Advances in Adaptive Data Analysis
Vol. 3, No. 4 (2011) 527–540

DOI: 10.1142/S1793536911000921

A REINTERPRETATION OF EMD BY CUBIC SPLINE INTERPOLATION

MINJEONG PARK
Department of Statistics, Seoul National University,
Seoul 151-747, Korea
mj.park.stat@gmail.com

DONGHOH KIM
Department of Applied Mathematics, Sejong University,
Seoul 143-747, Republic of Korea
donghohkim@sejong.edu

HEE-SEOK OH
Department of Statistics, Seoul National University,
Seoul 151-747, Korea
heeseok@stats.snu.ac.kr

Empirical mode decomposition (EMD) is a data-driven technique that decomposes a signal into several zero-mean oscillatory waveforms according to the levels of oscillation. Most of the studies on EMD have focused on its use as an empirical tool. Recently, Rilling and Flandrin [2008] studied theoretical aspects of EMD with extensive simulations, which allow a better understanding of the method. However, their theoretical results have been obtained by considering constraints on the signal such as equally spaced extrema and constant frequency. The present study investigates the theoretical properties of EMD using cubic spline interpolation under more general conditions on the signal. This study also theoretically supports modified EMD procedures in Kopsinis and Mclaughlin [2008] and developed for improving the conventional EMD. Furthermore, all analyses are preformed in the time domain where EMD actually operates; therefore, the principle of EMD can be visually and directly captured, which is useful in interpreting EMD as a detection procedure of hidden components.

Keywords: Cubic splines; empirical mode decomposition; frequency.

1. Introduction

A time-varying frequency, termed instantaneous frequency (IF), is a key characteristic of nonstationary signals [Boashash (1992)]. The empirical mode decomposition (EMD) proposed by Huang *et al.* [1998] is a data-driven algorithm for decomposing a signal into intrinsic mode functions (IMFs) in which the IF is well defined. Although EMD has been widely used for analyzing nonstationary or nonlinear signals [Huang *et al.* (1998); Huang and Shen (2005)], there has been a deficiency in theoretical studies on it.

Recently, the pioneering study of Rilling and Flandrin [2008] provided a remarkable discussion of the possibilities and limitations of EMD along with extensive simulations. They clarified the ranges of frequency and amplitude ratios between the higher frequency (HF) and the lower frequency (LF) components, which can be used to figure out which component determines the extrema of a signal. They then derived the model for EMD with those extrema based on Fourier representation according to different ranges. Moreover, Rilling and Flandrin [2008] extended their analysis on EMD to a general signal model that consists of simple waveforms defined by Fourier expansion.

However, their studies were carried out under the following stringent conditions: (1) the two components (HF and LF) of the signal are sinusoids having constant amplitudes and frequencies, (2) the extrema of the signal $x(t)$ are equally spaced in order to use the Dirac comb for extremum sampling, (3) the extrema of the signal are assumed to be the same as the extrema of HF or LF; in that case, the iteration of EMD process is not actually required because this assumption is the ultimate goal of EMD that is supposed to be achieved at the last iteration in the EMD process, and (4) in the generalization of Rilling and Flandrin [2008], the Fourier expansion is applied to an IMF, which might be a nonstationary component. Although Rilling and Flandrin [2008] provided excellent results of theoretical EMD aspects particularly for signals having two constant frequency components, their analysis was too specific for explaining the manner that EMD works for a general signal having time-varying frequency and amplitude. Furthermore, it cannot be visible to show the process, in which EMD works for a general signal, based on Fourier representation. Overall, their analysis for EMD provides a significant contribution, but it inherits all limitations of Fourier transform. Hence, it is essential to investigate an ideal analysis on EMD that should deal with a general type of signals; for example, signals having time-varying frequency.

In this paper, we discuss the principle of EMD according to the ranges studied in Rilling and Flandrin [2008], mainly focusing on how the method works and why it does not. For this purpose, we take advantage of the concept of cubic spline interpolation that provides a direct interpretation of the EMD principle. The main contribution of the paper is twofold. First, we significantly relax the assumptions on the signal in order to obtain theoretical interpretation results of EMD, which include those of Rilling and Flandrin [2008]. Our analysis is not limited to a certain number of components, a constant frequency or amplitude, a sinusoidal signal, or the type of expansion (e.g., Fourier expansion). For example, the proposed analysis can deal with the following chirp signal $x(t) = \mathrm{HF} + \mathrm{LF} = 2\cos 2\pi(0.2t + 0.1t^2) + 3\cos 2\pi(0.05t)$, whereas Rilling and Flandrin [2008] cannot handle it. The only requirement of our analysis is that the HF component should be an IMF, which is a trivial and essential assumption. Second, our interpretation is discussed in the time domain, where the sifting process of EMD is performed. Consequently, the steps in the decomposition process can be visually depicted, so that it is useful in acquiring a deep understanding of EMD as well as in explaining the EMD principle.

The rest of this paper is organized as follows. Section 2 briefly reviews the EMD algorithm and a framework for understanding the EMD process. In Sec. 3, we investigate important facts to show that EMD can be explained by cubic spline interpolation, and discuss the principle of EMD by considering three cases with different ranges of frequency and amplitude ratios. Finally, concluding remarks are presented in Sec. 4.

