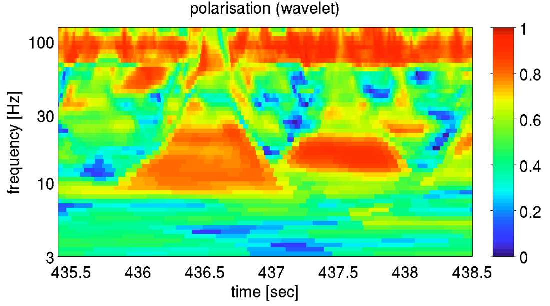
Polarisation

are these patches meaningful ?

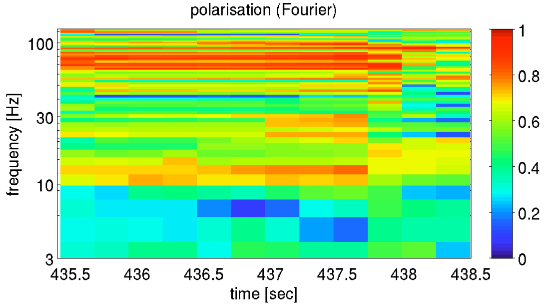
Fourier vs wavelet

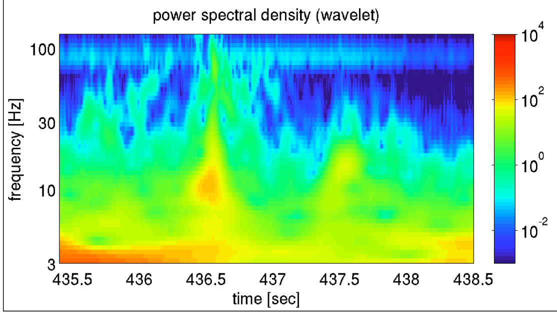
Wavelet analysis

Fourier analysis

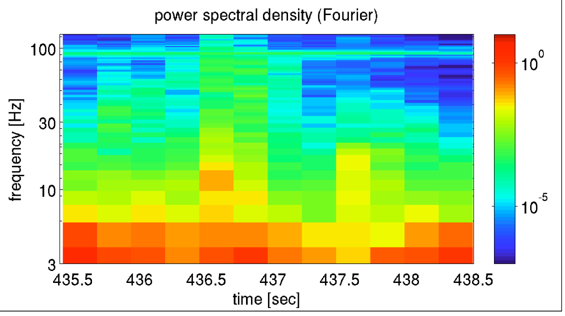

polarisation (wavelet)

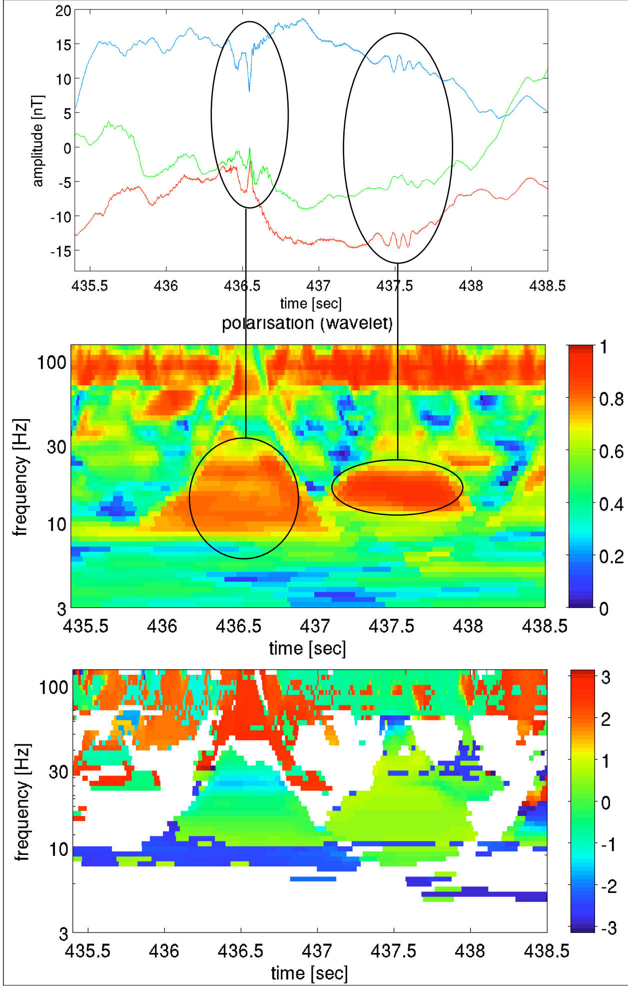
Polarisation


polarisation (Fourier)


power spectral density (wavelet)

