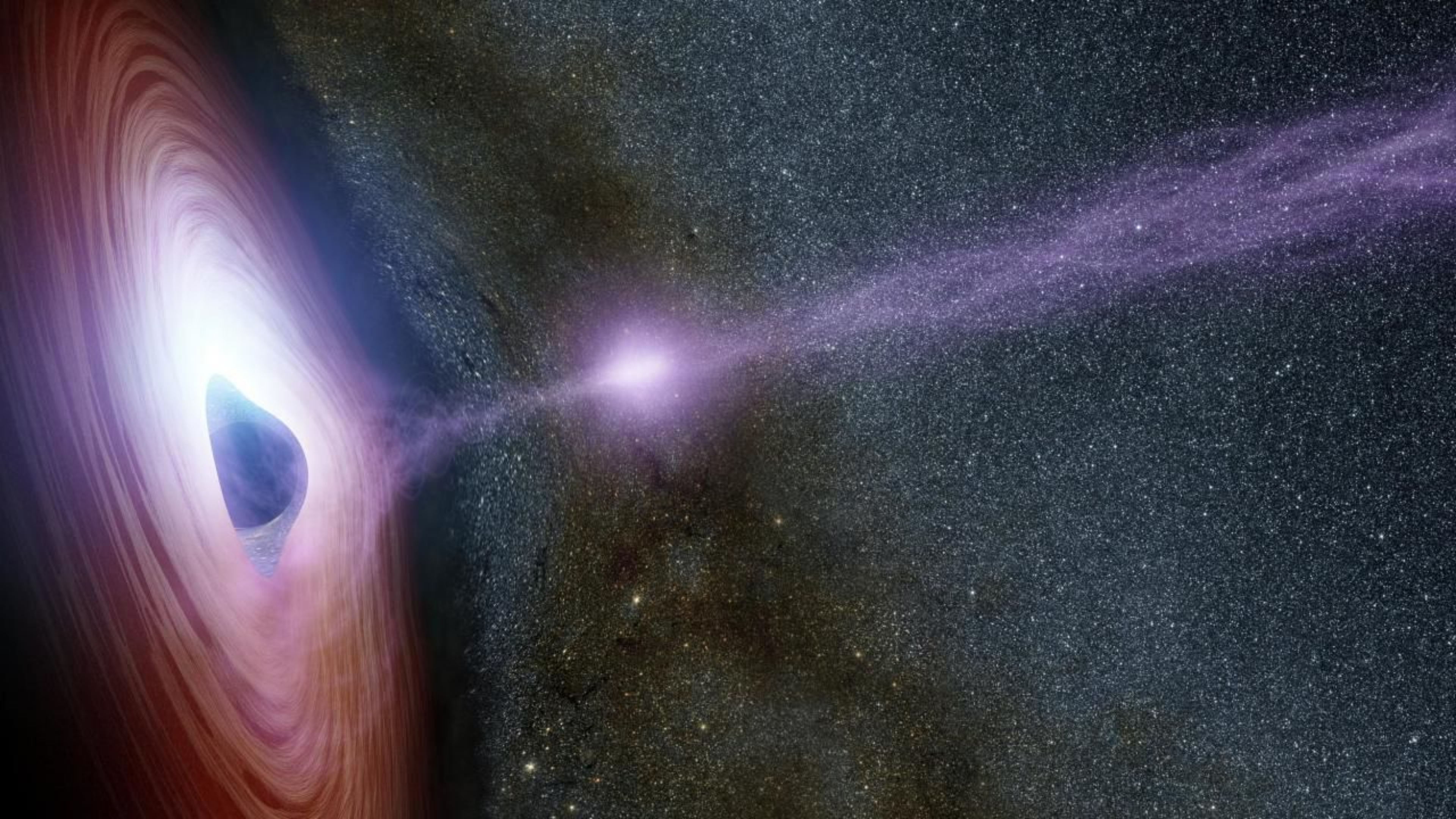
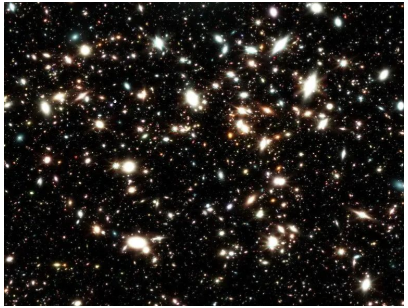
The Hilbert-Huang Transform

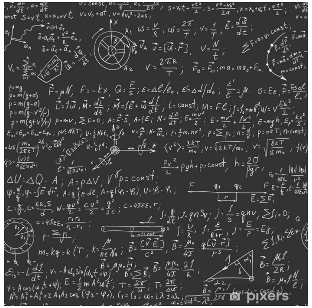
Isabella G. Pretto
California Institute of Technology