2. Background

2.1. *Empirical mode decomposition*

Huang *et al.* [1998] defined an oscillatory wave to be an IMF if it satisfied two conditions: (1) the number of extrema and the number of zero crossings should be equal or differ by one and (2) at any point, the mean of the upper envelope and the lower envelope should be zero. Hence, the IMF is a locally zero-symmetric function to which Hilbert transform can be correctly applied to obtain the IF. The EMD process for obtaining the IMF can be summarized as follows: (1) identify the local extrema of the given signal $x(t)$; (2) construct the upper (lower) envelope by applying cubic spline interpolation to all the local maxima (minima), and let μ_0 be the mean envelope; (3) compute the difference $c_1(t) = x(t) - \mu_0(t)$; (4) perform steps (1)–(3) by considering c_1 as the given signal, determine $c_k = c_{k-1} - \mu_{k-1}$, and repeat the above steps until c_k satisfies the above-mentioned IMF conditions or a stopping rule; and (5) extract IMFs repeatedly until a monotone residue remains. Thus, the signal is finally decomposed into n IMFs imf_i, $i = 1, \ldots, n$ and a residue r_n, i.e., $x(t) = \sum_{i=1}^{n} \mathrm{imf}_i(t) + r_n(t)$.

2.2. *Framework for understanding the behavior of EMD*

Here, we briefly review the framework used by Rilling and Flandrin [2008] in their analysis. Such a review will be useful for explaining the relationship between the extrema in the signal and the frequency and amplitude ratios of the two components (HF and LF).

Consider a signal with two components. Let one component has an amplitude, frequency, and phase of a_1, f_1, and φ_1, respectively; let these parameters be a_2, f_2, and φ_2, respectively, in the case of the other component. Without the loss of generality, we can use the following expressions: $a = a_2/a_1$, $f = f_2/f_1 (f \in (0, 1))$, and $\varphi = \varphi_2 - \varphi_1$. Then, we obtain the simple continuous-time model

$$x(t; a, f) = \cos 2\pi t + a \cos(2\pi f t + \varphi), \quad t \in \mathbb{R}.$$

The terms $\cos 2\pi t$ and $a \cos(2\pi f t + \varphi)$ denote the HF and LF components, respectively.

Now, by simply following the derivation made by Rilling and Flandrin [2008], we consider the three cases such as: (1) $1 > af > af^2$, (2) $af > af^2 > 1$, and (3) $af > 1 > af^2$. In the first (second) case, the HF (LF) component determines the extrema of the signal. In these two cases, each extremum of the signal is distant from the corresponding extremum of the HF (LF) component by at most $1/(2\pi)\sin^{-1}(af)(1/(2\pi f)\sin^{-1}(1/af))$. In the third case, the extrema cannot be properly detected close to the extrema of HF or LF. For details, refer to Rilling and Flandrin [2008].

3. EMD Principle Based on the Cubic Spline Interpolation

In this section, we present an analysis on the EMD principle based on the cubic spline interpolation in the time domain; such an analysis provides a clear understanding of the meaning of EMD outputs as well as the possibilities and limitations of EMD. Cubic spline interpolation plays an important role in this theoretical analysis. We first discuss how the characteristics of a signal can be represented by cubic spline interpolation, which allows the analysis to be applied to signals other than sinusoidal signals. We then explain the behavior of EMD based on cubic spline interpolation within the framework for the three cases mentioned in the previous section.

3.1. *EMD and cubic spline interpolation*

Let the signal $x(t)$ be expressed as:

$$x(t) = h(t) + \ell(t), \tag{1}$$

where $h(t)$ represents the HF component of $x(t)$ and $\ell(t)$ denotes the sum of LF components of it; the IF of $h(t)$ at any instant is greater than the highest IF among LF components. The component(s) $h(t)$ and/or $\ell(t)$ can have a time-varying frequency or amplitude, and the only requirement is that $h(t)$ satisfies the IMF conditions.

Let $S[K,x](t)$ be the cubic spline interpolation of $x(t)$ with the set of time points of knots $K = \{t_j : j = 1, 2, \ldots\}$. Define the set of time points of the maxima in $x(t)$ as $M^x = \{m_j^x : j = 1, 2, \ldots\}$ and the set of time points of the minima as $N^x = \{n_j^x : j = 1, 2, \ldots\}$. Then, $S[M^x, x](t)$ represents the upper envelope of $x(t)$ in the EMD process, and similarly, $S[N^x, x](t)$ denotes the lower envelope of $x(t)$.

Proposition 1. *The* visible *upper* (*lower*) *envelope is the sum of the* hidden *cubic spline interpolations of* $h(t)$ *and* $\ell(t)$ *with knots that are the maxima* (*minima*) *of* $x(t)$. *That is*,

$$S[M^x, x](t) = S[M^x, h](t) + S[M^x, \ell](t), \tag{2}$$

$$S[N^x, x](t) = S[N^x, h](t) + S[N^x, \ell](t). \tag{3}$$

Proof. The upper envelope of $x(t)$ on the interval $[m_j^x, m_{j+1}^x]$ is formed by the two adjacent maxima $x(m_j^x)$ and $x(m_{j+1}^x)$. Let $d_j^x = m_{j+1}^x - m_j^x$. The upper envelope can be constructed as:

$$\begin{aligned} S[M^x, x](t) = {} & \frac{m_{j+1}^x - t}{d_j^x}\left[x(m_j^x) - \frac{d_j^x}{6}\gamma^x(m_{j+1}^x)\left\{(t - m_j^x) + \frac{(t-m_j^x)^2}{d_j^x}\right\}\right] \\ & + \frac{t - m_j^x}{d_j^x}\left[x(m_{j+1}^x) - \frac{d_j^x}{6}\gamma^x(m_j^x)\left\{(m_{j+1}^x - t) + \frac{(m_{j+1}^x - t)^2}{d_j^x}\right\}\right], \end{aligned} \tag{4}$$

