Huang-Hilbert transform
description & application to turbulent data

J. de Patoul

Laboratoire de Physique des Plasmas
École Polytechnique, Palaiseau, France

Journal Club Meeting – October 10, 2013


For a signal $x(t) \in L^p$, the Hilbert transform is

$$H[x(t)] = \frac{1}{\pi} \lim_{\epsilon \to 0} \int_{-\infty}^{0-\epsilon} \frac{x(\tau)}{t-\tau} d\tau + \int_{0-\epsilon}^{+\infty} \frac{x(\tau)}{t-\tau} d\tau.$$ 

With the Hilbert transform, the analytical signal is defined as,

$$z(t) \equiv x(t) + iy(t) = a(t)e^{i\theta(t)} \quad \text{with} \quad y(t) \equiv H[x(t)]$$

where $a(t)$ is the instantaneous amplitude, and $\theta(t)$ is the phase function:

$$a(t) = \sqrt{x^2 + y^2} \quad \text{and} \quad \theta(t) = \arctan\left(\frac{y}{x}\right).$$

The instantaneous frequency can be defined as,

$$\omega = \frac{d\theta}{dt}.$$
Problem: A sensible instantaneous frequency cannot be found through this method for obtaining an arbitrary function.

⇒ Hilbert-Huang transform (HHT)
The general idea is to decompose the complicated signal into a finite series of function with simple oscillatory mode:

**The intrinsic mode function (IMF).**

The IMFs are (by construction),
- \# extrema = \# zero crossings
- upper and lower envelopes are symmetric with respect to zero
- a well-behaved Hilbert transform
Empirical Mode Decomposition (EMD)

EMD algorithm:

1. Initialization: \( R_{j-1}(t) = S(t) \)
2. Extract the IMF\(_j\):
   1. Identify local maxima and local minima
   2. Identify upper and a lower envelope by interpolation (cubic splines)
   3. Subtract the mean envelope from the signal
   4. Iterate until the stopping criteria is satisfied, i.e., “mean envelope = 0” (sifting)
3. Subtract the obtained mode from the signal: \( R_j(t) = S(t) - IMF_j(t) \)
4. Iterate on the residual: \( S(t) \rightarrow R_j(t) \), until the residual is a monotonic function.

\[
S(t) = IMF_1(t) + R_1(t) \\
= IMF_1(t) + IMF_2(t) + R_2(t) \\
= \ldots \\
= \sum_{j=1}^{N} IMF_j(t) + R_N(t)
\] (1)
Empirical Mode Decomposition (EMD)

Extraction of the IMF\(_1\) from the signal \(S(t)\)

Initialization: \(R_0(t) = S(t)\)

Iteration 0: upper envelope, \(e_{u0}(t)\)

Iteration 0: lower envelope, \(e_{l0}(t)\)

(Talk: Flandrin, 2010)
Empirical Mode Decomposition (EMD)

Iteration 0: mean envelope, $e_{m0}(t) = 0.5(e_{u0}(t) + e_{l0}(t))$

Iteration 0: $r_0(t) = R_0(t) - e_{m0}(t)$.

Iteration 1: upper/lower/mean-envelope on $r_0(t) \rightarrow r_1(t) = r_0(t) - e_{m1}(t)$.

(Talk: Flandrin, 2010)
Empirical Mode Decomposition (EMD)

**Iteration 5:** \( r_5(t) = r_4(t) - e_{m4}(t) \)

Here, the stopping criteria is mean envelope \( e_{m4}(t) \to 0 \)

**Extraction of the IMF\(_1\)**

Initialization: \( R_1(t) = R_0(t) - IMF_1(t) \)

Iterate on the residual: \( R_1(t) \ldots R_N(t) \), until the residual is a monotonic function
Empirical Mode Decomposition (EMD)

(Talk: Flandrin, 2010)
Hilbert spectral analysis (HSA)

For each IMF$_j$,

$$x(t) = \text{IMF}_j$$

$$y(t) = H[x(t)]$$  The Hilbert transform.