Power spectral density


power spectral density (Fourier)



Time series

Polarisation

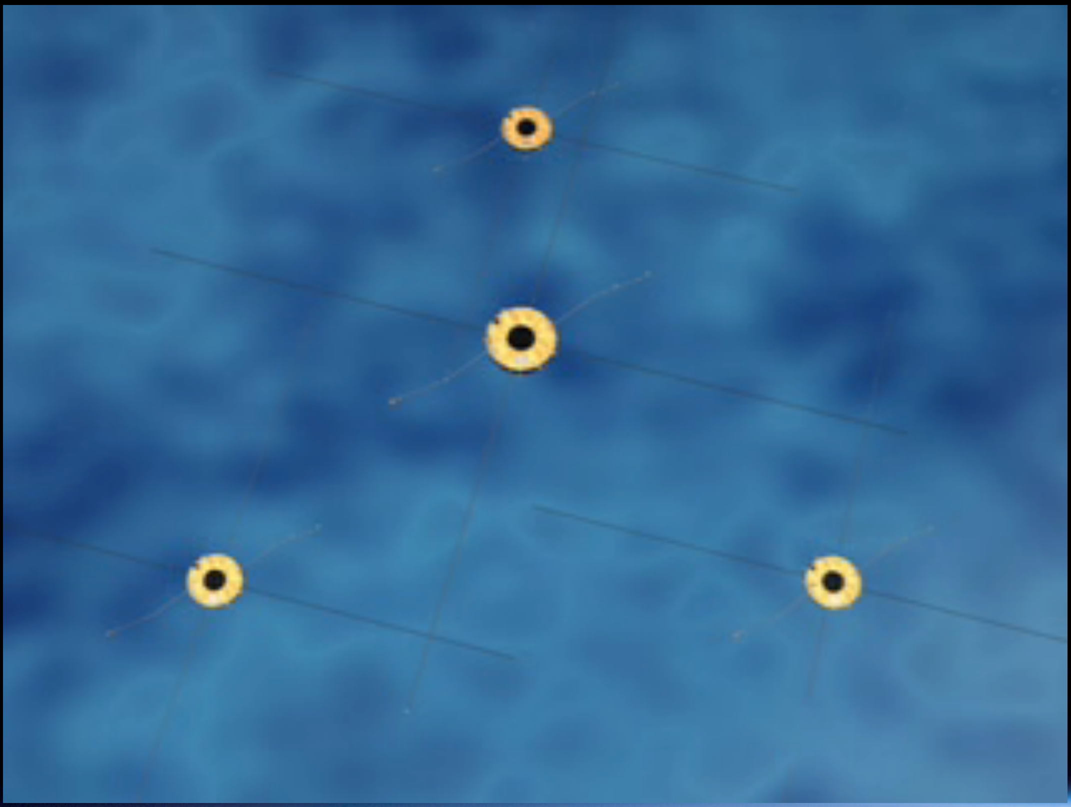
Angle of polarisation plane [rad]

Conclusion

- Wavelets are better suited for the characterization of **wave packets**
- This is important for turbulent wavefields in which different types of waves may coexist
- But the compromise between time and frequency resolution cannot be avoided