where γ^x denotes the second derivative of the cubic spline, and hence, $\gamma^x(m_j^x)$ is an element of $\boldsymbol{\gamma}^x = R^{-1}Q^T\boldsymbol{x}$, where $\boldsymbol{\gamma}^x = (.., \gamma^x(m_j^x), ..)^T$ and $\boldsymbol{x} = (.., x(m_j^x), ..)^T$. Here, the matrix Q is an $n \times (n-2)$ matrix with entries q_{ij}, for $i = 1, \ldots, n$ and $j = 2, \ldots, n-1$, given by $q_{j-1,j} = 1/d_{j-1}^x, q_{jj} = -(1/d_{j-1}^x + 1/d_j^x), q_{j+1,j} = 1/d_j^x$, and $q_{ij} = 0$ for $|i - j| \geq 2$, and the symmetric matrix R is an $(n-2) \times (n-2)$ matrix with elements r_{ij}, for i and j taking values from 2 to $(n-2)$, given by $r_{ii} = (d_{i-1}^x + d_i^x)/3, r_{i,i+1} = r_{i+1,i} = d_i^x/6$, and $r_{ij} = 0$ for $|i-j| \geq 2$. For details on cubic splines, refer to Green and Silverman [1994] and Unser [1999]. Then, from Eq. (1),

$$\boldsymbol{\gamma}^x = R^{-1}Q^T\boldsymbol{x} = R^{-1}Q^T\boldsymbol{h} + R^{-1}Q^T\boldsymbol{\ell} = \boldsymbol{\gamma}^h + \boldsymbol{\gamma}^\ell, \tag{5}$$

where $\boldsymbol{\gamma}^h = (.., \gamma^h(m_j^x), ..)^T$ and $\boldsymbol{\gamma}^\ell = (.., \gamma^\ell(m_j^x), ..)^T$. Hence, from Eqs. (1) and (5), we obtain Eq. (2) by substituting $x(\cdot)$ with $h(\cdot) + \ell(\cdot)$ and $\gamma^x(\cdot)$ with $\gamma^h(\cdot) + \gamma^\ell(\cdot)$ in Eq. (4). Similarly, Eq. (3) is obtained. □

Next, we discuss the manner that the EMD output after n iterations can be represented by the original components and a linear combination of the spline interpolations of h and ℓ for each c_i, $i = 1, 2, \ldots, n-1$, where c_i denotes the component extracted after i iterations of the sifting process. This representation of the output is crucial for understanding the behavior of EMD for the three cases.

To clarify the above statement, we define $E_1^i = M^{c_i}$ and $E_2^i = N^{c_i}$ as the set of time points of the maxima and minima of the extracted signal $c_i(t)$. Note that c_0 is identical to the original signal $x(t)$. Thus, $E_1^0 = M^x$ and $E_2^0 = N^x$. In addition, let μ_i be the mean envelope of c_i, and $S[K_2, S[K_1, y]] = S[K_2K_1, y]$ denotes the spline interpolation of $S[K_1, y]$ with the set of knots K_2.

Proposition 2. *The component c_n extracted after n iterations of the sifting process can be expressed in terms of the HF component, the LF component, and a linear combination of spline interpolations of the two components in each iteration. That is,*

$$\begin{aligned} c_n &= c_0 - \sum_{i=0}^{n-1}\mu_i = h + \ell - \sum_{i=0}^{n-1}\mu_i, \\ &= h - \xi_{n-1}^h + \ell - \hat{\ell}_{n-1}, \end{aligned} \tag{6}$$

where ξ^h_{n-1} and $\hat{\ell}_{n-1}$ are defined as:

$$\xi^h_{n-1} = \sum_{i=0}^{n-1} \frac{1}{2} \sum_{k=1}^{2} S[E^i_k, h] - \sum_{i_2>i_1}^{n-1} \frac{1}{2^2} \sum_{k_2=1}^{2} \sum_{k_1=1}^{2} S[E^{i_2}_{k_2} E^{i_1}_{k_1}, h]$$
$$+ \sum_{i_3>i_2>i_1}^{n-1} \frac{1}{2^3} \sum_{k_3=1}^{2} \sum_{k_2=1}^{2} \sum_{k_1=1}^{2} S[E^{i_3}_{k_3} E^{i_2}_{k_2} E^{i_1}_{k_1}, h]$$
$$+ \cdots - (-1)^n \frac{1}{2^n} \sum_{k_n=1}^{2} \cdots \sum_{k_1=1}^{2} S[E^{n-1}_{k_n} \cdots E^0_{k_1}, h] \tag{7}$$

and

$$\hat{\ell}_{n-1} = \sum_{i=0}^{n-1} \frac{1}{2} \sum_{k=1}^{2} S[E^i_k, \ell] - \sum_{i_2>i_1}^{n-1} \frac{1}{2^2} \sum_{k_2=1}^{2} \sum_{k_1=1}^{2} S[E^{i_2}_{k_2} E^{i_1}_{k_1}, \ell]$$
$$+ \sum_{i_3>i_2>i_1}^{n-1} \frac{1}{2^3} \sum_{k_3=1}^{2} \sum_{k_2=1}^{2} \sum_{k_1=1}^{2} S[E^{i_3}_{k_3} E^{i_2}_{k_2} E^{i_1}_{k_1}, \ell]$$
$$+ \cdots - (-1)^n \frac{1}{2^n} \sum_{k_n=1}^{2} \cdots \sum_{k_1=1}^{2} S[E^{n-1}_{k_n} \cdots E^0_{k_1}, \ell]. \tag{8}$$

