Advances in Adaptive Data Analysis
Vol. 1, No. 4 (2009) 561–571

CONVERGENCE OF A CONVOLUTION-FILTERING-BASED ALGORITHM FOR EMPIRICAL MODE DECOMPOSITION

CHAO HUANG and LIHUA YANG*

School of Mathematics and Computing Science
Sun Yat-sen (Zhongshan) University, P. R. China
**mcsylh@mail.sysu.edu.cn*

YANG WANG

Department of Mathematics
Michigan State University, USA

Lin *et al.* propose the *iterative Toeplitz filters* algorithm as an alternative iterative algorithm for Empirical Mode Decomposition (EMD). In this alternative algorithm, the average of the upper and lower envelopes is replaced by certain "moving average" obtained through a low-pass filter. Performing the traditional sifting algorithm with such moving averages is equivalent to iterating certain convolution filters (finite length Toeplitz filters). This paper studies the convergence of this algorithm for signals of continuous variables, and proves that the limit function of this iterative algorithm is an ideal high-pass filtering process.

Keywords: Convolution filter; iterative Toeplitz filter; EMD algorithm.

1. Introduction

Signal and data analysis has been playing a very important role in practical applications. It serves two purposes: to determine the parameters needed for the construction of a necessary model, and to conform if the model constructed represents the physical phenomenon. Traditional data analysis methods such as Fourier analysis, based on the linear stationary assumption have been shown to be efficient for processing of linear and stationary data. However, data from real systems, either natural or man-made ones, are often both nonlinear and nonstationary. Many studies have shown that the traditional data analysis methods are not suitable for analyzing nonlinear and nonstationary data. Other methods have been introduced to analyze nonstationary and nonlinear data. For example, wavelet analysis and Wigner–Ville distribution are designed for linear but nonstationary data. Meanwhile, various nonlinear time series analysis methods are designed for nonlinear but stationary and deterministic systems.

In 1998, Huang *et al.* presented a new time–frequency algorithm for nonlinear and nonstationary signal analysis: Hilbert–Huang Transform (HHT).[3,4] It consists of two procedures: the Empirical Mode Decomposition (EMD) and the Hilbert spectrum. With EMD, any complicated data set can be decomposed into a finite and often small number of components called Intrinsic Mode Functions (IMFs). With the Hilbert transform, the IMFs yield instantaneous frequencies as functions of time that give sharp identification of imbedded structures. The final presentation of the results is an energy–frequency–time distribution, designated as the Hilbert spectrum, which possesses high time–frequency locality. HHT is a method for signal analysis based on the local characteristic time scale of the data, rather than basis functions given in advance. Since it expands the data in a basis derived from the data adaptively, it often leads to very useful decompositions. Fourier analysis, wavelet analysis, Short-Time Fourier analysis and so on comparatively expand the data in certain basis chosen in advance, and do not have the adaptivity property. EMD produces basis adaptively according to the wave shape of signals during the process, and therefore is applicable to nonlinear and nonstationary processes. Its applications have spread from ocean science, earthquake research, biomedicine, physics and so on.[3,6,7]

The IMFs are obtained through the EMD algorithm which is an iteration process. The local maxima and minima of a function (signal) are, respectively, connected via cubic splines to form the so-called upper and lower envelops. The average of the two envelopes is then subtracted from the original data. This process is iterated to obtain the first IMF. The other IMFs are obtained by the same process on the residual signal. The EMD algorithm is highly adaptive. A small perturbation, however, can alter the envelopes dramatically, raising some questions about its stability. As powerful as EMD is in many applications, a mathematical foundation is virtually nonexistent. Many fundamental mathematical issues such as the convergence of the EMD algorithm have never been established.

In Ref. 1 the cubic splines were replaced by B-splines, which gave an alternative way for EMD. But the modification relying on the local extrema is still a highly nonlinear process, so again did not resolve those mathematical issues.

