*Proceedings of the 9th International Conference "Reliability and Statistics in Transportation and Communication" (RelStat'09), 21–24 October 2009, Riga, Latvia, p. 329-334. ISBN 978-9984-818-21-4*
*Transport and Telecommunication Institute, Lomonosova 1, LV-1019, Riga, Latvia*

# PSEUDOCEPSTRAL ANALYSIS OF ELECTROMAGNETIC TRANSIENTS USING EMPIRICAL MODE DECOMPOSITION

***Yury A. Krasnitsky***

*Transport and Telecommunication Institute*
*Lomonosova Str., 1, Riga,LV-1019, Latvia*
*Ph.: +371 67100608. Fax: +371 67100660. E-mail: krasn@tsi.lv*

An approach to process transients in order to estimate their inherent structure has been considered. The proposal is based on empirical mode decomposition (EMD). It permits to expand any signals (including non-stationary and non-linear ones) into a little amount of quasi-harmonic components named intrinsic mode functions (IMFs). Analysis of IMFs leads to an analogy with usual power spectrum calculated by Fourier transformation (FT). In this work it is proposed to apply EMD for expanding logarithm of power spectrum into new set of IMFs. Analysis of instantaneous frequency of the latter creates a certain "quasicepstrum", which can serve as an estimation of inherent structure of the learned signal. The set of Matlab programs is worked out to illustrate the proposed method and to compare it with FT. The results are presented for natural electromagnetic (EM) pulses processing.

***Keywords:*** *non-stationary signal, logarithmic spectrum, cepstrum, empirical mode decomposition*

## 1. Introduction

An increasing interest in non-stationary transient signals analysis has observed in areas such radar, communication systems, natural EM phenomena, medicine diagnostics, and others. A linear convolutional model is frequently introduced as a description of that signals [1]. It has supposed the signal is formed through linear combination of some pulses, which arrive to the observation point along non-coincided paths. Mutual delays of pulses depend on differences of these paths lengths, features of propagation trace for every pulse as well as properties of radiation source. Just the delays determine the inherent structure of the transient. Therefore, evaluation of delays from the observed time series is the primary step to decide goal inverse problem, namely reconstruction separate traces and radiator features using appropriate signal processing method.

Frequently, only unique non-repeated realization of the transient is available. Let the signal under consideration is $u(t)$. Collection of methods to estimate delays of internal structural components in it is rather poor. A common way used is to apply the cepstrum analysis [2]. It is conceptually based on three steps.

Omitted details concerning discretization of the signal, one can calculate as follows:

(a) direct Fourier transformation (FT) to find the signal spectrum

$$\dot{S}(\omega) = \int_{-\infty}^{+\infty} u(t)\exp(-i\omega t)dt = \mathrm{FT}[u(t)]; \tag{1}$$

(b) the logarithmic spectrum

$$\dot{S}_{\log}(\omega) = \log \dot{S}(\omega); \tag{2}$$

(c) inverse FT of the latter to compute the cepstrum

$$c(\tau) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \dot{S}_{\log}(\omega)\exp(i\omega t)d\omega = \mathrm{FT}^{-1}[\dot{S}_{\log}(\omega)]\cdot \tag{3}$$

The result (3) is known as the complex cepstrum of the signal $u(t)$. The values assigned as $\tau$ were named **quefrencies**. Physically, anyone from quefrencies has meaning as some delay. In many cases it is sufficiently to operate with power cepstrum or real cepstrum only [2].

Mathematically, the FT possesses known lacks itself. Particularly, FT is not adequate for non-stationary signal processing [1]. In addition, cepstral algorithms are worked satisfactorily for the linear models only, since convolutional components of signal are mapped onto logspectrum (2) additively. Moreover, they run successfully if the next conditions are kept: (i) signal has sufficient duration; (ii) inherent structure of it is periodical; (iii) signal-to-noise ratio should be sufficiently large. As a rule, natural EM transient phenomena have not complied with conditions (i) and (ii). Pulses (atmospherics) radiated by lightning discharges are just that very case.

In the sections below, an attempt is made to develop an alternative way from (1) – (3) for evaluating of delays in transients. The method proposed is based on the EMD.