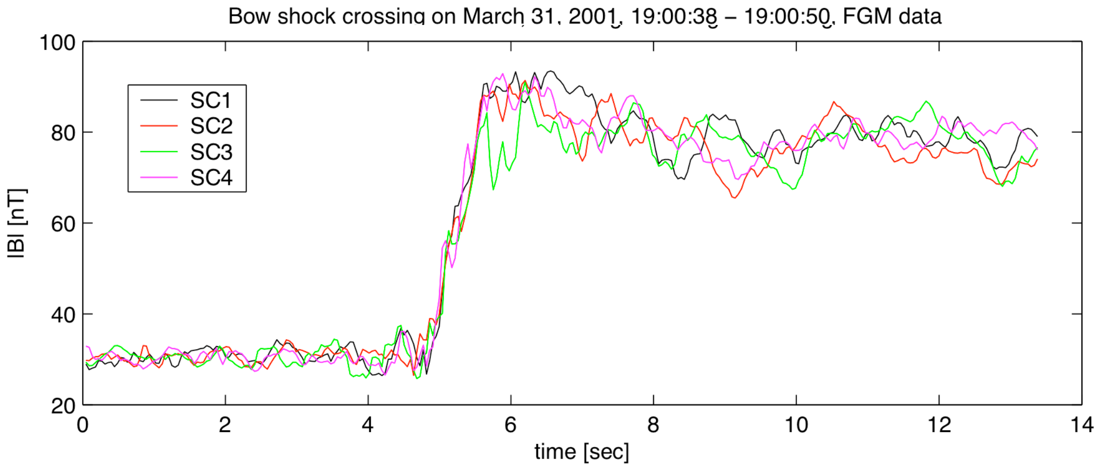
Example

Automated event timing multipoint data



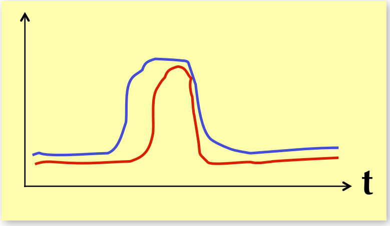
Timing

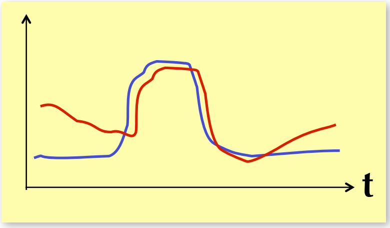
- A major fraction of multisatellite data analysis involves timing of specific events on the 4 satellites
- Timing = when does a given event occur on the 4 spacecraft ?

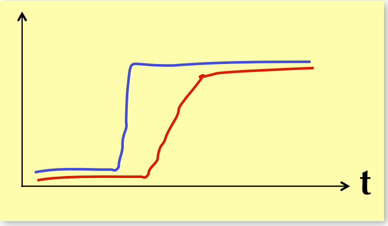


Timing

- But accurate timing can become quite difficult when one does not observe the same « pattern » on all four spacecraft
- Examples where there is no clear solution





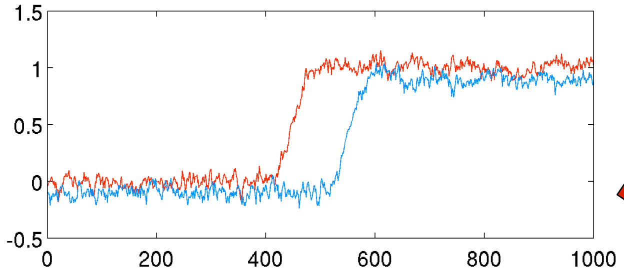


- This becomes a pattern identific

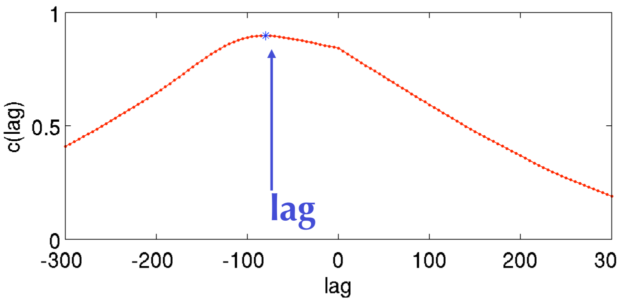
Standard approach

- The usual approach involves computation of the cross-correlation function

$$c(\tau) = \sum_{t=-T}^{T} \frac{[u_1(t) - \bar{u}_1][u_2(t-\tau) - \bar{u}_2]}{\sigma_{u_1}\sigma_{u_2}}$$