We remark that ξ^h_{n-1} consists of the interpolating cubic splines of $h, S[E^i_k, h]$ and repeated interpolating cubic spline results of the spline with the newly given knots in each iteration, $S[E^{i_2}_{k_2} E^{i_1}_{k_1}, h], S[E^{i_3}_{k_3} E^{i_2}_{k_2} E^{i_1}_{k_1}, h], \ldots, S[E^{n-1}_{k_n} \cdots E^0_{k_1}, h]$. $\hat{\ell}_{n-1}$ does similarly. We also note that each sum of the coefficients of ξ^h_{n-1} and $\hat{\ell}_{n-1}$ is ${}_nC_1 - {}_nC_2 + \cdots - (-1)^n\ {}_nC_n = 1$, and hence, the sums of the coefficients of $h - \xi^h_{n-1}$ and $\ell - \hat{\ell}_{n-1}$ are zero, respectively. The fact that the sums of coefficients are equal to zero is very important for providing an interpretation for the iterations during the EMD process. As we will discuss later, ξ^h_{n-1} denotes an error for an estimator of zero function and $\hat{\ell}_{n-1}$ is an estimator of ℓ or vice versa according to the cases.

Proof. From Proposition 1, the EMD output c_1 can be expressed as:

$$\begin{aligned} c_1 &= c_0 - \mu_0 \\ &= h - \frac{S[M^x, h] + S[N^x, h]}{2} + \ell - \frac{S[M^x, \ell] + S[N^x, \ell]}{2} \\ &= h - \xi^h_0 + \ell - \hat{\ell}_0. \end{aligned} \tag{9}$$

A term μ_1 is the mean envelope of $c_1 (= c_0 - \mu_0)$, which is the output of the first sifting process, and hence, it is defined as $\mu_1 = \frac{1}{2}\{S[M^{c_1}, c_1] + S[N^{c_1}, c_1]\}$. On the basis of Proposition 1, it follows that

$$\mu_1 = \frac{1}{2}\{S[M^{c_1}, c_0] - S[M^{c_1}, \mu_0] + S[N^{c_1}, c_0] - S[N^{c_1}, \mu_0]\}.$$

After applying Proposition 1 to each term of μ_1, we obtain $S[M^{c_1}, c_0] = S[M^{c_1}, x] = S[M^{c_1}, h] + S[M^{c_1}, \ell]$ and

$$\begin{aligned} S[M^{c_1}, \mu_0] &= \frac{1}{2}\{S[M^{c_1}, S[M^x, x] + S[N^x, x]]\} \\ &= \frac{1}{4}\{S[M^{c_1}, S[M^x, h]] + S[M^{c_1}, S[M^x, \ell]] \\ &\quad + S[M^{c_1}, S[N^x, h]] + S[M^{c_1}, S[N^x, \ell]]\}. \end{aligned}$$

Considering the similarity between the expressions of $S[N^{c_1}, c_0]$ and $S[N^{c_1}, \mu_0]$ and rearranging the terms, we obtain the extracted component c_2 as:

$$\begin{aligned} c_2 = h &- \left[\frac{1}{2}\{S[M^x, h] + S[N^x, h]\} + \frac{1}{2}\{S[M^{c_1}, h] + S[N^{c_1}, h]\}\right. \\ &- \frac{1}{4}\{S[M^{c_1}, S[M^x, h]] + S[M^{c_1}, S[N^x, h]] \\ &+ \left. S[N^{c_1}, S[M^x, h]] + S[N^{c_1}, S[N^x, h]]\}\right] \\ &+ \ell - \left[\frac{1}{2}\{S[M^x, \ell] + S[N^x, \ell]\} + \frac{1}{2}\{S[M^{c_1}, \ell] + S[N^{c_1}, \ell]\}\right. \\ &- \frac{1}{4}\{S[M^{c_1}, S[M^x, \ell]] + S[M^{c_1}, S[N^x, \ell]] \\ &+ \left. S[N^{c_1}, S[M^x, \ell]] + S[N^{c_1}, S[N^x, \ell]]\}\right] \\ &= h - \xi_1^h + \ell - \hat{\ell}_1. \end{aligned}$$

By letting $S[E_1^1 E_1^0, x] = S[M^{c_1}, S[M^x, x]]$ and repeatedly applying the above derivation for the next $n-1$ iterations, we obtain c_n. □

3.2. *Behavior of EMD in the framework*

Here, we explain the EMD principle by considering the following three cases that were taken into consideration by Rilling and Flandrin [2008] in their analysis. We note that our analysis allows a time-varying frequency and amplitude, i.e., a and f mean $a(t)$ and $f(t)$, respectively, and hence, each case represents local properties.

3.2.1. *Case $1 > af > af^2$ where the HF component determines the extrema of the signal*

Since EMD extracts the HF component, this case can be considered as *a normal case*. We perform an extensive analysis on EMD principle to figure out how the method works in this case. Moreover, we also investigate the situation where EMD does not work properly and the reason why it does not perform well.

3.2.1.1. Understanding of EMD at the first iteration

Since m_j^x is close to m_j^h in this case, the hidden cubic spline $S[M^x, h]$ in Eq. (9) can efficiently represent the upper envelope of $h, S[M^h, h]$. Similarly, $S[N^x, h]$ in Eq. (9) describes the lower envelope of h properly. On the other hand, the hidden components $S[M^x, \ell]$ and $S[N^x, \ell]$ in Eq. (9) are cubic spline interpolations of the LF component with knots in $\{m_j^x\}$ and $\{n_j^x\}$, respectively. The number of knots for the LF component representation is approximately given by $1/f$ for one oscillation of LF component in this case, and hence, these cubic spline interpolations can function as efficient estimators of the LF component for a low f value, $1/f > 3$. However, when constructing spline interpolation with a high f value, an undersampling problem happens.