## 2. Principles of EMD

Empirical mode decomposition proposed by Huang et al. [3] deals with both linear and nonlinear or non-stationary signals. Unlike the FT, the basis functions for EMD called intrinsic mode functions (IMFs) derive directly and adaptively from the signal itself. The IMFs obtained from the decomposition of the signal by EMD must obey two general assumptions:

a) have the same number of extrema and zero crossings or differ at most by one;

b) be symmetric with respect to the local zero mean.

Thus, the *sifting* procedure to obtain IMFs of the signal $u(t)$ is described as follows.

1) Identify all the maxima and the minima in the signal $u(t)$.

2) Generate its upper and lower envelopes using cubic spline interpolation.

3) Compute the point by point local mean $m_1$ from upper and lower envelopes.

4) Extract the details

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

5) Check the properties of $h_1$ and iterate $k$ times, then

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

becomes the IMF once it satisfies some stopping criterion. It is designated as first IMF

$$c_1 = h_{1k} \ .$$

6) Repeat steps 1) to 5) on the extracted data

$$r_1 = u(t) - c_1 \ .$$

7) The step 6) is repeated until all the IMFs and residuals are obtained.

As the stopping criterion, or the sifting threshold, the normalized squared difference between two successive sifting operations is to be set: That is

$$\sum_{k=0}^{K} \left| h_{n-1}[k] - h_n[k] \right|^2 \Big/ h_{n-1}^2[k] \le thresh \ . \tag{4}$$

*Thresh* value is generally set between 0.2 and 0.3. The decomposed signal can be represented as

$$u(t) = \sum_{n=1}^{N} c_n(t) + r_N \tag{5}$$

where N is the total number of IMFs and $r_N$ is the final residue.

EMD can provide a certain time space filtering. For instance, elimination from (5) some IMFs and residue $r_N$ which can be either the mean trend or a constant leads to new signal

$$\breve{u}(t) = \sum_{n=1}^{M} c_n(t) , \tag{6}$$

where $M < N$, and $c_n$ are preserved IMFs.

Let's try to apply the EMD algorithm described in 1) – 7) for decomposition of natural EM transients. Fig. 1 (upper row) shows the original atmospherics *u(t)* observed at the distance about 600 km from a lightning return stroke. EMD creates the set of IMFs from $c_1$ to $c_5$ and residue $r_N$.

