Transport and Telecommunication, 2009, Volume 10, No 3, 4–9
Transport and Telecommunication Institute, Lomonosova 1, Riga, LV-1019, Latvia

PSEUDOCEPSTRAL ANALYSIS OF TRANSIENT PULSES BASED ON EMPIRICAL MODE DECOMPOSITION

Yury A. Krasnitsky

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

It is proposed the method to estimate an inherent structure of transient pulses 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-harmonics named intrinsic mode functions (IMFs). The IMFs concept leads to an analogy with spectrum calculated by Fourier transformation (FT). In this work it is proposed to apply EMD for expanding logarithm of that spectrum into new set of IMFs. Instantaneous frequencies analysis of the latter creates a certain "quasicepstrum" which can serve as an estimation of inherent structure of the learned signal. The Matlab codes are worked out to illustrate the proposed method and to compare it with FT. The results are presented for processing natural electromagnetic (EM) pulses.

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

1. Introduction

A noticeable interest in transient signals analysis has observed in such areas as radar, communication, natural EM phenomena, or medicine diagnostics. Frequently, only unique non-repeated realization of the transient is available. The convolutional model [1] has supposed the transient is a combination of some pulses arrived to the observation point along non-coincided paths. Mutual delays of pulses depend on differences of these paths lengths; propagation traces features as well as properties of radiation source. Just the delays determine the inherent structure of the transient. Therefore, evaluation of delays from observed signal is the way to decide goal inverse problem, namely reconstruction separate traces and radiator features using signal processing method.

Let's consider that the signal is $u(t)$. The methods to estimate delays of internal structural components in it are rather poor. A common way used in this investigation – is to apply the cepstrum analysis [2]. It is conceptually based on three steps:

(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

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

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

Mathematically, the FT possesses the 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 following 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 transients have not complied with conditions (i) and (ii).

Pulses (atmospherics) radiated by lightning discharges are just that very case. The really received (non-modelled) atmospherics *u(t)* from a lightning return stroke at 600 km distance as well as real part of logspectrum (2) and power cepstrum of it are shown on Figure 1. It is seen that estimation of delays from cepstrum in this case is rather impossible.

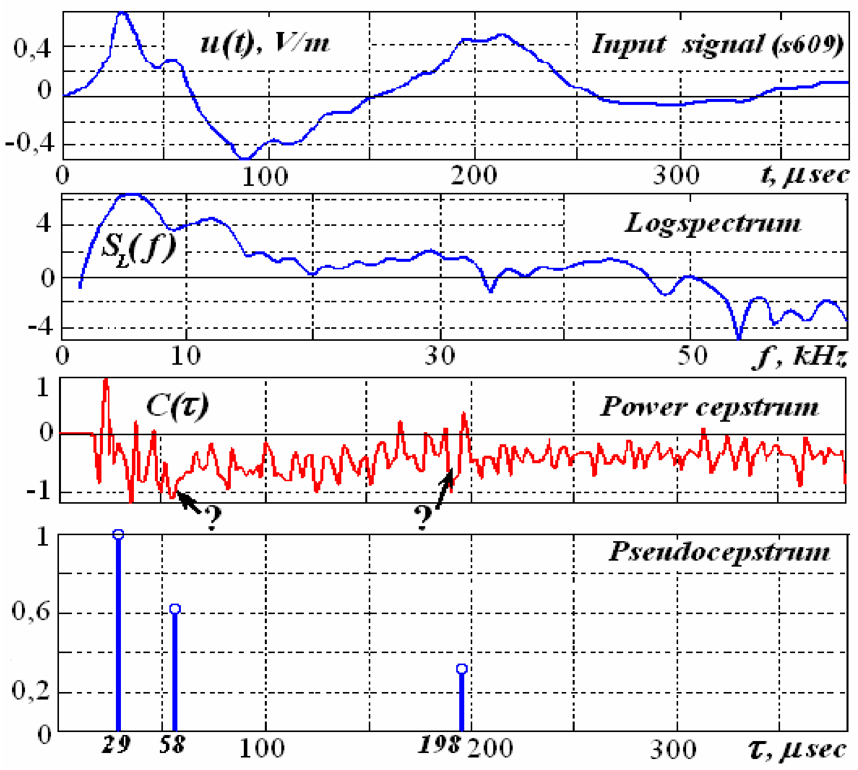

Figure 1. Analysis of the atmospherics

In the sections followed, 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 created 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, which 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

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

5) Check the properties of h1 and iterate k times, then

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

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

$c_1(t) = h_{1k}(t)$.

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

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

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

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

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

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

where N is the total number of IMFs and $r_N(t)$ 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 1)-7) for decomposition of natural EM transients. Figure 2 shows the atmospherics *u(t)* (the upper row) and the set of IMFs (2[th] and the lower rows).