The first column in Fig. 1 shows how EMD works well at the first iteration for a simulated signal whose HF is a linear chirp satisfying $1 > af > af^2$ with a low f value. As one can see in the figure, the interval between two consecutive knots decreases over time. We estimate and observe the red upper envelope of the signal, as shown in the first row. It is composed of two invisible spline estimators, $S[M^x, h]$ and $S[M^x, \ell]$, which are plotted in red color in the second and third panels. In order to identify the spline estimators of the LF component corresponding to the maxima and minima of $x(t), S[M^x, \ell]$ and $S[N^x, \ell]$, we separately plot them in the third and fourth panels. It is apparent that the cubic splines of the HF component represent their (true) symmetric envelopes well. Since f is fairly small in this example, the number of knots used for the spline interpolation of the LF component is sufficient to make the spline a good estimator of LF.

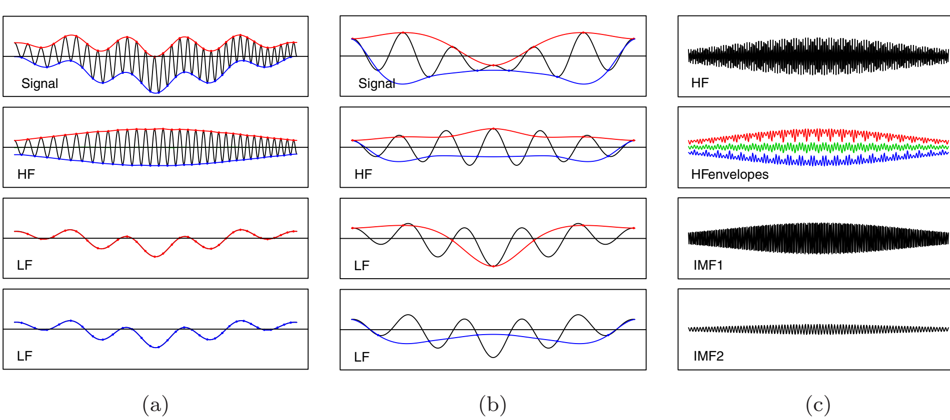

Fig. 1. Case $1 > af > af^2$: (a) EMD works well at the first iteration. (b) and (c) EMD does not work properly. The red and blue lines represent interpolating cubic splines with knots in M^x and N^x, respectively.

3.2.1.2. Averaging the envelopes of the signal

The assumption that $h(t)$ satisfies the IMF conditions implies that $S[M^h, h]$ and $S[N^h, h]$ are symmetric and their sum is zero. In order to remove the invisible spline estimators for the HF component, $S[M^x, h]$ and $S[N^x, h]$, the EMD process averages the upper and lower envelopes of $x(t)$ and determines $\xi_0^h(t)$ in $c_1(t)$ to be close to zero as shown by the green line in the second panel of the first column in Fig. 1. Thus, the spline interpolations of the HF component are removed with remaining the error, $\xi_0^h(t)$, which is caused by the difference between the extrema of $x(t)$ and those of the HF component. For the LF component, the EMD process produces the average of the LF estimators, $\hat{\ell}_0(t)$, as shown in the third and fourth panels. If, as in the present case, both $S[M^x, \ell]$ and $S[N^x, \ell]$ are good estimators of the LF component, then $\ell(t) - \hat{\ell}_0(t)$ must be close to zero.

3.2.1.3. Iteration effects

In the next iteration, the sifting process for $c_1(t) = h(t) + \ell(t) - \hat{\ell}_0(t) - \xi_0^h(t)$ offers the following advantages. First, the newly given value af between $h(t)$ (HF) and $\ell(t) - \hat{\ell}_0(t) - \xi_0^h(t)$ (LF) is relatively small, and hence, $m_j^{c_1}$ moves closer to m_j^h as iteration proceeds. Thus, the sum of $S[M^{c_1}, h]$ and $S[N^{c_1}, h]$ remains smaller error than that in the previous iteration. Second, the LF component in $c_1(t), \ell(t) - \hat{\ell}_0(t) - \xi_0^h(t)$, approaches a constant. Hence, when performing the cubic spline interpolation of the LF component in $c_1(t)$, the problem caused by undersampling, if any, can be relatively released. Therefore, by iteration, the EMD process continues to remove the LF component in $c_{n-1}(t), \ell(t) - \hat{\ell}_{n-2}(t) - \xi_{n-2}^h(t)$, more accurately.

3.2.1.4. Mathematical meaning of iterations

If we consider only the terms of i_1 in Eq. (7), each term consists of the average of envelopes. When $E_1^{i_1}$ and $E_2^{i_1}$ are appropriately detected, which means to be close enough to M^h and N^h, respectively, the envelopes obtained at the i_1th iteration ($S[E_1^{i_1}, h]$ and $S[E_2^{i_1}, h]$) and any sum of their repeated interpolating splines ($S[E_1^{i_k} \cdots E_1^{i_1}, h]$ and $S[E_1^{i_k} \cdots E_2^{i_1}, h]$) are symmetric each other, so that each term of i_1 is very close to zero by the IMF definition.