To build a mathematical foundation, an alternative approach for EMD was proposed by Lin *et al.*[5] Instead of the average of the upper and lower envelopes, this new approach uses certain Toeplitz operators. These operators are low-pass convolution filters that yield a "moving average" similar to the mean of the envelopes in the original EMD algorithm. It was demonstrated in Ref. 5 that this new approach often leads to comparable EMD as the classical EMD, and in general it serves as a useful alternative or complement. For a periodic signal the convergence is completely characterized in Ref. 5. However, the convergence for a nonperiodic signal in $l^\infty(\mathbb{Z})$ is a much more difficult problem. Convergence results in this setting were obtained by Wang and Zhou.[9]

This paper studies the corresponding model of continuous variables. Let $\mathbb{R}$ denote the set of all real numbers and $\mathbb{Z}$ the set of all integers, set $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$. $\mu_I(t)$ is a measure on I defined as $d\mu_I(t) := dt$ for $I = \mathbb{R}$ and $d\mu_I(t) := (1/2\pi)dt$

for $I = \mathbb{T}$. For $1 \le p \le \infty$, $L^p(I)$ stands for the space of functions whose pth powers are integrable on I endowed with the following norm

$$\|f\|_{L^p(I)} := \begin{cases} \left(\int_I |f(t)|^p d\mu_I(t)\right)^{1/p}, & 1 \le p < \infty, \\ \mathrm{essup}_{t\in I}|f(t)|, & p = \infty. \end{cases} \tag{1.1}$$

For $a \in L^1(I)$, define the convolution operator by

$$T_a f(x) := (a * f)(x) := \int_I a(t) f(x-t) d\mu_I(t), \quad \forall\, f \in L^p(I). \tag{1.2}$$

We know that, $\forall\, f \in L^p(I)$, there holds $\|T_a f\|_p \le \|a\|_1 \cdot \|f\|_p$, so T_a is a bounded linear operator on $L^p(I)$.

In this paper, we study the convergence of $\{(I - T_a)^n f\}$ in $L^p(I)$ and the corresponding limit functions. In Sec. 2, we study the convergence for periodic functions, that is $I = \mathbb{T}$. In Sec. 3, we study the convergence for non-periodic functions, that is $I = \mathbb{R}$. We give convergence conditions in L^p and corresponding limit functions, and analyze the convergence rate.

2. Convergence for Periodic Signals

Given a sequence $\{c_k\}_{k\in\mathbb{Z}}$, we define

$$\|\{c_k\}\|_p := \begin{cases} \left(\sum_{k\in\mathbb{Z}} |c_k|^p\right)^{1/p} < \infty, & 1 \le p < \infty; \\ \sup_{k\in\mathbb{Z}} |c_k|, & p = \infty. \end{cases} \tag{2.1}$$

Denote by $l^p(\mathbb{Z})$ the space of all sequences $\{c_k\}_{k\in\mathbb{Z}}$ satisfying $\|\{c_k\}\|_p < +\infty$. It is well-known that $l^p(\mathbb{Z})$ is a Banach space endowed with the norm (2.1).

For any $f \in L^1(\mathbb{T})$, we call

$$\hat{f}(k) := \frac{1}{2\pi}\int_{\mathbb{T}} f(t)\, e^{-ikt} dt, \quad k \in \mathbb{Z}, \tag{2.2}$$

the kth Fourier coefficient of f. Let $f, a \in L^1(\mathbb{T})$, then $\forall\, k \in \mathbb{Z}$, there holds $(T_a f)\hat{}(k) = \hat{a}(k)\hat{f}(k)$, which yields

$$[(I - T_a)^n f]\hat{}(k) = [1 - \hat{a}(k)]^n \hat{f}(k). \tag{2.3}$$

Next we discuss the convergence of $\{(I - T_a)^n f\}_{n\in\mathbb{N}}$. For simplicity, given an $a \in L^1(\mathbb{T})$, set

$$\begin{cases} \Lambda_{-1} := \{k \in \mathbb{Z}\colon |1 - \hat{a}(k)| < 1\}, \\ \Lambda_0 := \{k \in \mathbb{Z}\colon \hat{a}(k) = 0\}, \\ \Lambda_1 := \{k \in \mathbb{Z}\colon |1 - \hat{a}(k)| \ge 1,\ \hat{a}(k) \ne 0\}. \end{cases} \tag{2.4}$$