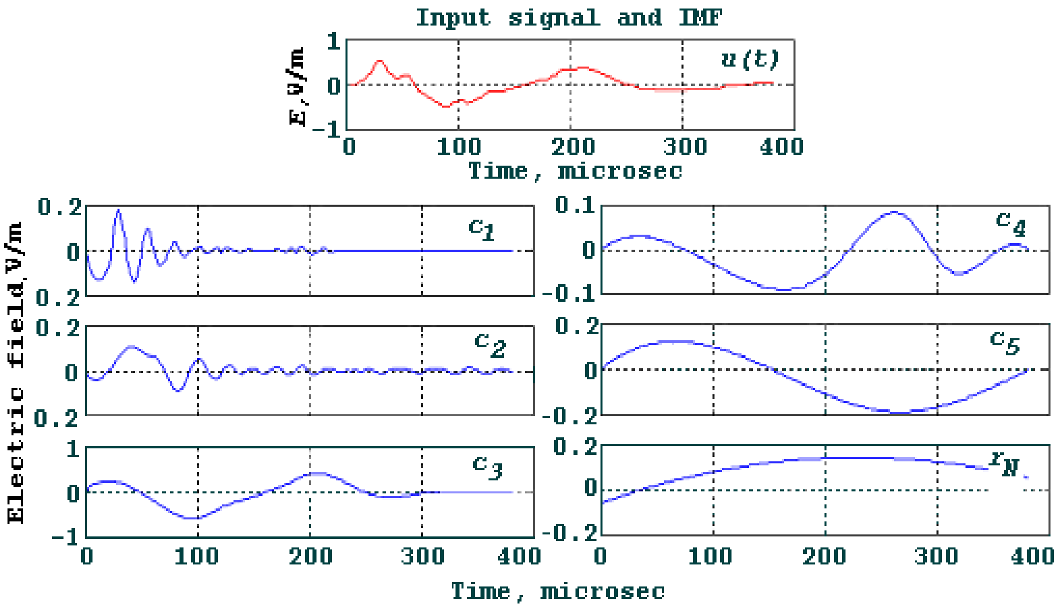


*Fig. 1.* Empirical mode decomposition of the atmospherics

One can observe oscillatory features of IMFs in point of different frequency contents in those functions. It is the ground to filtering the signal in accordance of eq. (6). Elimination $c_1$ and $c_2$ from (5) is equivalent of low-pass filtering for *u(t)* and it can work as noise reduction. On the contrary, excluding $c_3$, $c_4$ and $c_5$ preserves high frequency components in (6). The residual $r_N$ is the slowest function in expansion (6). It has to consider as a non-linear trend.

## 3. Hilbert's Spectrum

Empirical mode decomposition proposed by Huang et al. [3] deals with both linear and nonlinear, and/or nonstationary signals. The oscillatory nature of every IMF permits to map the latter onto complex space using the Hilbert transform (HT). The result is analytic signal

$$w_n(t) = c_n(t) + i\,\mathrm{HT}[c_n(t)] = A_n(t)\exp[i\varphi_n(t)] \tag{7}$$

using this IMF as the real part and its HT as the imaginary part,
where

$$A_n(t) = \sqrt{\mathrm{Re}^2\, w_n(t) + \mathrm{Im}^2\, w_n(t)} \tag{8}$$

is the amplitude (envelope) and

$$\varphi_n(t) = \mathrm{arc\,tg}\,[\mathrm{Im}\, w_n(t)/\mathrm{Re}\, w_n(t)] \tag{9}$$

is the phase of $n^{\mathrm{th}}$ analytic IMF.

Then, omitted the residue and perform the HT on each IMF, we get the complex signal

$$U(t) = \sum_{n=1}^{N} A_n(t)\exp[i\varphi_n(t)]\,. \tag{10}$$

It can be examined instead of real signal *u(t)*.

The instantaneous frequency

$$\omega_n(t) = \mathrm{d}\varphi_n(t)/\mathrm{d}t \tag{11}$$

for every IMF follows from (9). Subsequently, the signal (10) can be expressed as

$$U(t) = \sum_{n=1}^{N} A_n(t) \exp[i \int_0^T \omega_n(t) dt]. \tag{12}$$

Temporal functions (11) plotted on the common field leads to the 2-*d* picture in coordinates "time-frequency". This picture can be extended by supplement of dependence of the amplitude (8) versus time *t* and frequencies (11). Resulting 3-*d* distribution "time-frequency-amplitude" is known as the Hilbert amplitude spectrum $S_H(\omega, t)$ of the signal *u(t)*. Combining algorithms for EMD and Hilbert spectrum calculation the new powerful signal processing method was created called the Hilbert-Huang transform (HHT) [3].

One also can define the marginal average power spectrum as

$$S_{ma}(\omega) = \frac{1}{T} \int_0^T S_H(\omega, t) dt. \tag{13}$$

This spectrum can serve as a measure of average power contribution from each frequency value. It is noted that the frequency in the Hilbert spectrum has a totally different meaning from it in Fourier spectrum [4]. To comparison, we were calculated both the logarithmic FT power spectrum with (2) and, by analogy, the logarithmic marginal spectrum from eq. (13). Fig. 2 shows these two spectra of the same atmospherics *u(t)* which has been already presented on Fig. 1. Everyone can see that these spectra are not radically far in average, excepting fine details observed in the marginal spectrum. On the other hand, the slow components of these spectra are nearly similar.