ARC Seminar
January 23, 2024

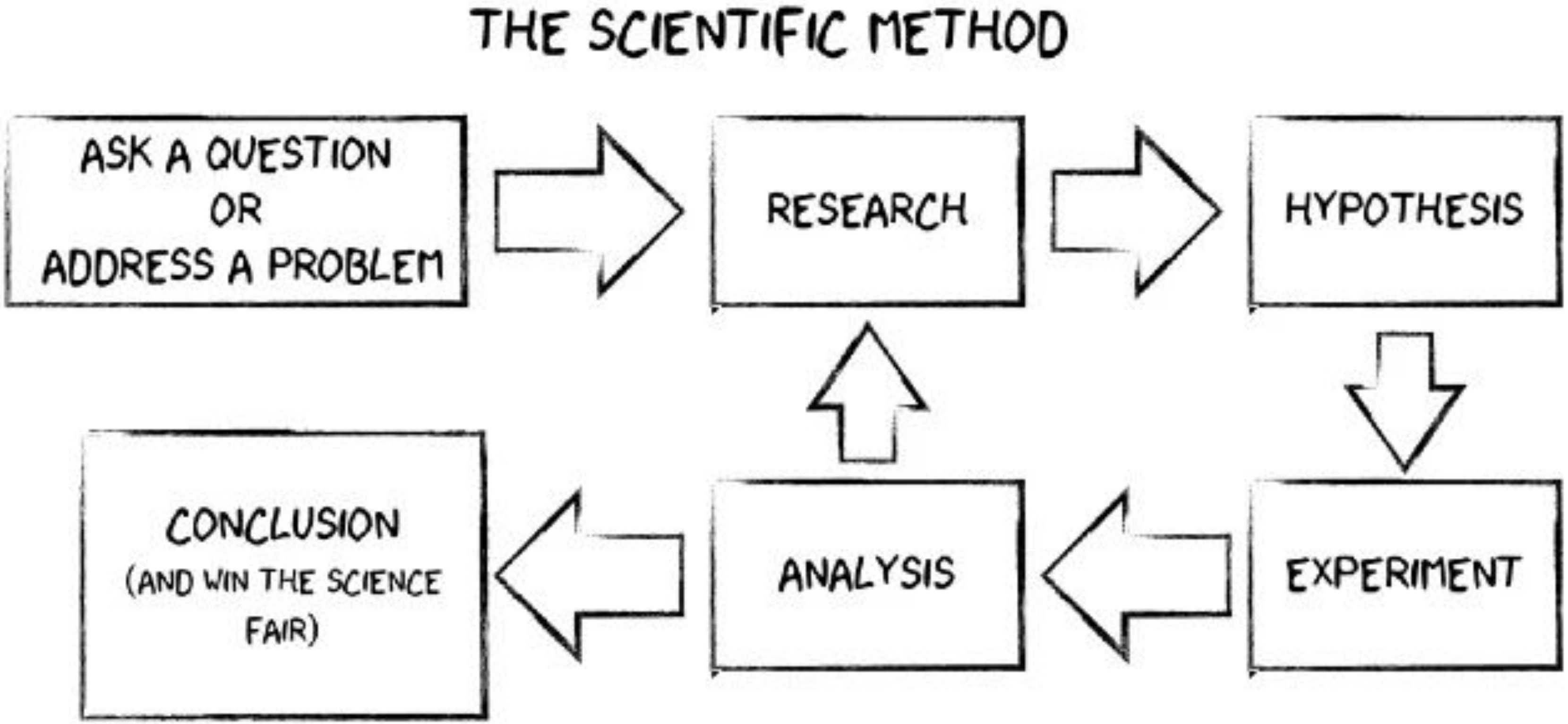




?



Isabella G. Pretto



Isabella G. Pretto

What’s up?

The need for an adaptive basis

- Traditionally, assume stationary and/or linearity
- Real world?

- A priori basis
- Fourier transform: trigonometric functions
- Convolution computation

- A posteriori basis
- Adaptive — let the data speak for themselves!
- Arbitrary selection of basis ↔ representation of a physical process

The Hilbert-Huang Transform (HHT)

- Designed for linear and nonstationary data
- HHT = Empirical Mode Decomposition (EMD) + Hilbert Transform
- Data $\xrightarrow{\text{EMD}}$ Intrinsic Mode Functions (IMFs)
- Physically meaningful instantaneous frequency and amplitude