2.1. Convergence of $\{(I - T_a)^n f\}$ in $L^2(\mathbb{T})$

It is well-known that $\{e^{ikt}\}_{k\in\mathbb{Z}}$ constitutes an orthonormal basis of $L^2(\mathbb{T})$, and that if $f \in L^2(\mathbb{T})$ then it has the Fourier expansion:

$$f(t) = \sum_{k\in\mathbb{Z}} \hat{f}(k)e^{ikt}, \tag{2.5}$$

in $L^2(\mathbb{T})$. With this expansion, $f \mapsto \{\hat{f}(k)\}_{k\in\mathbb{Z}}$ is an isomorphism from $L^2(\mathbb{T})$ to $l^2(\mathbb{Z})$, and the Parseval equality $\|f\|_{L^2(\mathbb{T})} = \|\{\hat{f}(k)\}\|_2$ holds.

Theorem 2.1. *Let $f \in L^2(\mathbb{T})$ and $a \in L^1(\mathbb{T})$. Then $\{(I - T_a)^n f\}_{n\in\mathbb{N}}$ converges in $L^2(\mathbb{T})$ if and only if $\hat{f}(k) = 0$ holds for any $k \in \Lambda_1$. In this case,*

$$\lim_{n\to\infty} (I - T_a)^n f = \sum_{k\in\Lambda_0} \hat{f}(k)e^{ikt}.$$

Proof. The convergence of $\{(I - T_a)^n f\}_{n\in\mathbb{N}}$ in $L^2(\mathbb{T})$ is equivalent to the convergence of its Fourier coefficients $\{\{[(I - T_a)^n f]\hat{}(k)\}_{k\in\mathbb{Z}}\}_{n\in\mathbb{N}}$ in $l^2(\mathbb{Z})$. $\forall\, n \in \mathbb{N}, k \in \mathbb{Z}$, we have

$$[(I - T_a)^n f]\hat{}(k) = [1 - \hat{a}(k)]^n \hat{f}(k).$$

If $\{(I-T_a)^n f\}_{n\in\mathbb{N}}$ converges in $L^2(\mathbb{T})$, then $\{\{[1-\hat{a}(k)]^n \hat{f}(k)\}_{k\in\mathbb{Z}}\}_{n\in\mathbb{N}}$ converges in $l^2(\mathbb{Z})$. It follows that $\forall\, k \in \mathbb{Z}$, the sequence of number $\{[1 - \hat{a}(k)]^n \hat{f}(k)\}_{n\in\mathbb{Z}}$ converges. Consequently, the difference between two adjacent terms of this sequence of number converges to 0, that is, $\forall\, k \in \mathbb{Z}$, there holds

$$[1 - \hat{a}(k)]^n \hat{a}(k)\hat{f}(k) \to 0, \quad n \to \infty.$$

Thus $\hat{a}(k)\hat{f}(k) = 0$ $(\forall\, k \in \Lambda_0 \cup \Lambda_1)$, which implies $\hat{f}(k) = 0$ $(\forall\, k \in \Lambda_1)$.