- And looking for the lag τ that maximizes $c(\tau)$



Standard approach

$$c(\tau) = \sum_{t=-T}^{T} \frac{[u_1(t) - \bar{u}_1][u_2(t-\tau) - \bar{u}_2]}{\sigma_{u_1}\sigma_{u_2}}$$

This cross-correlation approach has many drawbacks

- The width T of the interval greatly influences the result. How should its value be chosen ?
- It is biased by large scales
- It is easily misled by offsets and trends

We use instead a multiresolution approach that has been developed in the frame of stereoscopic vision

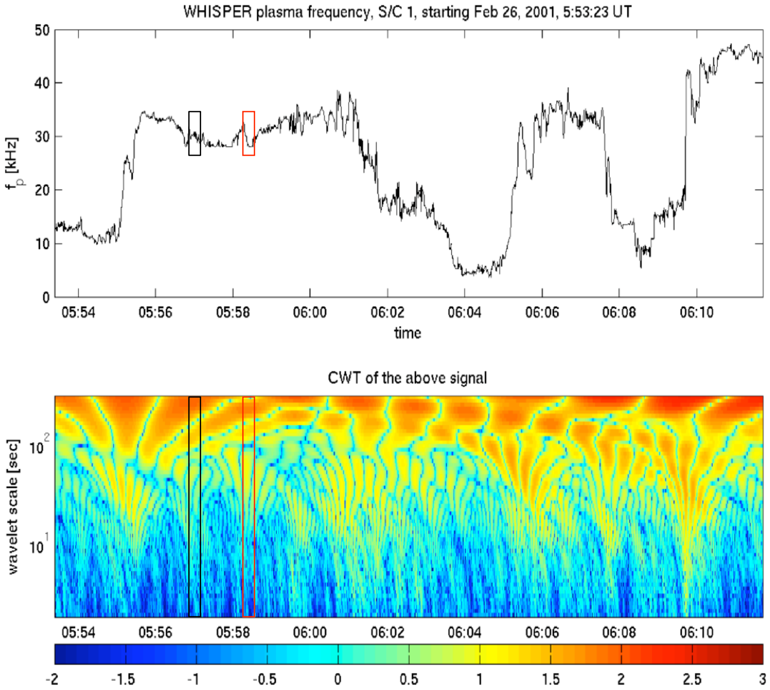
(Perrin & Torresani, 2001)

Multiresolution approach

At each time t, the wavelet transform of u(t) uniquely describes u(t) in a local neighbourhood : amplitude, derivative, texture, ...

➔ similar patterns should therefore have similar wavelet coefficients

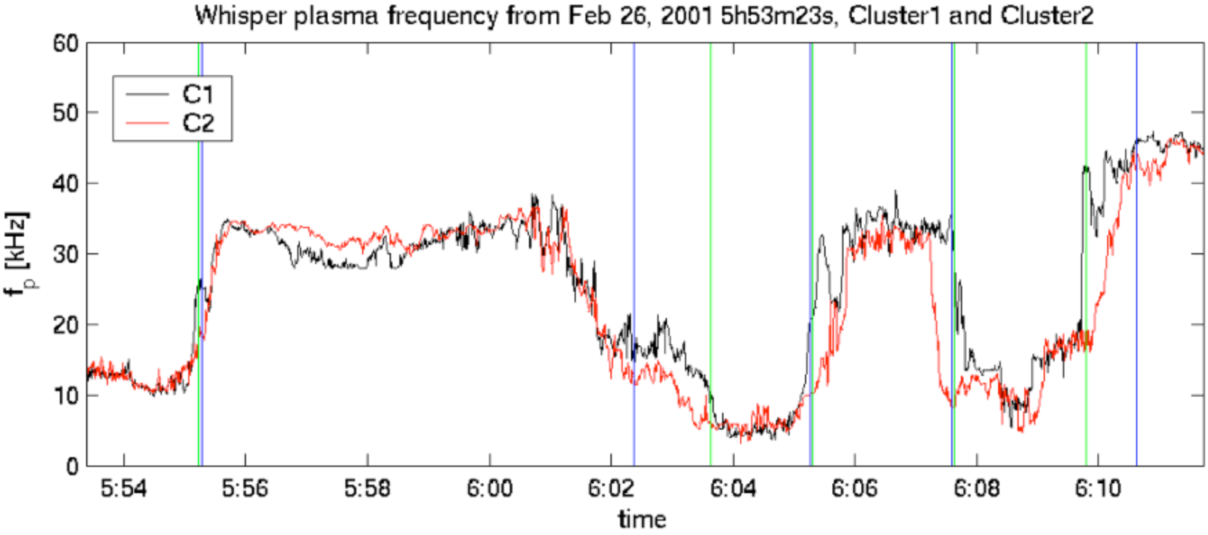
➔ instead of comparing the time series, we compare their wavelet coefficients