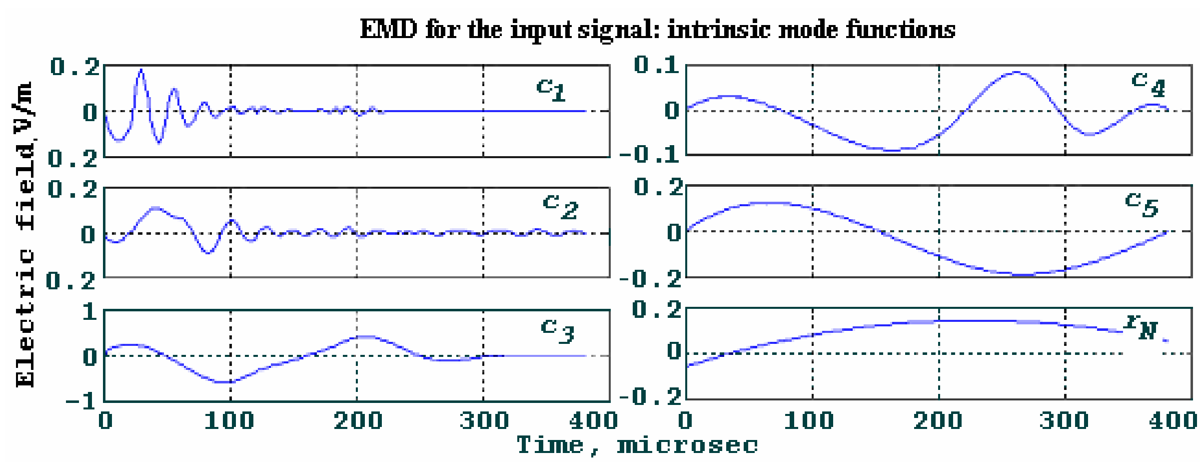

Figure 2. Empirical mode decomposition of the atmospherics

One can observe oscillatory features of IMFs in point of different frequency contents of 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 considered as a non-linear trend.

3. Hilbert Spectrum

Empirical mode decomposition 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

$$w_{\mathrm{n}}(t) = c_{\mathrm{n}}(t) + i\,\mathrm{HT}[\,c_{\mathrm{n}}(t)] = A_{\mathrm{n}}(t)\exp[\,i\varphi_{\mathrm{n}}(t)] \tag{7}$$

having this IMF as the real part and its HT as the imaginary part

where

$$A_{\mathrm{n}}(t) = \sqrt{\mathrm{Re}^{2}\, w_{\mathrm{n}}(t) + \mathrm{Im}^{2}\, w_{\mathrm{n}}(t)} \tag{8}$$

is amplitude (envelope) and

$$\varphi_{\mathrm{n}}(t) = \mathrm{arc\ tg}\,[\mathrm{Im}\; w_{\mathrm{n}}(t)/\mathrm{Re}\; w_{\mathrm{n}}(t)] \tag{9}$$

is phase of n^{th} analytic IMF.

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

$$U(t) = \sum_{\mathrm{n}=1}^{\mathrm{N}} A_{\mathrm{n}}(t)\exp[\,i\varphi_{\mathrm{n}}(t)]\,, \tag{10}$$

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

The instantaneous frequency

$$\omega_{\mathrm{n}}(t) = \mathrm{d}\varphi_{\mathrm{n}}(t)/\mathrm{d}t \tag{11}$$

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

$$U(t) = \sum_{\mathrm{n}=1}^{\mathrm{N}} A_{\mathrm{n}}(t)\exp[\,i\int_{0}^{T}\omega_{\mathrm{n}}(t)\mathrm{d}t]. \tag{12}$$

Temporal functions (12) 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).

One also can define the marginal average power spectrum as

$$S_{\mathrm{ma}}(\omega) = \frac{1}{T}\int_{0}^{T} S_{\mathrm{H}}(\omega,t)\,\mathrm{d}t. \tag{13}$$

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 Fourier spectrum [4]. To comparison, we have calculated both the logarithmic FT power spectrum with (2) and, by analogy, the logarithmic marginal spectrum from Eq. (13). It could be seen on Figure 3 that these spectra are not radically far in average, with the exception of fine details visible in the marginal spectrum. However, the slow components of these spectra are nearly similar.