Conversely, suppose $\hat{f}(k) = 0$ $(\forall\, k \in \Lambda_1)$. Denote $c_k := \hat{f}(k)\chi_{\Lambda_0}(k)$ $(\forall\, k \in \mathbb{Z})$, where χ_{Λ_0} stands for the characteristic function of set Λ_0. Then

$$\sum_{k\in\mathbb{Z}} |[1 - \hat{a}(k)]^n \hat{f}(k) - c_k|^2 = \sum_{k\in\Lambda_{-1}} |[1 - \hat{a}(k)]^n \hat{f}(k)|^2.$$

$\forall\, \epsilon > 0$, choose $K \in \mathbb{N}$ such that $\sum_{|k|>K} |\hat{f}(k)|^2 < \epsilon$. Then

$$\begin{aligned}\sum_{k\in\mathbb{Z}} |[1 - \hat{a}(k)]^n \hat{f}(k) - c_k|^2 &< \sum_{k\in\Lambda_{-1}, |k|\le K} |[1 - \hat{a}(k)]^n \hat{f}(k)|^2 + \epsilon \\ &\le \lambda^{2n} \|\{\hat{f}(k)\}\|_2^2 + \epsilon,\end{aligned}$$

where $\lambda := \max_{k\in\Lambda_{-1}, |k|\le K} |1 - \hat{a}(k)| < 1$. Thus $\{\{[1 - \hat{a}(k)]^n \hat{f}(k)\}_{k\in\mathbb{Z}}\}_{n\in\mathbb{N}}$ converges to $\{c_k\}_{k\in\mathbb{Z}}$ in $l^2(\mathbb{Z})$, which implies $\{(I - T_a)^n f\}_{n\in\mathbb{N}}$ converges to $\sum_{k\in\mathbb{Z}} c_k e^{ikt} = \sum_{k\in\Lambda_0} \hat{f}(k)e^{ikt}$ in $L^2(\mathbb{T})$. $\square$

Corollary 2.2. *Let $a \in L^1(\mathbb{T})$ with $|1 - \hat{a}(k)| < 1$ or $\hat{a}(k) = 0$ for all $k \in \mathbb{Z}$. Then $\{(I - T_a)^n f\}$ converges to $\sum_{k\in\Lambda_0} \hat{f}(k)e^{ikt}$ in $L^2(\mathbb{T})$ for all $f \in L^2(\mathbb{T})$.*

2.2. *Convergence of $\{(I - T_a)^n f\}$ in $L^p(\mathbb{T})$, $1 \le p \le \infty$*

We first discuss the necessary condition for the convergence of $\{(I - T_a)^n f\}$ in $L^p(\mathbb{T})(1 \le p \le \infty)$.

Theorem 2.3. *Let $f \in L^p(\mathbb{T})$ where $1 \le p \le \infty$ and $a \in L^1(\mathbb{T})$. If $\{(I - T_a)^n f\}$ converges in $L^p(\mathbb{T})$, then $\hat{f}(k) = 0$ holds for any $k \in \Lambda_1$.*

Proof. Since the convergence of $\{(I - T_a)^n f\}$ in $L^p(\mathbb{T})$ implies its convergence in $L^1(\mathbb{T})$, without loss of generality we assume that $p = 1$. $\forall\, k \in \mathbb{Z}$, we have

$$\begin{aligned}|[1 - \hat{a}(k)]^{n+1}\hat{f}(k) - [1 - \hat{a}(k)]^n\hat{f}(k)| &= |[(I - T_a)^{n+1}f]\hat{\ }(k) - [(I - T_a)^n f]\hat{\ }(k)| \\ &\le \|(I - T_a)^{n+1}f - (I - T_a)^n f\|_{L^1(\mathbb{T})} \to 0\end{aligned}$$

as $n \to \infty$, that is $|[1 - \hat{a}(k)]^n\hat{a}(k)\hat{f}(k)| \to 0$. It follows that $\hat{f}(k) = 0$ for all $k \in \Lambda_1$. □

Below is a sufficient condition for convergence:

Theorem 2.4. *Let $a \in L^1(\mathbb{T})$ and $f \in L^p(\mathbb{T})$ with $1 \le p \le \infty$ such that $\hat{f}(k) = 0$ for all $k \in \Lambda_1$ and $\sum_{k\in\Lambda_{-1}} |\hat{f}(k)| < \infty$. Then there exists an $f_a \in L^p(\mathbb{T})$, such that $\hat{f}_a(k) = \hat{f}(k)\chi_{\Lambda_0}(k)$ and*

$$\|(I - T_a)^n f - f_a\|_{L^\infty(\mathbb{T})} \to 0, \quad (n \to \infty).$$

Proof. For all $n \in \mathbb{N}$, denote

$$f_n(x) := \sum_{k\in\Lambda_{-1}} [1 - \hat{a}(k)]^n\hat{f}(k)e^{ikx},$$

then f_n is a continuous function on $\mathbb{T}$, and $\hat{f}_n(k) = [1-\hat{a}(k)]^n\hat{f}(k)\chi_{\Lambda_{-1}}(k)$ $(\forall\, k \in \mathbb{Z})$. Thus $\forall\, k \in \mathbb{Z}$, we have

$$[(I - T_a)^n f - f_n]\hat{\ }(k) = [1 - \hat{a}(k)]^n\hat{f}(k)[1 - \chi_{\Lambda_{-1}}(k)] = \hat{f}(k)\chi_{\Lambda_0}(k).$$

The right side of the above equation is clearly independent of n, so

$$[(I - T_a)^n f - f_n]\hat{\ }(k) = [(I - T_a)f - f_1]\hat{\ }(k), \quad \forall\, k \in \mathbb{Z},\ n \in \mathbb{Z}.$$

By the uniqueness of Fourier coefficients of functions in $L^p(\mathbb{T})$, we have $(I - T_a)^n f - f_n = (I - T_a)f - f_1$ $(\forall\, n \in \mathbb{N})$. Now denote $f_a := (I - T_a)f - f_1$, then $f_a \in L^p(\mathbb{T})$ and $\hat{f}_a(k) = \hat{f}(k)\chi_{\Lambda_0}(k)$. There holds

$$\begin{aligned}\|(I - T_a)^n f - f_a\|_{L^\infty(\mathbb{T})} &= \|f_n\|_{L^\infty(\mathbb{T})} \\ &\le \sum_{k\in\Lambda_{-1}} |1 - \hat{a}(k)|^n|\hat{f}(k)| \to 0 \quad (n \to \infty).\end{aligned}$$

□

Remark.

(1) For $1 < p \leq \infty$, Carleson's theorem shows that for any function in $L^p(\mathbb{T})$, its Fourier series converges to itself almost everywhere, so $f_a(x) = \sum_{k\in\Lambda_0} \hat{f}(k)e^{ikx}$ a.e. $x \in \mathbb{T}$. For $p = 1$, $s_n(x) := \sum_{k\in\Lambda_0,|k|\leq n} \hat{f}(k)e^{ikx}$ is the partial Fourier sum of f_a, and $\sigma_n(x) := 1/(n+1)\sum_{k=0}^{n} s_n(x) = \sum_{k\in\Lambda_0,|k|\leq n}(1 - |k|/(n+1))\hat{f}(k)e^{ikx}$ is called the (C, 1) sum of f_a. Classical Fourier analysis points out that any function in $L^1(\mathbb{T})$ can be expressed as the pointwise limit of its (C, 1) sum almost everywhere (see more details in Ref. 8). Thus, $f_a(x) = \lim_{n\to\infty} \sum_{k\in\Lambda_0,|k|\leq n} (1 - |k|/(n+1))\hat{f}(k)e^{ikx}$ a.e. $x \in \mathbb{T}$.

(2) Under the condition of the theorem, it is obvious that $\|(I-T_a)^n f - f_a\|_{L^p(\mathbb{T})} \to 0$ $(n \to \infty)$.

(3) By the expression of the limit function f_a, we conclude that the consequence of iteration is a filtering process using the ideal filter χ_{Λ_0}. When a is low-pass filter, the consequence of iteration is an ideal high-pass filtering process; when a is high-pass filter, the consequence of iteration is an ideal low-pass filtering process.

When a is a trigonometric polynomial, $\{k \in \mathbb{Z} : \hat{a}(k) \neq 0\}$ is clearly a set with finite elements, thus so is its subset Λ_{-1}. At this time, $\sum_{k\in\Lambda_{-1}} |\hat{f}(k)| < +\infty$ holds naturally, so we have the following corollary.