The Hilbert Spectral Analysis

- To deal with nonstationary/nonlinearity, allow for instantaneous frequency and amplitude
- For the time series x(t), define the Hilbert transform as

$$y(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{x}{(t-\tau)} d\tau,$$

- Define a complex signal

$$z(t) = x(t) + iy(t) = a(t)e^{i\theta(t)}$$

$$\theta(t) = \tan^{-1}(\frac{y}{x})$$

$$a(t) = (x^2 + y^2)^{1/2}$$

The Hilbert Spectral Analysis

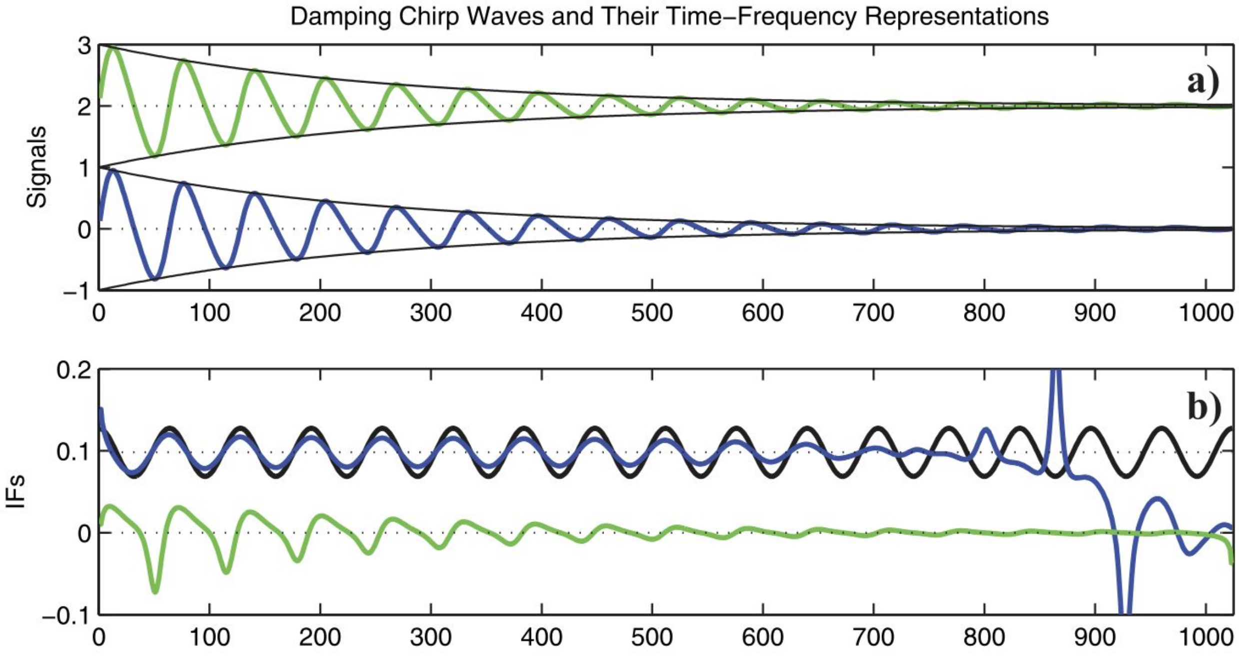
- Instantaneous frequency

$$\omega(t) = \frac{d\theta}{dt}$$

- Energy

$$h(\omega) = \int_0^T H(\omega, t)dt,$$

Limitations



$$x(t) = e^{-t/256} \sin\left[\frac{\pi t}{32} + 0.3 \sin\left(\frac{\pi t}{32}\right)\right]$$

Limitations

- Huang: need a purely oscillatory function with a zero reference level for the Hilbert transform to work