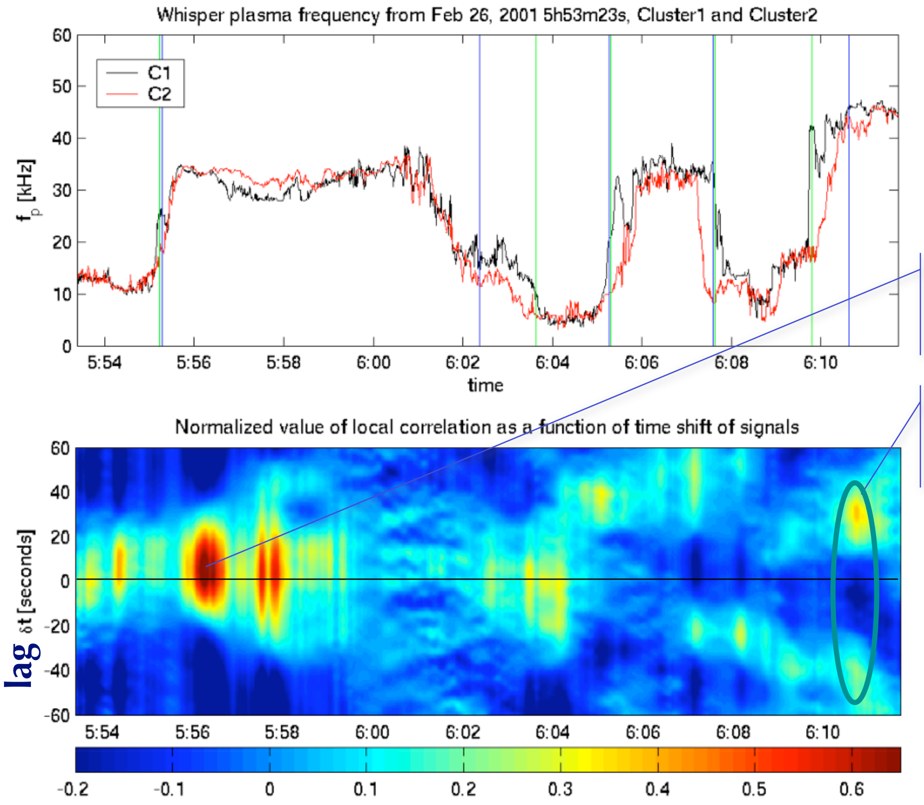
Example

A test case : electron density measurements

(Soucek et al., 2004)



Multiresolution approach



Time series

lag is well defined

ambiguous answer

Cross-correlation between their wavelet coefficients

Why is this approach better ?