Corollary 2.5. *Let $a \in L^1(\mathbb{T})$ be a trigonometric polynomial and $f \in L^p(\mathbb{T})$, $1 \leq p \leq \infty$. Then $\{(I - T_a)^n f\}_{n\in\mathbb{N}}$ converges in $L^p(\mathbb{T})$ if and only if $\hat{f}(k) = 0$ holds for any $k \in \Lambda_1$. Particularly, if $|1 - \hat{a}(k)| < 1$ or $\hat{a}(k) = 0$ hold for all $k \in \mathbb{Z}$, then $\{(I - T_a)^n f\}_{n\in\mathbb{N}}$ converges in $L^p(\mathbb{T})$ for all $f \in L^p(\mathbb{T})$.*

2.3. *Convergence rate of* $\{(I - T_a)^n f\}_{n\in\mathbb{N}}$

Theorem 2.6. *Let $a \in L^1(\mathbb{T})$ be an even real function satisfying either $|1-\hat{a}(k)| < 1$ or $\hat{a}(k) = 0$ for all $k \in \mathbb{Z}$. Assume that $f \in L^p(\mathbb{T})$, $1 \leq p \leq \infty$, satisfies $\sum_{k\in\Lambda_{-1}} |\hat{f}(k)| < +\infty$. Then there exists an $f_a \in L^p(\mathbb{T})$ such that $\hat{f}_a(k) = \hat{f}(k)\chi_{\Lambda_0}(k)$. Furthermore, there exists $0 < \delta < 1$ such that for all $n \in \mathbb{N}$,*

$$\|(I - T_a)^n f - f_a\|_{L^\infty(\mathbb{T})} \leq (1-\delta)^n \sum_{k\in\Lambda_{-1}} |\hat{f}(k)| + \sum_{k\in\mathbb{Z},0<\hat{a}(k)<\delta} |\hat{f}(k)|.$$

Proof. By the condition of the theorem, it is clear that $0 \leq \hat{a}(k) < 2$ $(\forall\, k \in \mathbb{Z})$. Since $\lim_{|k|\to\infty} \hat{a}(k) = 0$, there exists $0 < \delta < 1$, such that $0 \leq \hat{a}(k) \leq 2-\delta$ $(\forall\, k \in \mathbb{Z})$. By the proof of theorem 2.4, we have

$$\begin{aligned}\|(I - T_a)^n f - f_a\|_{L^\infty(\mathbb{T})} &\leq \sum_{k\in\Lambda_{-1}} |1 - \hat{a}(k)|^n |\hat{f}(k)| \\ &\leq (1-\delta)^n \sum_{k\in\Lambda_{-1}} |\hat{f}(k)| + \sum_{k\in\mathbb{Z},0<\hat{a}(k)<\delta} |\hat{f}(k)|.\end{aligned}$$

□

This theorem shows that the less information f possesses in the frequency domain $\{k \in \mathbb{Z} : 0 < \hat{a}(k) < \delta\}$, the faster convergence rate the iteration has. If the set $\{k \in \mathbb{Z} : \hat{a}(k) \neq 0\}$ is of finite elements, as long as δ is sufficiently small, $\{k \in \mathbb{Z} : 0 < \hat{a}(k) < \delta\}$ is an empty set, thus $\|(I - T_a)^n f - f_a\|_{L^\infty(\mathbb{T})}$ decays in exponential order $(1-\delta)^n$.