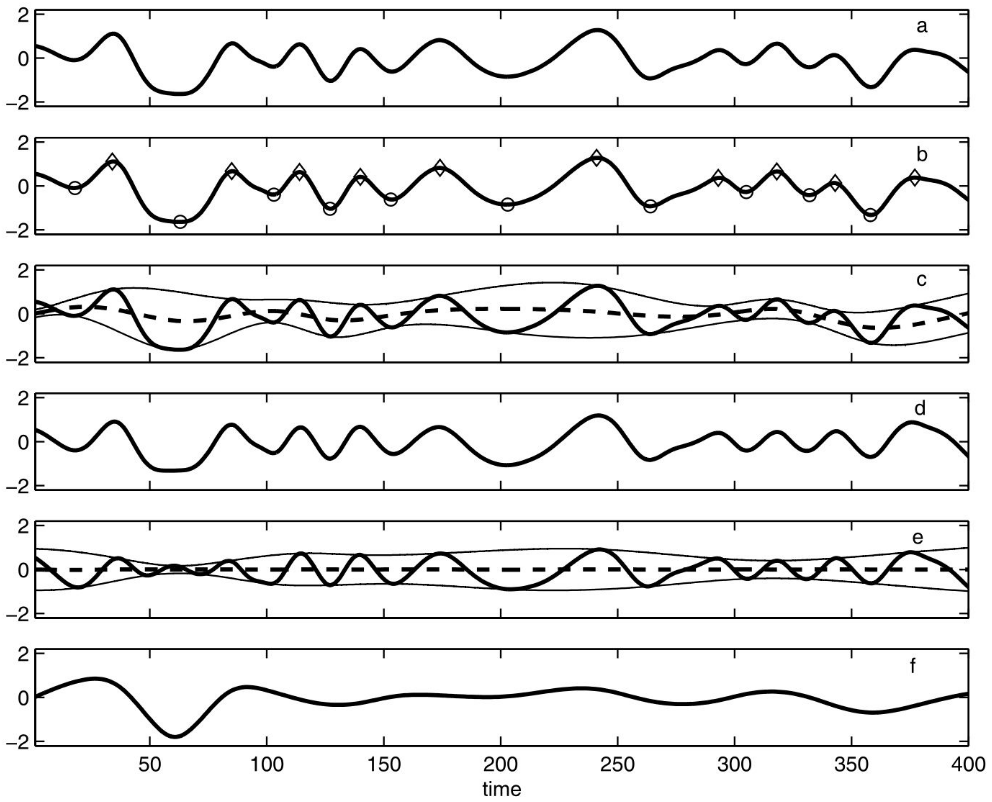
- How to decompose an arbitrary function $x(t)$ in terms of functions with meaningful Hilbert transform?

- Development of the empirical mode decomposition (EMD)

Empirical Mode Decomposition

- Assumes that the data is a superposition of oscillatory modes
- Intrinsic mode function (IMF)
  - Number of extrema and zero crossings must either be equal or differ by one
  - Mean value of the envelopes defined using the local extrema is zero
- Decompose a function through a sifting process