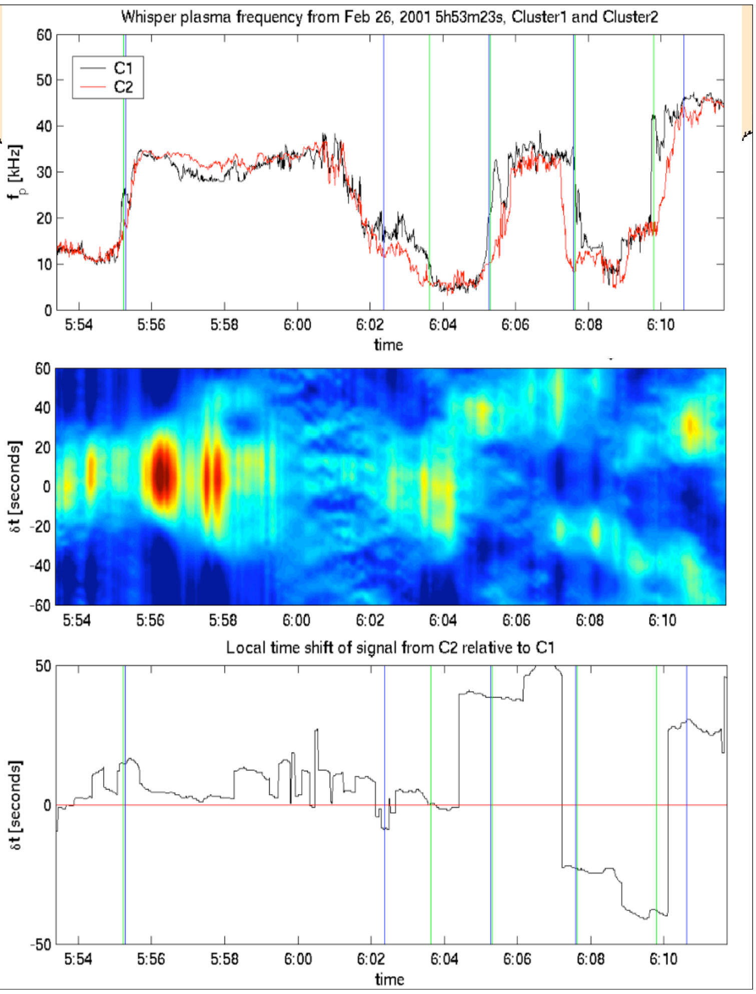
Multiresolution analysis offers several advantages :

good resolution : its resolution can be better than the sampling period (for correlation analysis, it is fixed by the window width)

robustness : it can handle patterns that do not look exactly the same

it is **data-adaptive** : the method selects itself the scales at which the corelation is the highest (like in artificial vision)

Final result



Time series

Cross-correlation between wavelet coefficients

Estimated lag

Example : comparison of EIT images

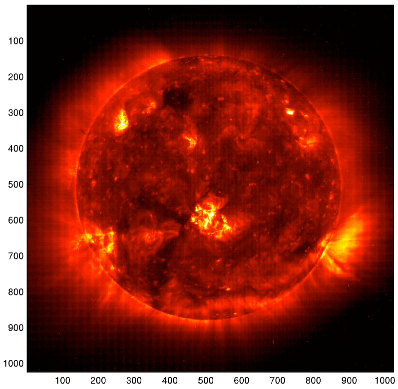


Image at 1:38 = right image

disparity due to solar rotation varies from 0 to 6 pixels

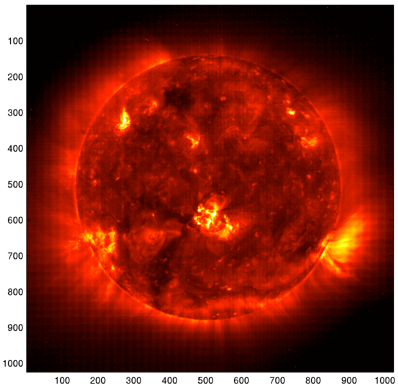
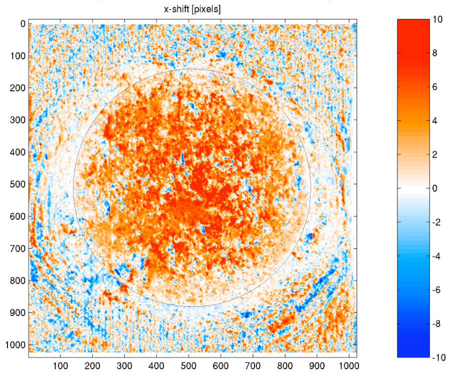


Image at 3:20 = left image



horizontal disparity between 2 images (in pixels)