With the Hilbert transform, the analytical signal is defined as,

$$z(t) \equiv x(t) + iy(t) = a_j(t)e^{i\theta_j(t)}$$

The instantaneous frequency is,

$$\omega_j = \frac{d\theta_j}{dt}.$$  

After performing the Hilbert transform, the signal can be expressed in the following form,

$$S(t) = \Re[\sum_j a_j(t)e^{i\int \omega_j(t)dt}].$$

The Hilbert Spectrum:

$$\mathcal{H}(\omega, t) = \sum_j \int p(\omega_j, a_j) a_j^2 da_j$$

where $p(\omega_j, a_j)$ is the PDF of the instantaneous frequency $\omega$ and $a$ for all the IMFs.
Hilbert spectral analysis (HSA)

(Talk: Flandrin, 2010)
Figure 1. The daily LOD data, extracted from the comb2000-daily.eop series constructed by Gross (2001).
Other Example, Huang et al. (2003)

(Time, every 5 years)
### Some conclusions – Summary

<table>
<thead>
<tr>
<th></th>
<th>Fourier</th>
<th>Wavelet</th>
<th>Hilbert</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basis</strong></td>
<td><em>a priori</em></td>
<td><em>a priori</em></td>
<td>adaptive</td>
</tr>
<tr>
<td><strong>Frequency</strong></td>
<td>convolution:</td>
<td>convolution:</td>
<td>differentiation:</td>
</tr>
<tr>
<td></td>
<td>global uncertainty</td>
<td>regional uncertainty</td>
<td>local, certainty</td>
</tr>
<tr>
<td><strong>Presentation</strong></td>
<td>energy-frequency</td>
<td>energy-time-frequency</td>
<td>energy-time-frequency</td>
</tr>
<tr>
<td><strong>Nonlinear</strong></td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td><strong>Nonstationary</strong></td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td><strong>Feature Extraction</strong></td>
<td>no</td>
<td>discrete: no; continuous: yes</td>
<td>yes</td>
</tr>
<tr>
<td><strong>Theoretical base</strong></td>
<td>theory complete</td>
<td>theory complete</td>
<td>empirical</td>
</tr>
</tbody>
</table>

Huang (2000); Huang & Shen (2005)
Exemples:

1. De-trending and De-noising.
   (e.g., Boudraa & Cexus, 2005; Flandrin et al., 2004)

2. Correlation between different time series.
   (e.g., Barnhart & Eichinger, 2011)

3. HHT for Turbulence data.
   (Huang et al., 2008, 2009)

1.1. De-trending and De-noising, Flandrin et al. (2004)

Figure 3: Denoising. *An example of an amplitude modulated low frequency oscillation embedded in fractional Gaussian noise of Hurst exponent $H = 0.3$ is plotted in (b). The estimated energies of the 7 IMFs are plotted in (a) as the thick line, together with the “noise only” model corresponding to $H = 0.3$ (thin line) and the 99% confidence interval (dotted line). The partial reconstruction obtained by adding the EMD residual and IMFs 5 to 7 (the only ones whose energy exceeds the threshold in (a)) is plotted in (c) as the full line, and this denoised signal is superimposed to the actual signal component (dotted line). The partial reconstruction of IMFs 1 to 4 (noise estimate) is plotted in (d).*

Figure 4: Detrending of the Heart-Rate Variability signal of Figure 1. *Top diagram: Signal spectrum in log-log coordinates (thin line), with a linear fit in the mid-frequency range (thick line). Bottom diagram: model-based detrending. Left: estimated energy of the 9 IMFs, plotted as the thick line, together with the “noise only” model (thin line) and the 95% confidence interval (dotted line). Top right: original signal. Middle right: estimated trend obtained from the partial reconstruction with IMFs 5 to 9 (the only ones whose energy exceeds the threshold in the left diagram) and the residual. Bottom right: detrended signal obtained from the partial reconstruction with IMFs 1 to 4.*
Fig. 2. Noisy signals (SNR=2dB; SNR=-9dB for "ECG").

\[
\begin{align*}
\tau_j &= \bar{\sigma}_j \sqrt{2 \log_e(T)} \\
\bar{\sigma}_j &= \text{MAD}_j / 0.6745 \\
\text{MAD}_j &= \text{Median} \{ |IMF_j(t) - \text{Median} \{ IMF_j(t') \} | \}
\end{align*}
\]

\[
\tilde{h}^{(i)}(t) = \begin{cases} 
  h^{(i)}(t), & |h^{(i)}(t)| > T_i \\
  0, & |h^{(i)}(t)| \leq T_i 
\end{cases}
\]

Fig. 3. Denoising results in SNR (dB) of test signals corrupted by Gaussian noise.
2. Correlation between different time series, Barnhart (2011)

Table 3
Correlation coefficients ($r$)—Total Solar Irradiance and Sunspot from 1749 to 2009.