Sifting Process



First protomode

$$h_1 = x(t) - m_1.$$

Sifting Process

- Iterate until protomode satisfies IMF definition

$$h_{1(k-1)} - m_{1k} = h_{1k}$$

$$c_1 = h_{1k}$$

- Purpose
  - Eliminate background waves
  - Make wave profiles more symmetric

Decomposition into IMFs

$$x(t) - c_1 = r_1$$

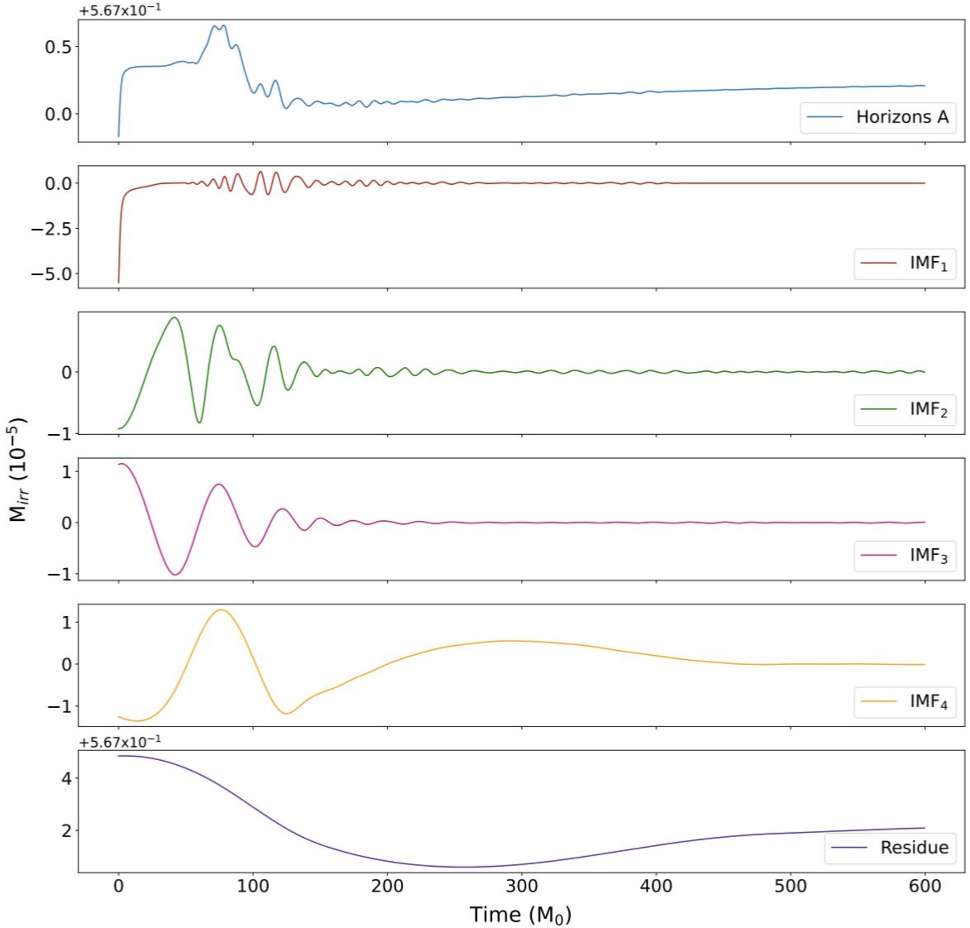
$$r_1 - c_2 = r_2$$

$$\cdots$$

$$r_{n-1} - c_n = r_n$$

$$x(t) = \sum_{j=1}^{n} c_j + r_n.$$

→ Stop when the residue becomes a function that is either monotonic or has only one extremum



Isabella G. Pretto

Stopping criterion

- When to stop sifting?
- Cauchy type of convergence test

$$\mathrm{SD}_k = \frac{\sum_{t=0}^{T} |h_{k-1}(t) - h_k(t)|^2}{\sum_{t=0}^{T} h_{k-1}^2(t)}$$

$$\mathrm{SD}_k = \frac{\sum_{t=0}^{T} |m_{1k}(t)|^2}{\sum_{t=0}^{T} |h_{1k}(t)|^2}$$

Stopping criterion

- S stoppage: stop sifting when the number of zero crossings and extrema

  (1) are equal or differ by at most 1

  (2) stay the same for S consecutive times

- Empirically determined: $3 < S < 8$

Properties of the EMD

- IMFs usually satisfy, empirically, requirements of convergence, completeness, orthogonality and uniqueness

- Express the data as

$$x(t) = \mathrm{Re}\left[\sum_{j=1}^{n} a_j(t) e^{i\omega_j(t)dt}\right]$$

  - Fourier representation

$$x(t) = \mathrm{Re}\sum_{j=1}^{\infty} a_j e^{i\omega_j t}$$

Properties of the EMD

- A ‘generalized’ Fourier transform
- Time-dependent frequency and amplitude
  - Accommodates nonlinear and nonstationary behavior
- Adaptive basis
  - Not subject to any limitations on time/frequency resolutions

Normalized Hilbert Transform

- Variations in the phase function >> variations in the amplitude

$$x(t) = a(t)\cos\theta(t)$$

- Physically meaningful instantaneous frequency should come only from the phase function

$$\mathrm{H}[a(t)\cos\theta(t)] = a(t)\mathrm{H}[\cos\theta(t)]$$

Normalized Hilbert Transform

- Bedrosian Theorem

$$H[f(t)h(t)] = f(t)H[h(t)]$$

True, only if

(1) Fourier spectra of f(t) and h(t) do not overlap, and
(2) Frequency range of the spectrum of h(t) >> frequency range of the spectrum of f(t)

Normalized Hilbert Transform

- Separate IMF into amplitude (AM) and frequency (FM) modulation parts
- Apply the Hilbert transform to the FM part only
- Algorithm

(1) Find the absolute value of the IMF

(2) Identify the maxima

(3) Define an envelope through the maxima

(4) Normalize: $$f_1(t) = \frac{x(t)}{e_1(t)}$$

Normalized Hilbert Transform

- Iterate

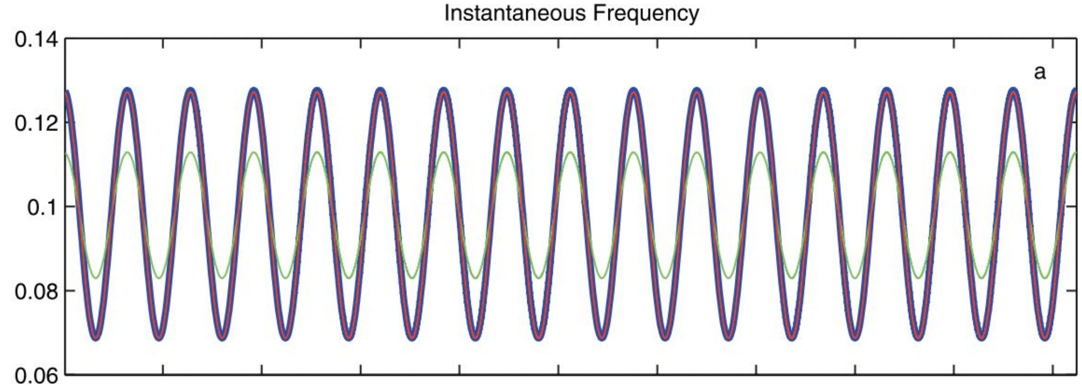
$$f_2(t) = \frac{f_1(t)}{e_2(t)}; \quad \ldots \; f_n(t) = \frac{f_{n-1}(t)}{e_n(t)}$$