On the other hand, if $E_1^{i_1}$ and $E_2^{i_1}$ are not appropriate, especially when $i_1 = 0$, each term of i_1 produces some error. The EMD procedure has an internal system to be designed to cancel the errors by iteration. For the error by the sum of envelopes at the first i_1th iteration denoted by $\epsilon_{i_1} = (S[E_1^{i_1}, h] + S[E_2^{i_1}, h])/2$, an estimator of it, which is given by averaging re-estimated cubic splines of the first envelopes, $\hat{\epsilon}_{i_1} = (S[E_1^{i_2} E_1^{i_1}, h] + S[E_1^{i_2} E_2^{i_1}, h] + S[E_2^{i_2} E_1^{i_1}, h] + S[E_2^{i_2} E_2^{i_1}, h])/2^2$, is subtracted at the second iteration. At the third iteration, the estimators of the i_1th envelopes, $S[E_{k_3}^{i_3} E_{k_1}^{i_1}]$ and $S[E_{k_3}^{i_3} E_{k_2}^{i_3} E_{k_1}^{i_1}]$, are made but the sum of their coefficients is zero. Finally, it is expected that the error ϵ_{i_1} and its all estimators cancel out each other as iteration goes. The infinite sum of the error and its estimators could be nonzero in some cases. It could be a clue for analyzing the stopping rule for Huang *et al.* [1998] or enhancing EMD procedure, but this is beyond the scope of the paper.

For the LF component with an appropriate value of f, all terms in Eq. (8) obtained by applying spline interpolation recursively can be accurate estimators of LF. The terms of estimators in Eq. (8) are added or subtracted by iterations and the sum of their coefficients is 1. Consequently, ξ^h_{n-1} is designed to approximate zero, and $\hat{\ell}_{n-1}$ appropriately estimates ℓ in the case of $1 > af > af^2$. Hence, EMD works properly by making c_n be close to h in Eq. (6).

3.2.1.5. Discussion of the cases that EMD does not work properly

The undesirable EMD performance is mainly due to the poor approximation of the LF component or the incorrect representations of the envelopes of the HF component. We investigate these observations explicitly using the following two cases. First, when f is close to 1, the number of knots for one oscillation of LF can also decrease close to 1, and M^x and N^x at the first iteration do not provide sufficient number of knots for the spline approximation of $\ell(t)$. The middle column in Fig. 1 illustrates an example for insufficient number of knots. We generate a signal satisfying $1 > af > af^2$ with a high f value. The interpolating splines of the hidden LF component (the third and fourth panels) cannot be good estimators of the LF component because of undersampling problem for the spline construction. Even when we obtain the exact extrema of HF that is finally achieved in the EMD process, the value of f (and the number of knots for LF estimation) will be unlikely changed. Therefore, any single term in Eq. (8) cannot be the proper LF estimator. That is, EMD cannot appropriately work for a high f value because of undersampling problem in the LF estimation. Consequently, EMD cannot remove ℓ by its incorrect estimator $\hat{\ell}_{n-1}$, and, as an output, produces $h+\ell-\hat{\ell}_{n-1}$ instead of h even in the best case of $\xi^h_{n-1} = 0$. Second, if the frequency value of the HF component f_1 is very high, m^x_j cannot be appropriately close to m^h_j. The very tiny difference between m^x_j and m^h_j makes the spline component $S[M^x, h]$ a poor estimator of the upper envelope of the HF component $S[M^h, h]$ because of rapid decreasing or increasing of $h(t)$ (see the third column in Fig. 1). The term $\ell(t) - \hat{\ell}_0(t) - \xi^h_0(t)$ does not approximate zero because of the average of inaccurately estimated envelopes of the HF component, $\xi^h_0(t)$, which is shown by the green line in the second panel. After incorrectly extracting $\text{imf}_1 = h(t) - \xi^h_n(t) + \ell(t) - \hat{\ell}_n(t) \approx h(t) - \xi^h_n(t)$ as $h(t)$, the EMD process can extract $\xi^h_0(t)$ from $x(t) - \text{imf}_1 \approx \ell(t) + \xi^h_0(t)$ as imf_2 or imf_3, as shown in the third and fourth panels.

Moreover, employing smoothing spline instead of spline interpolation can facilitate a more accurate estimation of the envelopes of HF or lead to the envelopes being more symmetric to each other, and hence, in the EMD process, $\mu_0(t)$ can be considered as the proper estimator of the LF component; $\xi^h_0(t)$ is correctly determined as a zero function in the first iteration.

3.2.2. *Case $af > af^2 > 1$ where the LF component determines the extrema of the signal*

In this case, m_j^x is close to m_j^ℓ instead of m_j^h, and hence, the hidden cubic spline component $S[M^x, \ell]$ approximates $S[M^\ell, \ell]$, the upper envelope of $\ell(t)$. Thus, the term $\hat{\ell}_{n-1}$ of Eq. (6) should be replaced by ξ_{n-1}^ℓ when ℓ is an IMF. On the other hand, another hidden component, $S[M^x, h]$, is a spline estimator of the HF component with knots in M^x. Therefore, ξ_{n-1}^h of Eq. (6) is replaced with $\hat{h}_{n-1}$, and hence, the extracted signals can be expressed as $c_n = h - \hat{h}_{n-1} + \ell - \xi_{n-1}^\ell$ for the monocomponent ℓ. In this case, the number of knots in M^x is always less than the number of knots required for the proper spline interpolation of HF. Hence, $\hat{h}_{n-1}$ is obtained as a weighted average of spline estimators for HF with insufficient number of knots.