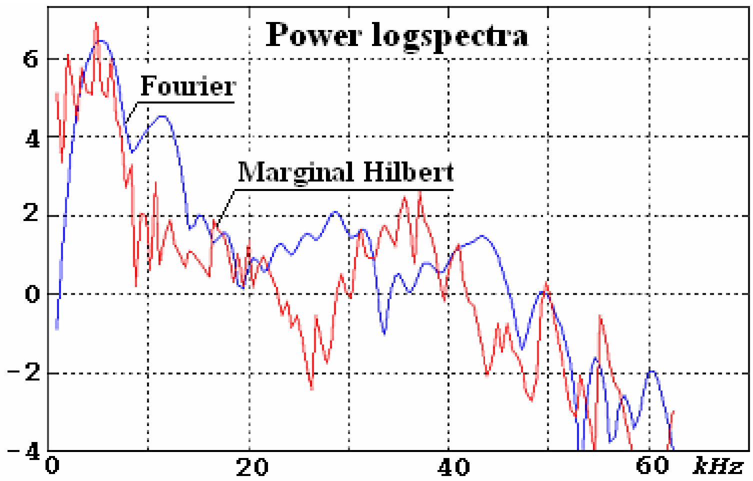

Figure 3. Comparison of the spectra

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 disjunctive to the separate items. Unlike it, for any signals, including non-stationary and non-linear ones, EMD of the marginal logspectrum (13) allows the *explicit* additive representation as the "palpable" sum of the separated IMFs. Fig. 4 shows the IMFs calculated from the marginal logspectrum.

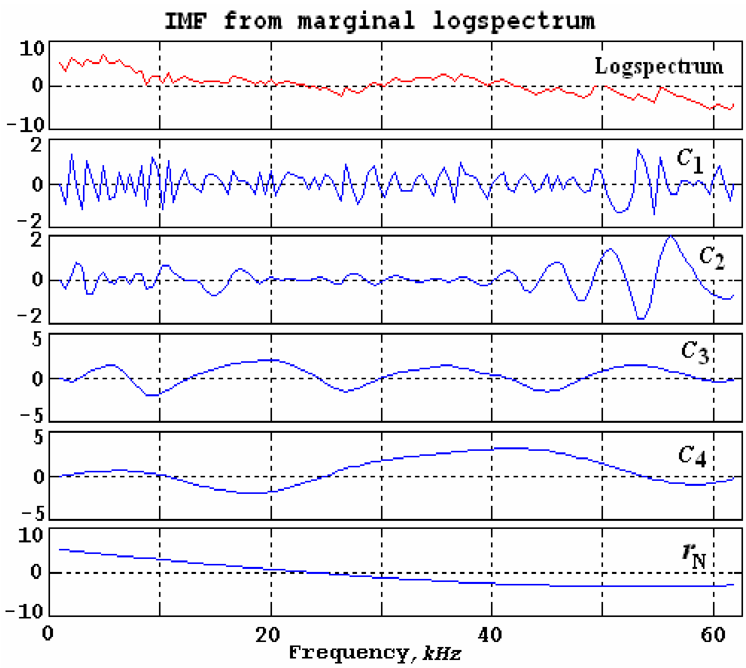

Figure 4. The IMFs given rise to the pseudocepstrum.

If the input signal *u(t)* is a group of delayed pulses which are various in common case, then every IMF should be conform to the definite pulse from this group. The delays of pulses determine the structure of the signal.

An evident chance for filtering of this logspectrum is seen. One can suppose that main features of the *u(t)* are mapped onto the IMFs signed c_3 and c_4. Indeed, these IMFs distinct from the others with quasi-harmonic behaviour. The c_2 would be inserted too to give completeness. The IMF c_1 is contaminated by noise and eliminated from analysis, as well as the residue r_N .

Quasi-harmonic behaviour of logspectral components permits to apply for estimation of delays the methods based on calculation of IMFs eigenvalues. Using algorithm by Pisarenko [5] one can compute the data followed in the table:

IMFs from Figure 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. 1, below) can be constructed using the table. It would be recognized as some *pseudocepstrum* because the data about delays are arrived from frequency domain by the same way as in Eq. (3). This graph can be considered as set of δ-impulses with determined times of arrival and relative amplitudes of the pulses that dictate the structure of the transient.

Conformably to the atmospherics the hop propagation model is frequently used [6]. The observed signal is considered as result of interaction of ground wave with sky waves reflected from ionosphere. The δ-pulses positions showed on the lowest plot on Figure 1 and indicate consequently on the following moments:

1) the electric current pulse propagating along lightning channel is reached the top of the channel; EM radiation from the channel is ended;
2) one-hop ionospheric wave is arrived to the observation point;
3) two-hop ionospheric wave is arrived.

Proceeding 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 is presented based on EMD. A concept of pseudocepstrum and the appropriate procedure for computing is 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 Modelling.* 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, Vol. 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: *34 th Annual Precise Time and Time Interval (PTTI) Meeting.* 2003, pp. 501-510.
5. Marple, S.L., Jr. *Digital Spectral Analysis with Applications.* N.Y.: Prentice-Hall, Inc., 1987.
6. Krasnitsky, Yu.A. Evaluation of Lightning Current Pulse Parameters from Spherics Waveforms, *J. Geophys. Res.*, Vol. 99, No D5, 1994, pp. 10723-10725.