- Define the FM and AM components as

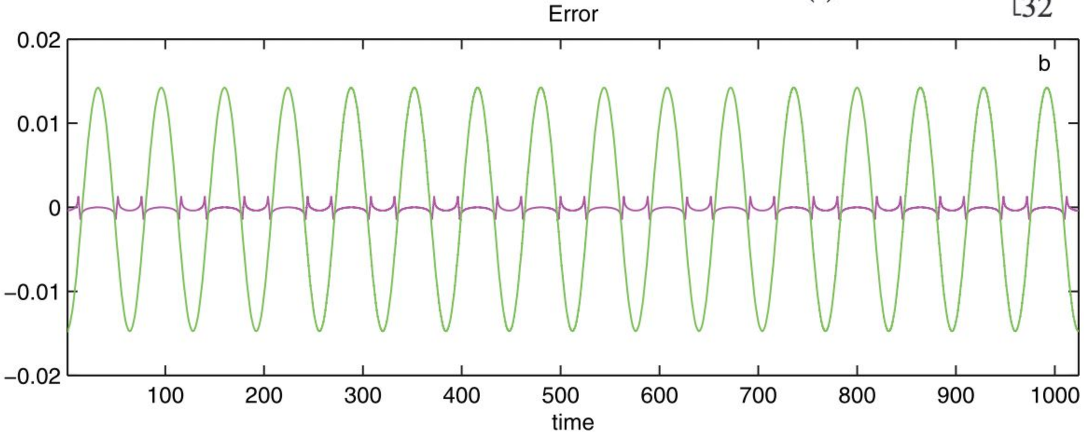
$$F(t) = f_n(t)$$

$$A(t) = \frac{x(t)}{F(t)}$$

- Normalized Hilbert transform = normalization + Hilbert transform applied to the FM part



$$x(t) = e^{-t/256} \sin\left[\frac{\pi t}{32} + 0.3 \sin\left(\frac{\pi t}{32}\right)\right]$$