<table>
<thead>
<tr>
<th></th>
<th>Sunspot</th>
<th>IMF 1</th>
<th>IMF 2</th>
<th>IMF 3</th>
<th>IMF 4</th>
<th>IMF 5</th>
<th>IMF 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSI</td>
<td>0.85</td>
<td>0.50</td>
<td>0.82</td>
<td>0.27</td>
<td>0.28</td>
<td>0.33</td>
<td>0.20</td>
</tr>
<tr>
<td>IMF 1</td>
<td>0.21</td>
<td>0.54</td>
<td>0.26</td>
<td>0.04</td>
<td>-0.08</td>
<td>0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>IMF 2</td>
<td>0.61</td>
<td>0.54</td>
<td>0.95</td>
<td>0.10</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>IMF 3</td>
<td>0.36</td>
<td>0.10</td>
<td>0.31</td>
<td>0.74</td>
<td>0.10</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>IMF 4</td>
<td>0.26</td>
<td>-0.02</td>
<td>0.01</td>
<td>0.19</td>
<td>0.79</td>
<td>-0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>IMF 5</td>
<td>0.48</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.01</td>
<td>0.31</td>
<td>0.84</td>
<td>0.25</td>
</tr>
<tr>
<td>IMF 6</td>
<td>0.59</td>
<td>-0.03</td>
<td>-0.05</td>
<td>-0.01</td>
<td>-0.001</td>
<td>0.59</td>
<td>0.96</td>
</tr>
</tbody>
</table>
2. Correlation between different time series, Barnhart (2011)

Table 4: Correlation coefficients (r) – Total Solar Irradiance and Global Mean Temperature from 1880 to 1945.

<table>
<thead>
<tr>
<th>TSI</th>
<th>IMF 1</th>
<th>IMF 2</th>
<th>IMF 3</th>
<th>IMF 4</th>
<th>IMF 5</th>
<th>IMF 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.28</td>
<td>0.02</td>
<td>-0.03</td>
<td>-0.34</td>
<td>0.13</td>
<td>0.28</td>
</tr>
<tr>
<td>IMF 1</td>
<td>-0.13</td>
<td>-0.03</td>
<td>-0.08</td>
<td>-0.09</td>
<td>-0.04</td>
<td>-0.04</td>
</tr>
<tr>
<td>IMF 2</td>
<td>-0.18</td>
<td>0.004</td>
<td>-0.02</td>
<td>-0.34</td>
<td>-0.07</td>
<td>-0.04</td>
</tr>
<tr>
<td>IMF 3</td>
<td>0.09</td>
<td>0.02</td>
<td>-0.02</td>
<td>-0.40</td>
<td>0.30</td>
<td>0.10</td>
</tr>
<tr>
<td>IMF 4</td>
<td>0.11</td>
<td>0.07</td>
<td>0.08</td>
<td>0.04</td>
<td>-0.06</td>
<td>-0.26</td>
</tr>
<tr>
<td>IMF 5</td>
<td>0.77</td>
<td>0.03</td>
<td>0.02</td>
<td>0.20</td>
<td>0.66</td>
<td>0.86</td>
</tr>
<tr>
<td>IMF 6</td>
<td>0.67</td>
<td>0.01</td>
<td>-0.04</td>
<td>-0.04</td>
<td>0.27</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Table 5: Correlation Coefficients (r) – Total Solar Irradiance and Global Mean Temperature from 1945 to 2009.

<table>
<thead>
<tr>
<th>TSI</th>
<th>IMF 1</th>
<th>IMF 2</th>
<th>IMF 3</th>
<th>IMF 4</th>
<th>IMF 5</th>
<th>IMF 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.13</td>
<td>0.01</td>
<td>0.09</td>
<td>0.39</td>
<td>0.27</td>
<td>-0.06</td>
</tr>
<tr>
<td>IMF 1</td>
<td>0.02</td>
<td>0.004</td>
<td>0.13</td>
<td>0.17</td>
<td>0.01</td>
<td>0.005</td>
</tr>
<tr>
<td>IMF 2</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.06</td>
<td>0.43</td>
<td>0.08</td>
<td>-0.10</td>
</tr>
<tr>
<td>IMF 3</td>
<td>0.10</td>
<td>-0.003</td>
<td>0.03</td>
<td>0.08</td>
<td>0.67</td>
<td>-0.05</td>
</tr>
<tr>
<td>IMF 4</td>
<td>0.27</td>
<td>0.02</td>
<td>0.01</td>
<td>-0.05</td>
<td>0.13</td>
<td>0.59</td>
</tr>
<tr>
<td>IMF 5</td>
<td>-0.85</td>
<td>0.004</td>
<td>0.06</td>
<td>-0.006</td>
<td>-0.28</td>
<td>0.81</td>
</tr>
<tr>
<td>IMF 6</td>
<td>0.83</td>
<td>-0.04</td>
<td>-0.04</td>
<td>0.007</td>
<td>0.33</td>
<td>0.26</td>
</tr>
</tbody>
</table>
3. HHT for Turbulence data, Huang et al. (2008, 2009)