In general, consider the case that LF consists of several components as $\ell(t) = \ell_1(t) + \ell_2(t)$, where $\ell_1(t)$ and $\ell_2(t)$ are the HF component and the sum of LF components of $\ell(t)$, respectively, with satisfying $1 > af > af^2$, which represents the normal case. We now represent $c_n = h - \hat{h}_{n-1} + \ell_1 - \xi_{n-1}^{\ell_1} + \ell_2 - \hat{\ell}_{2,n-1}$. As explained in the normal case, $-\xi_{n-1}^{\ell_1} + \ell_2 - \hat{\ell}_{2,n-1}$ is supposed to approximate zero. On the other hand, $h - \hat{h}_{n-1} + \ell_1$ moves to achieve the IMF conditions by iterations because the sum of envelopes in the nth iteration approximates zero. That is, $S[E_1^n, h - \hat{h}_{n-1} + \ell_1] + S[E_2^n, h - \hat{h}_{n-1} + \ell_1] \approx 0$ when $E_k^{n-1} \approx \cdots \approx E_k^0 \approx E_k^{\ell_1}, k = 1, 2$. More specifically, in the example of the second iteration,

$$\begin{aligned} & S[M^{c_1}, h - \hat{h}_0 + \ell_1] + S[N^{c_1}, h - \hat{h}_0 + \ell_1] \\ & = \frac{1}{2}\{2S[M^{c_1}, h] - S[M^{c_1}M^x, h] - S[N^{c_1}M^x, h] \\ & \quad + 2S[N^{c_1}, h] - S[M^{c_1}N^x, h] - S[N^{c_1}N^x, h]\} \\ & \quad + S[M^{c_1}, \ell_1] + S[N^{c_1}, \ell_1] \\ & \approx 0, \end{aligned}$$

when $M^{c_1} \approx M^x \approx M^{\ell_1}$ and $N^{c_1} \approx N^x \approx N^{\ell_1}$. Since a spline estimator of HF with undersampled knots is very less oscillatory, the first three terms and the following three terms can be close to zero. The last two terms can approximate zero as ℓ_1 is an IMF. Finally, $h - \hat{h}_{n-1} + \ell_1$ becomes imf_1 and $x - c_n = \ell_2 + \hat{h}_{n-1}$ goes to the next step for extracting the second IMF. In other words, $h(t)$ cannot be separated from $\ell_1(t)$ (see the first column of Fig. 2). For better visualization, we put $\ell_2 = 0$. The green line in the second panel shows that $\hat{h}_0$ has the almost identical frequency to that of imf_2 in the fourth panel. The imf_1 is given as the result of $h + \ell_1 - \hat{h}_{n-1}$. When $\ell_2 \neq 0$, the component $\hat{h}_{n-1}$ can be extracted after imf_2 according to the new value af between $\hat{h}_{n-1}$ and $\ell_2(t)$.

Before concluding this case, we emphasize that the current analysis can provide a simple interpretation of the case where $\hat{h}_0$ (and finally $\hat{h}_{n-1}$) closely achieves zero and hence $h + \ell_1$ is obtained as the first IMF for simple sinusoids h and ℓ with

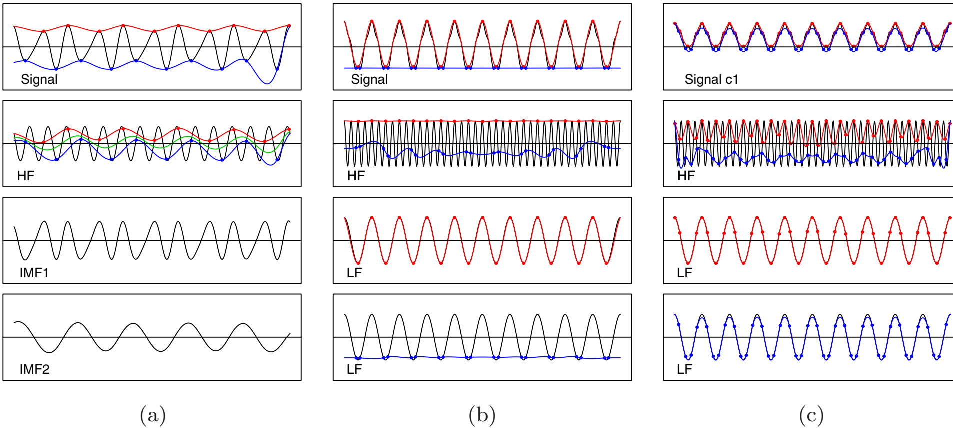

Fig. 2. (a) Case $af > af^2 > 1$: EMD does not separate HF. (b) and (c) Case $af > 1 > af^2$: EMD can have refinement effects — the first iteration vs. the second iteration. The panels for HF and IMF_2 are enlarged for visualization.

$f_2/f_1 = k$. In fact, Rilling and Flandrin [2008] performed a complicated and heavy analysis to interpret this case. In our analysis, it is required to check that

$$\frac{1}{2}\{\cos(2\pi f_1 m_j^x) + \cos(2\pi f_1(m_j^x + T_2))\} + \cos\left(2\pi f_1\left(m_j^x + \frac{T_2}{2}\right)\right) = 0,$$

where T_2 denotes the period of ℓ_2. The first two terms imply the linear interpolation of HF values at two consecutive knots given by the signal's maxima, and the last term denotes the HF value with knots given by the signal's minima between two maxima. We estimate the HF value at the minima using the HF value at the middle point of two maxima. Then, the above equation can be considered as a minimal necessary condition for $\hat{h}_0$ to be zero, and we obtain $\cos(\pi/k) = -1$ from simple calculations, which implies that $1/k$ should be odd number. Hence, when $1/k$ is close to odd number, $\hat{h}_{n-1}$ may not appear as an IMF component. This analysis can substitute the explanation for the simulation results obtained by Rilling and Flandrin [2008].