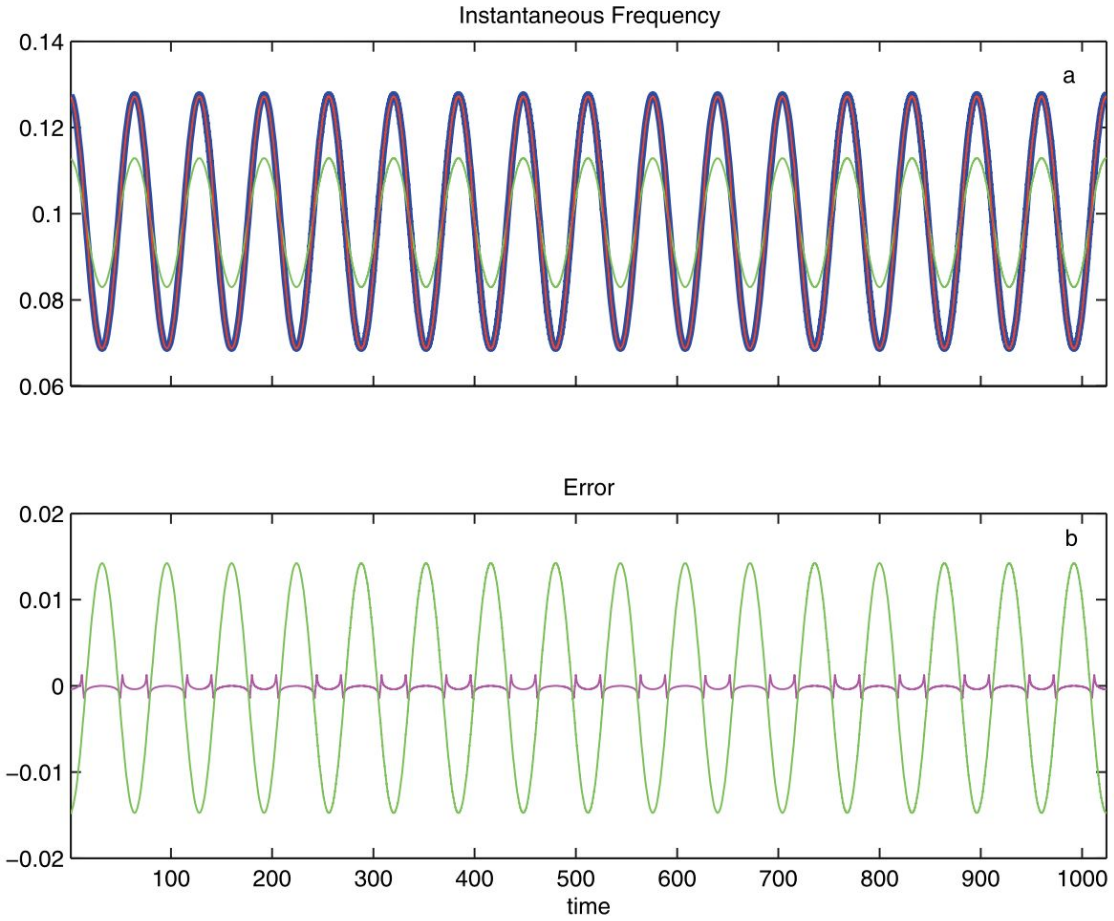
Direct Quadrature Method

- Hilbert transform underestimates the true frequency

- Direct Quadrature Method

- What if we just take advantage of the good properties of the IMF?

- Data ➜ IMFs ➜ AM + FM ➜ arccos of phase function

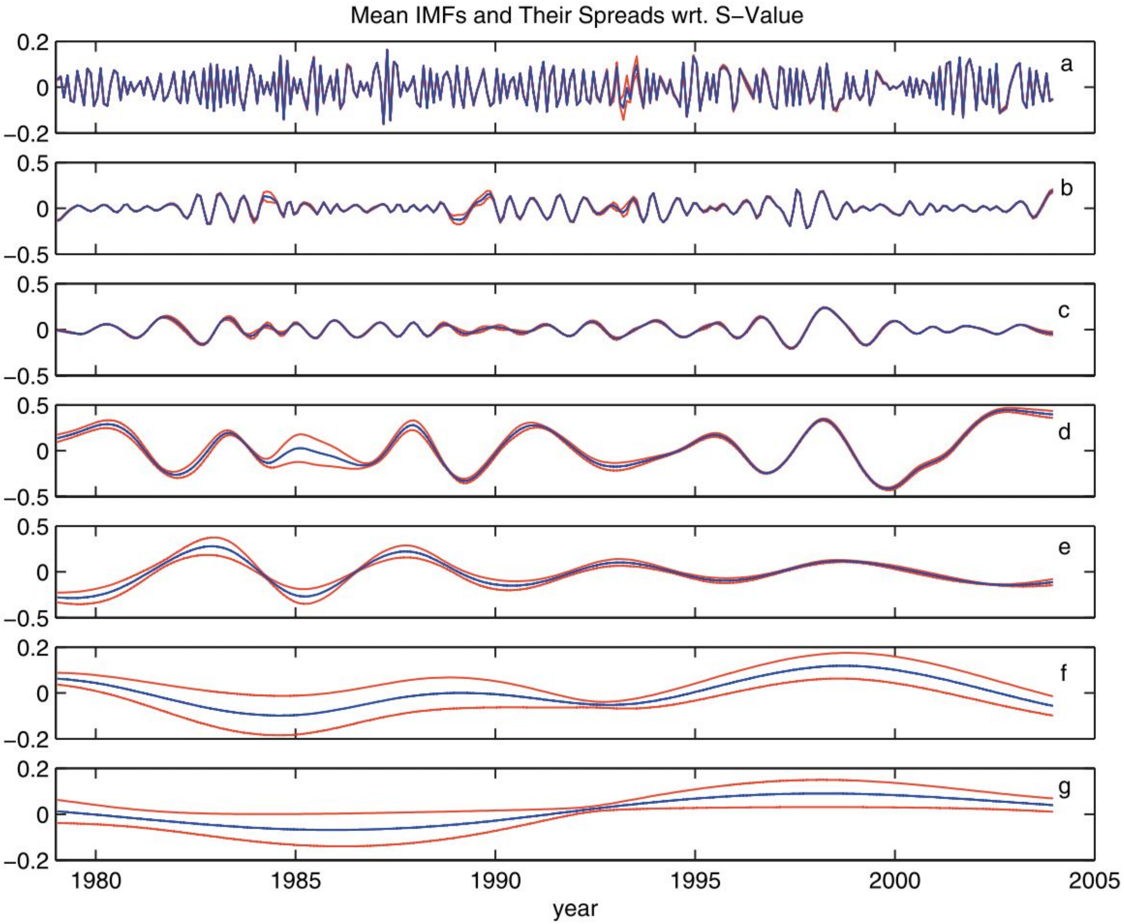


Isabella G. Pretto

Confidence Limit

- Need a criterion to measure the reliability of the decomposition
- How sensitive are the results to the stopping criterion?
- Check for ‘convergence’ of IMFs by using different values of stopping criterion
- Calculate the mean and the spread for different sets of n IMFs

4 < S < 13



Isabella G. Pretto

Statistical Significance

- Data is usually noisy

- How to know whether an IMF is a component of noise?