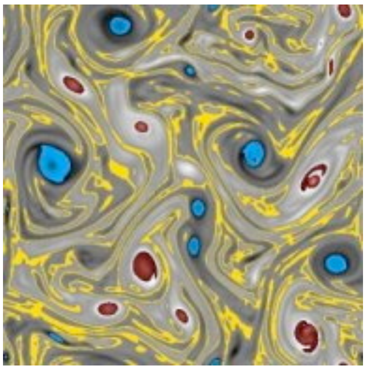
Example

Extracting coherent structures from
turbulent flows

Turbulent flows

2 key issues in the analysis of 2D (or 3D) turbulence are:

- how to identify and isolate coherent structures ?



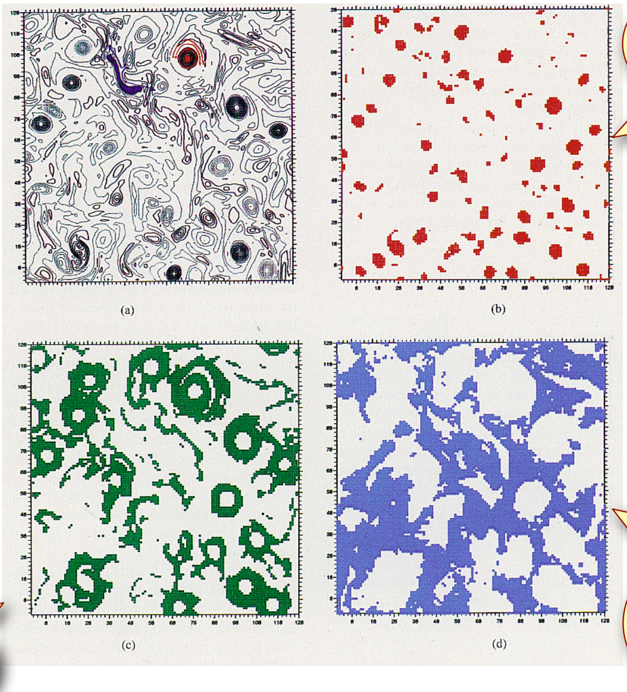
Vorticity field in 2D turbulence
(Farge et al., 1985)

- can Navier-Stokes simulations be improved using spectral methods with better basis functions than Fourier modes ?

$$\frac{\partial \vec{v}}{\partial t} + \left(\vec{v} \cdot \vec{\nabla}\right) \vec{v} = -\frac{1}{\rho}\vec{\nabla} p + \nu \nabla^2 \vec{v} + \vec{f}$$

Turbulent flows

2D numerical simulation of vorticity field (Farge et al., 1996)



elliptic regions (dominated by rotation) : coherent vortices

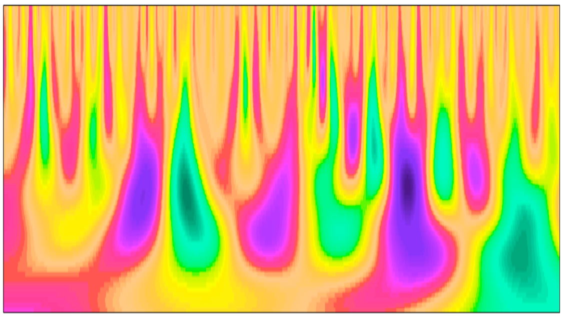
hyperbolic regions (dominated by deformation) : incoherent background

shear layers

Pitfalls

Beware of pitfalls

Example : Arneodo et al. (Nature, 1991) claimed to have seen the Richardson cascade in fully developed turbulence, simply by looking at the intricate structure of the wavelet transform