3. Convergence for Non-Periodic Signals

For any $f \in L^1(\mathbb{R})$ we shall use

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(t)\, e^{-i\xi t} dt, \quad \xi \in \mathbb{R}, \tag{3.1}$$

to denote the Fourier transform of f. Utilizing the fact that $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is a dense linear subspace of $L^2(\mathbb{R})$, the definition of the Fourier transform can be extended to $L^2(\mathbb{R})$, and furthermore, to $L^1(\mathbb{R}) + L^2(\mathbb{R})$. Since $L^p(\mathbb{R}) \subset L^1(\mathbb{R}) + L^2(\mathbb{R}) := \{f + g : f \in L^1(\mathbb{R}),\ g \in L^2(\mathbb{R})\}$, the Fourier transform of $f \in L^p(\mathbb{R})$ is well-defined for $1 \le p \le 2$, with the Hausdorff–Young inequality $\|\hat{f}\|_{L^q(\mathbb{R})} \le (2\pi)^{1/q} \|f\|_{L^p(\mathbb{R})}$ holds, where $2 \le q \le \infty$ satisfying $p^{-1} + q^{-1} = 1$.[8]

For all $a \in L^1(\mathbb{R})$ and $f \in L^p(\mathbb{R})$ with $1 \le p \le 2$ we have $(T_a f)\hat{} = \hat{a}\hat{f}$. Thus

$$[(I - T_a)^n f]\hat{}(\xi) = [1 - \hat{a}(\xi)]^n \hat{f}(\xi). \tag{3.2}$$

For simplicity, given an $a \in L^1(\mathbb{R})$ we set

$$\begin{cases} E_{-1} := \{\xi \in \mathbb{R}\colon |1 - \hat{a}(\xi)| < 1\}, \\ E_0 := \{\xi \in \mathbb{R}\colon \hat{a}(\xi) = 0\}, \\ E_1 := \{\xi \in \mathbb{R}\colon |1 - \hat{a}(\xi)| \ge 1,\ \hat{a}(\xi) \neq 0\}. \end{cases} \tag{3.3}$$

3.1. *Convergence of $\{(I - T_a)^n f\}$ in $L^2(\mathbb{R})$*

Below we discuss the convergence of $\{(I - T_a)^n f\}$ in $L^2(\mathbb{R})$. We have the following result.

Theorem 3.1. *Let $a \in L^1(\mathbb{R})$ and $f \in L^2(\mathbb{R})$. Then $\{(I - T_a)^n f\}$ converges in $L^2(\mathbb{R})$ if and only if $\hat{f}(\xi) = 0$ a.e. $\xi \in E_1$. In this case, the limit of $\{(I - T_a)^n f\}$ is the inverse Fourier transform of $\hat{f}\chi_{E_0}$, where χ_{E_0} denotes the characteristic function of set E_0.*

Proof. Since Fourier transform is an isomorphism on $L^2(\mathbb{R})$, the convergence of $\{(I - T_a)^n f\}$ is equivalent to the convergence of its Fourier transform $\{[(I - T_a)^n f]\hat{}\}$.

If $\{[1 - \hat{a}]^n \hat{f}\}$ converges in $L^2(\mathbb{R})$, by Riesz's theorem,[2] there exists a subsequence $\{[1 - \hat{a}]^{n_k} \hat{f}\}_{k=1}^\infty$ which converges almost everywhere on $\mathbb{R}$. Clearly, $\{[1 - \hat{a}]^{n_k+1} \hat{f}\}_{k=1}^\infty$ converges in $L^2(\mathbb{R})$, so again there exists a subsequence of $\{[1 - \hat{a}]^{n_k+1} \hat{f}\}_{k=1}^\infty$ which converges almost everywhere on $\mathbb{R}$, without loss of generality we assume that

$\{[1-\hat{a}]^{n_k+1}\hat{f}\}_{k=1}^{\infty}$ converges almost everywhere on $\mathbb{R}$. Then, the difference between $[1-\hat{a}]^{n_k+1}\hat{f}$ and $[1-\hat{a}]^{n_k}\hat{f}$ converges to 0 almost everywhere on $\mathbb{R}$. That is

$$|[1-\hat{a}(\xi)]^{n_k}\hat{a}(\xi)\hat{f}(\xi)| \to 0, \quad \text{a.e. } \xi \in \mathbb{R}.$$

It follows that $\hat{a}(\xi)\hat{f}(\xi) = 0$ a.e. $\xi \in E_0 \cup E_1$, thus $\hat{f}(\xi) = 0$ a.e. $\xi \in E_1$.