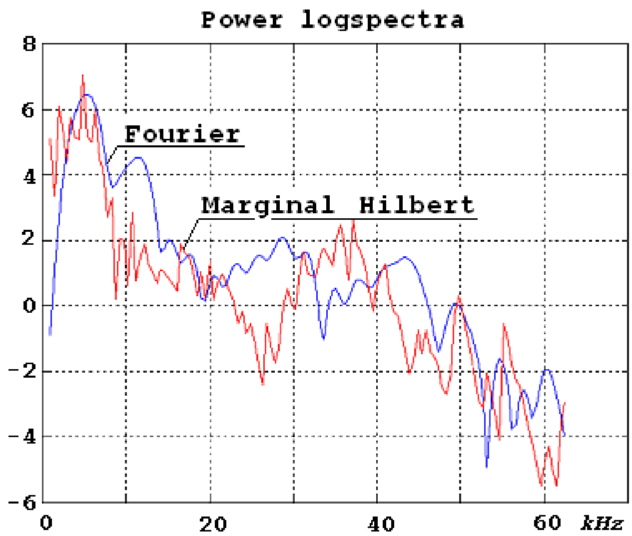


*Fig. 2.* Comparison of the spectra

## 4. Evaluation of Delays

For convolutional model of signal, the cepstral processing method (1)-(3) after taking the logarithm of spectrum should be led to *implicit* additive relations among signal components, since this sum is not disunited to the separate items. Unlike it, for any signals, including non-stationary and non-linear ones, EMD of the marginal logarithmic spectrum (13) allows the *explicit* additive representation as the "palpable" sum of the separated IMFs.

Fig. 3 shows the IMFs calculated from the marginal logspectrum (the upper row). If to assume that the signal *u(t)* is composed as a group of delayed pulses who are various in common case, then every IMF in Fig. 3 should be conformed to the definite pulse from this group. The delays, or the times of arrivals, of these pulses determine the structure of the signal.

An evident chance for filtering of this logspectrum is seen now. One can suppose that main features of the signal *u(t)* (Fig.1, upper plot) are mapped in Fig. 3 onto the IMFs signed $c_3$ and $c_4$. Indeed, these intrinsic functions distinct from the others with nearly quasi-harmonic behaviour. Probably, the $c_2$ would be inserted too to give some completeness. The IMF $c_1$ is contaminated by noise, and should be eliminated from analysis in consequence of it, as well as the residue $r_N$.

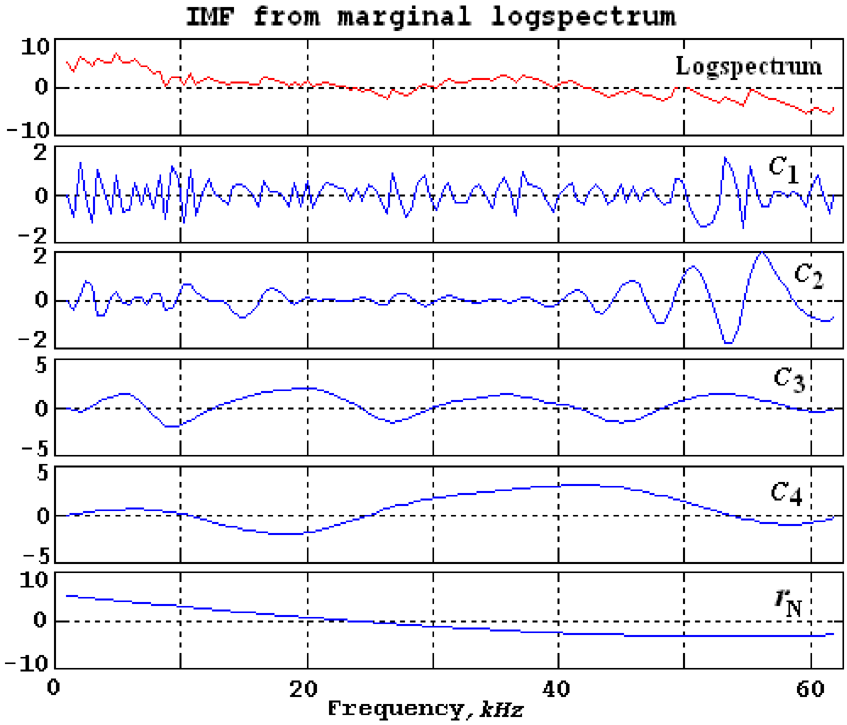


*Fig. 3.* Empirical mode decomposition of the marginal logspectrum

In some cases it is helpful to equalize the selected IMFs in order to do their forms more similar to sinusoidal functions. Quasi-harmonic behaviour permits to apply for estimation of delays in that forms the methods based on calculation of IMFs eigenvalue. Using Pisarenko's algorithm [5] one can compute the data followed in the table:

| **IMFs from Fig. 3** | $c_4$ | $c_3$ | $c_2$ |
|---|---|---|---|
| **Delays,** ***microsec*** | **30.06** | **58** | **197.9** |
| **Amplitudes,** ***relative units*** | **1.00** | **0.617** | **0.314** |