Hilbert marginal spectrum:

\[ h(\omega) = \int_0^{+\infty} H(\omega, t) dt. \]

Fig. 1: Comparison of the Hilbert marginal energy spectrum (solid line) and Fourier spectrum (dashed line, vertically shifted). The slope of the reference line is \(-5/3\). Both the second-order Hilbert and Fourier spectra indicate the same inertial subrange, \(10 < f \text{ (or } \omega) < 1000 \text{ Hz}\). The insert shows the compensated spectra. The HHT spectra estimated using two different algorithms are shown for comparison (dot-dashed line for the Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) method and solid line for the more classical spline method), indicating a stability of the spectrum with respect to the algorithm used.

Fig. 2: (a) Mean frequency \(vs.\) mode number for the turbulent velocity time series. There is an exponential decrease with a slope very close to 1. This indicates that EMD acts as a filter bank which is almost dyadic. (b) Fourier spectrum of each mode (from 1 to 12) showing that they are narrow-banded. The slope of the reference line is \(-5/3\) corresponding to the inertial-range Kolmogorov spectrum.
3. HHT for Turbulence data, Huang et al. (2008, 2009)

Hilbert marginal spectrum:

\[ h(\omega) = \int_0^{+\infty} H(\omega, t) dt = \int_0^{+\infty} p(\omega, a) a^2 da. \]

Generalization to arbitrary order:

\[ L_q(\omega) = \int_0^{+\infty} p(\omega, a) a^q da. \]

In case of scale invariance,

\[ L_q(\omega) \sim \omega^{-\zeta(q)} \]

Compare to the structure function analysis \( \zeta(q) - 1 = \zeta(q) \):

\[ \langle \Delta x^q \rangle \sim I^\zeta(q) \]

Fig. 8: Comparison of the scaling exponents \( \zeta(q) - 1 \) (diamond) with the classical \( \zeta(q) \) obtained from structure functions analysis with the ESS method (dashed line) and K41 \( q/3 \) (solid line). The insert shows the departure from the K41 law.
3. HHT for Turbulence data, Huang et al. (2008, 2009)

**Fig. 1.** The river flow discharge time series of (a) Seine river, recorded from 1 January 1976 to 28 April 2008, (b) Wimereux river, recorded from 1 January 1981 to 27 May 2006. The data illustrate clear strong annual cycles with huge fluctuations. The total lengths are 11,828 and 9278 data points for the Seine river and the Wimereux river, respectively.

**Fig. 6.** Representation of the cross-correlation $\rho_{ij}$ between IMF modes from the Seine and Wimereux rivers. The data span is taken from 1 January 1981 to 27 May 2006 for both series. For convenience, we consider the coefficient value $\log_{10}(\rho_{ij})$. As expected, the annual cycle shows a strong correlation with a coefficient $\rho_{9,8} = 0.426$. The coefficient of the most correlated modes is $\rho_{11,11} = 0.579$. These two strong correlations are then marked by $\square$. 

References

Spectral analysis

Hilbert-Huang transform

IMF

EMD

HSA

example

Summary

Examples

ex 1

ex 2

ex 3

Conclusion
3. HHT for Turbulence data, Huang et al. (2008, 2009)

![Graph showing comparison of scaling exponents](image)

**Figure 5.6:** (a) A portion of fBm data with (bottom) and without (top) a sine wave perturbation (middle), and (b) the corresponding Fourier power spectrum.

![Graph showing influence of single scale](image)

**Figure 5.7:** Influence of a single scale on (a) the second order structure function, and (b) the second order Hilbert marginal spectrum with various intensities $I$. The vertical solid line indicates location of disturbance.
“The Hilbert-Huang transform (HHT) is an empirically based data-analysis method. Its basis of expansion is adaptive, so that it can produce physically meaningful representations of data from nonlinear and non-stationary processes.”

Huang & Shen (2005).

The Hilbert-Huang transform (HHT) is,

1. A new analysis technique proposed by Huang et al. (1998, 1999)
3. An adaptive decomposition operating in the time domain.
4. Design for nonlinear and non-stationary.
5. An energy/frequency/time representation (such as wavelet).
6. Based on
   Empirical method decomposition (EMD) \(\Rightarrow\) intrinsic mode function (IMF),
   Hilbert spectral analysis (HSA) \(\Rightarrow\) instantaneous frequency.