3.2.3. *Case $af > 1 > af^2$ where both components cannot determine the extrema of the signal*

This case represents that M^x or N^x follows neither the extrema of the HF component nor those of the LF component. In the normal case, we have mentioned that the undersampling problem for the LF component estimation is difficult to solve. Hence, we focus on cases where the knots do not cause undersampling problem in the LF estimation, which implies that f is sufficiently low.

From Eq. (9), the sifting process at the first iteration actually produces four hidden spline estimators. At the next iteration, each one is estimated again by spline interpolation with the updated extrema sets as explained in Sec. 3.2.1.4. for the normal case. The sum of coefficients of all estimators for $M^x = E_1^0$ after n (≥ 2) iterations in Eq. (7) is zero, because

$$\sum_{p=0}^{n-2} (-1)^p \ {}_{n-2}C_p \frac{1}{2^{p+1}} \ 2^p = \frac{1}{2} \sum_{p=0}^{n-2} (-1)^p \ {}_{n-2}C_p = 0, \quad n \geq 2.$$

Hence, each inaccurately estimated spline can be canceled by its repeatedly estimated splines with remaining negligible errors. For example, when $S[M^x, \ell]$ is incorrectly estimated, its effects on $\hat{\ell}_{n-1}$ is not significant once the EMD process iterates more than twice. On the other hand, the value af between h (HF) and $\ell - \xi_{i-1}^h - \hat{\ell}_{i-1}$ (LF) may finally drop down under 1 with iterations. Once af falls below 1, $S[E_k^i, h], S[E_k^i, \ell], k = 1, 2$, and their re-estimated splines after the ith iteration can perform as in the normal case. Therefore, in Eqs. (7) and (8), the incorrectly estimated splines are canceled by their repeatedly estimated splines with remaining negligible errors as iteration proceeds, whereas the appropriately estimated splines after the ith iteration perform well. The second and third columns in Fig. 2 illustrate this analysis. $S[M^x, h]$ and $S[M^x, \ell]$ are correctly estimated, but $S[N^x, h]$ and $S[N^x, \ell]$ are failed. In the second iteration, increasing number of knots by iteration enhances the estimators of LF, $S[E_1^1, \ell]$ and $S[E_2^1, \ell]$ to be accurate as shown in the third and fourth panels of the rightmost column. Incorrectly estimated splines obtained from each iteration can be eliminated at the subsequent iterations. Therefore, we observe some refinement effects for a low f value in the third case.

Finally, we state that, in the analysis for the normal case, we have approximately suggested $1/f > 3$ as the condition for obtaining the proper LF estimation by spline interpolation. For the high f value which provides the proper spline estimator according to knot locations, however, we can expect some refinement effects as explained in the third case. Hence, the simulation results for $1/3 < f < 2/3$ in Rilling and Flandrin [2008] can be explained in our analysis. A detailed analysis of f is beyond the scope of the present study.

4. Conclusion

In this paper, we provide a direct and simple interpretation of the EMD process by cubic spline interpolation. Using the framework involving three cases suggested by Rilling and Flandrin [2008], we discuss how EMD works and why EMD does not work properly. In addition, we explain EMD outputs as linear combinations of spline estimators and reveal where the undesirable components among EMD outputs are from. Furthermore, since the interpretation is performed in the time domain, it provides a direct insight into the EMD process. The summary of the analysis is given in Table 1. As a further aspect of this study, it seems plausible to extend the current analysis to two-dimensional EMD by incorporating thin-plate spline interpolation.

Table 1. Summary.

Case	Underlying principle of EMD	Details
$1 > af > af^2$	• Cubic spline interpolation provides proper estimation of LF and symmetric estimation of HF envelopes.	• There exist a certain range for f and an upper bound for f_1 where EMD works well. • When f is too high, EMD cannot work properly. • When f_1 is ultra high, an undesired IMF component ξ^h_{n-1} appears.
$af > af^2 > 1$	• Cubic spline interpolation represents LF envelopes effectively, but provides a poor estimation of HF due to undersampled knots.	• The sum of HF and LF, which is subtracted by incorrect estimator of HF, $h + \ell_1 - \hat{h}_{n-1}$, is misidentified as imf_1. • The incorrect estimator of HF with undersampled knots, $\hat{h}_{n-1}$, can appear as a new IMF component.
$af > 1 > af^2$	• By refinement effects obtained by iterations, EMD can extract the HF component as desired.	• For an appropriate range of the frequency ratio f, an incorrectly estimated term at an iteration can be removed at the subsequent iterations.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0076223 and 2009-0086944).

References

Boashash, B. (1992). Estimating and interpreting the instantaneous frequency of a signal — Part 1: fundamentals. *Proc. IEEE*, **80**: 519–538.

Green, P. J. and Silverman, B. W. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman & Hall, London.

Huang, N. E. and Shen, S. S. P. (2005). *Hilbert-Huang Transform and Its Applications*. World Scientific, Singapore.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung C. C. and Liu, H. H. (1998). The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. R. Soc. Lond. A*, **454**: 903–995.

Kopsinis, Y. and Mclaughlin, S. (2008). Improved EMD using double-iterative sifting and high order spline interpolation. *EURASIP J. Adv. Signal Process.*, **2008**, Article ID 128293, 1–8.

Rilling, G. and Flandrin, P. (2008). One or two frequencies? The empirical mode decomposition answers. *IEEE Trans. Signal Process.*, **56**: 85–95.

Unser, M. (1999). Splines: a perfect fit for signal processing. *IEEE Signal Process. Mag.*, **16**: 22–38.