The graph (Fig. 4 below) can be constructed using the table. It would be recognized as some pseudocepstrum because the data about delays were arrived from frequency domain by the same way as in eq. (3). This graph can be considered as a set of $\delta$-impulses with determined times of arrivals and relative amplitudes of the pulses that dictate the structure of the transient.

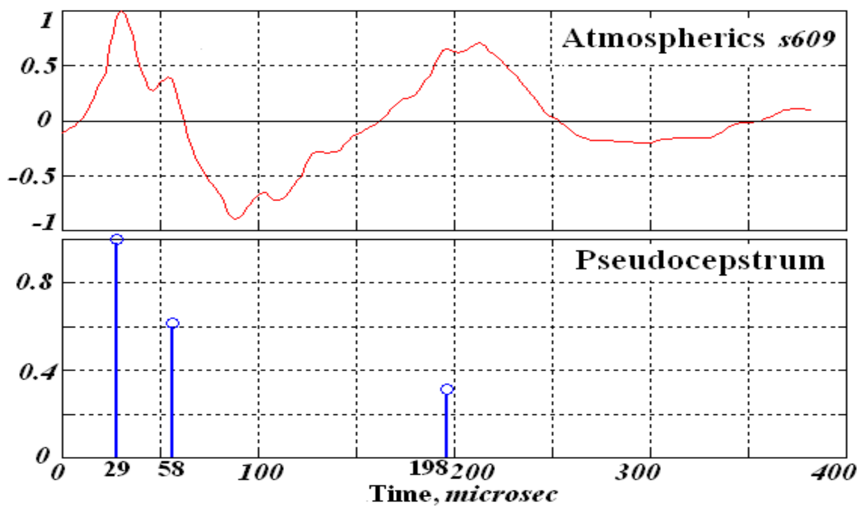


*Fig. 4.* The atmospherics and its pseudocepstrum

Conformably to the atmospherics, the hop propagation model is frequently used [6]. The observed signal (Fig. 4, upper plot) is considered as result of interaction of ground wave with sky waves reflected from ionosphere. The $\delta$-pulses positions in Fig. 4 indicate consequently these moments: (i) EM radiation from lightning channel is ended; (ii) one-hop ionospheric wave is arrived to the observation point; (iii) two-hop ionospheric wave is arrived. Proceed from these data, the inverse problem concerning geometry of waveguide "Earth – ionosphere" would be decided. It is important contribution in theory of single-station systems for passive location of lightning strokes on the basis of the analysis of own EM radiation.

## 5. Conclusions

In this paper, the new method for the analysis of inherent structure of transients had presented based on EMD. A concept of pseudocepstrum and the appropriate procedure for computing has been proposed. As an example, processing of an atmospherics is discussed. All the procedures described-above have been realized in Matlab.

## References

1. Hayes, M.H. *Statistical digital signal processing and modeling*. N.Y.: J.Wiley, 1996.
2. Oppenheim, A.V. *Applications of digital signal processing*. N.Y.:Prentice-Hall, Inc., 1978.
3. Huang, N.E. et al. The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis, *Proc.Roy.Soc.London*, A,**454**, 1998, pp. 903-995.
4. Peng, H.M., Chang, P.C., Chang, F.R. Hilbert spectrum for time-domain measurement data and its application. In: *Proceedings of the 34 th Annual Precise Time and Time Interval (PTTI) Meeting,* pp. 501-510.
5. Marple, S.L., Jr. *Digital spectral analysis with applications.* N.Y.:Prentice-Hall, Inc., 1987.
7. Krasnitsky Yu.A. (1994) Evaluation of lightning current pulse parameters from spherics waveforms. - *Journ. of Geophys. Res.*, **99**, No. D5, 10723-10725.