- Spectral properties of noise in the EMD

- *Flandrin et al., Flandrin and Golçalves*: Gaussian noise

- *Wu and Huang:* white noise

Statistical Significance

- Common properties

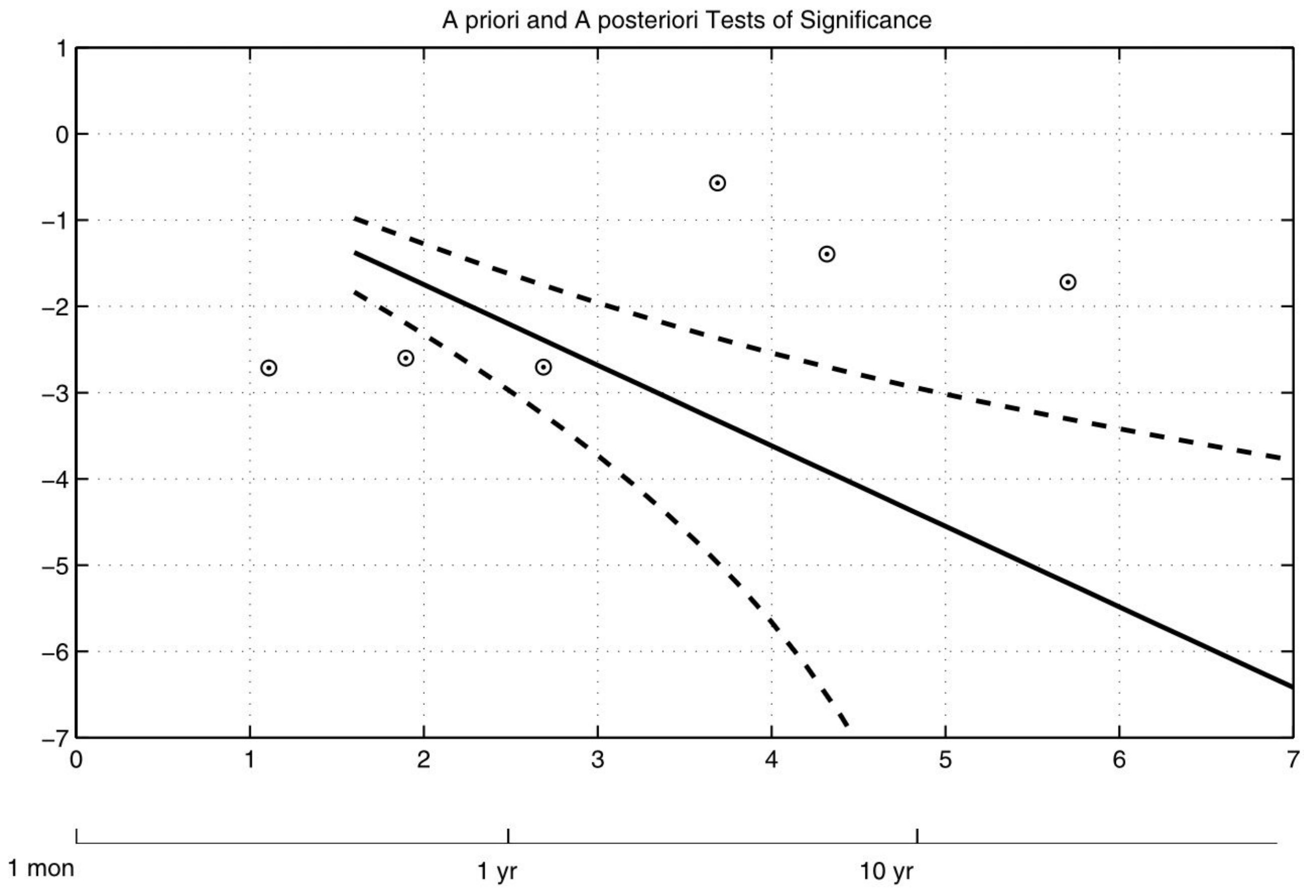
- Spectra of all IMFs, except the first one, collapse to a single shape, with the appropriate amplitude scaling

- Center frequencies of neighboring IMFs are halved

- Derive the expected energy distribution of IMFs of a noise series

- Central Limit Theorem: E follows a $\square^2$ distribution ➔ analytic form for the spread S

- Expected energy distribution + spread ➔ statistical significance test



Isabella G. Pretto

Statistical Significance

- Monte Carlo test

(1) Generate a large sample of the noise series

(2) Decompose into IMFs

(3) Find the statistical distribution of the metric of interest

(4) Compare the location of the same metric evaluated on the data with this distribution