Torrence

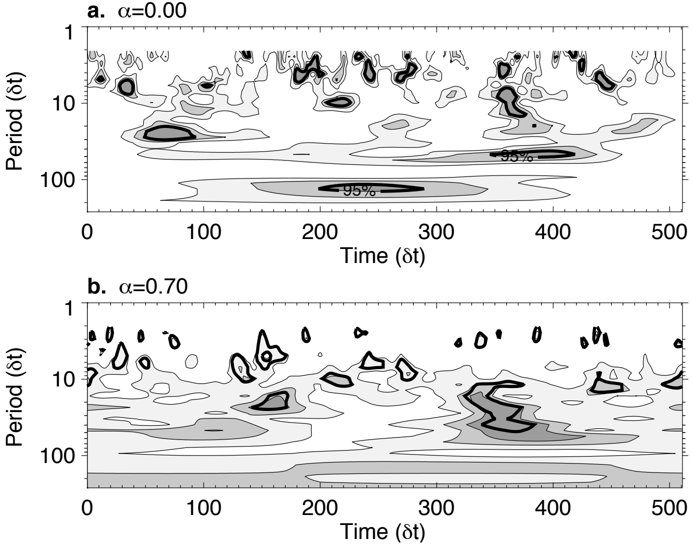
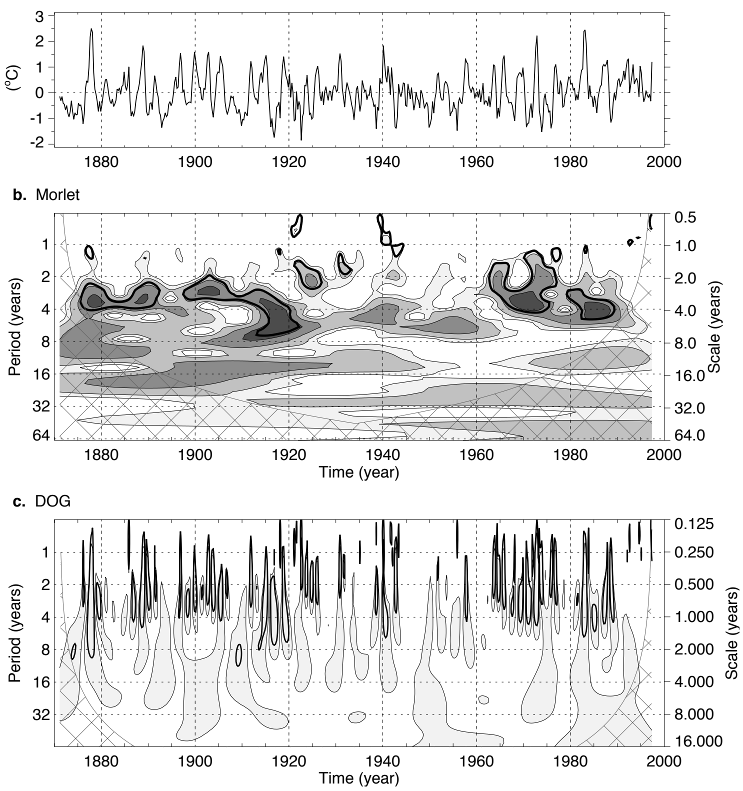


FIG. 4. (a) The local wavelet power spectrum for a Gaussian white noise process of 512 points, one of the 100 000 used for the Monte Carlo simulation. The power is normalized by $1/\sigma^2$, and contours are at 1, 2, and 3. The thick contour is the 95% confidence level for white noise. (b) Same as (a) but for a red-noise AR(1) process with lag-1 of 0.70. The contours are at 1, 5, and 10. The thick contour is the 95% confidence level for the corresponding red-noise spectrum.



Some useful references

- B. Burke Hubbard, ***The world according to wavelets: the story of a mathematical technique in the making*** (Peters, 1998): a friendly introduction
- I. Daubechies, ***Ten lectures on wavelets*** (SIAM, 1992): more mathematical, a classic
- S. Mallat, ***A wavelet tour of signal processing*** (Academic Press, 2002) : excellent reference on wavelets
- J.-L. Starck & F. Murtagh, ***Astronomical image and data analysis*** (Springer, 2006): wavelets for astronomical images
- ***Wavelets and turbulence***: lots of articles by M. Farge, K. Schneider, et al., see `http://wavelets.ens.fr/`
- G. Paschmann, et al. ***Analysis Methods for Multi-Spacecraft Data*** (ISSI, Bern, 2000), can be downloaded from `http://www.issi.unibe.ch/PDF-Files/analysis_methods_1_1a.pdf`