Conversely, suppose $\hat{f}(\xi) = 0$ a.e. $\xi \in E_1$. Denote $\hat{f}_a := \hat{f}\chi_{E_0}$. Then

$$\|(I-T_a)^n f - f_a\|_{L^2(\mathbb{R})}^2 = \|[1-\hat{a}]^n\hat{f} - \hat{f}_a\|_{L^2(\mathbb{R})}^2 = \int_{E_{-1}} |[1-\hat{a}(\xi)]^n|^2 \cdot |\hat{f}(\xi)|^2\, d\xi.$$

By the Lebesgue dominated convergence theorem, there holds $\lim_{n\to\infty} \|(I-T_a)^n f - f_a\|_{L^2(\mathbb{R})} = 0$. □

Based on the theorem above, we immediately have the following corollary.

Corollary 3.2. *Let $a \in L^1(\mathbb{R})$ such that either $|1-\hat{a}(\xi)| < 1$ or $\hat{a}(\xi) = 0$ a.e. $\xi \in \mathbb{R}$. Then $\{(I-T_a)^n f\}$ converges to the inverse Fourier transform of $\hat{f}\chi_{E_0}$ in $L^2(\mathbb{R})$ for all $f \in L^2(\mathbb{R})$.*

When the iterative algorithm is convergent, by the expression of the limit function, the consequence of iteration is essentially a reverse filtering process with regard to a. That is, when a is a low-pass filter, the consequence of iteration is a high-pass filtering process; conversely, when a is a high-pass filter, the consequence of iteration is a low-pass filtering process.

3.2. *Convergence of $\{(I-T_a)^n f\}_{n\in\mathbb{N}}$ in the Fourier domain*

The theorem below discusses the convergence of $\{(I-T_a)^n f\}_{n\in\mathbb{N}}$ in Fourier domain, that is, the convergence of $\{[(I-T_a)^n f]\hat{}\,\}_{n\in\mathbb{N}}$.

Theorem 3.3. *Let $a \in L^1(\mathbb{R})$ and $f \in L^p(\mathbb{R})$ with $1 \le p \le 2$. Then $\{[(I-T_a)^n f]\hat{}\,\}_{n\in\mathbb{N}}$ converges in $L^q(\mathbb{R})$, $p^{-1}+q^{-1} = 1$, if and only if $\hat{f}(\xi) = 0$ a.e. $\xi \in E_1$. In this case the limit is $\hat{f}\chi_{E_0}$.*

Proof.

Necessity: the proof is similar to the proof of necessity of Theorem 3.1.

Sufficiency: If $\hat{f}(\xi) = 0$ a.e. $\xi \in E_1$, it is easy to check that

$$[(I-T_a)^n f]\hat{}\,(\xi) - \hat{f}(\xi)\chi_{E_0}(\xi) = [1-\hat{a}(\xi)]^n\hat{f}(\xi)\chi_{E_{-1}}(\xi), \quad \text{a.e. } \xi \in \mathbb{R}.$$

Since $\hat{f} \in L^q(\mathbb{R})$, it follows from the Lebesgue dominated convergence theorem that $\|[(I-T_a)^n f]\hat{}\, - \hat{f}\chi_{E_0}\|_{L^q(\mathbb{R})} \to 0$ $(n \to \infty)$. □

Remark. It is easy to see that Theorem 3.1 is a special case of this theorem.

Corollary 3.4. *Let $a \in L^1(\mathbb{R})$ and $f \in L^p(\mathbb{R})$ with $1 \le p \le 2$. If $\{(I-T_a)^n f\}_{n\in\mathbb{N}}$ converges in $L^p(\mathbb{R})$, then $\hat{f}(\xi) = 0$ a.e. $\xi \in E_1$.*

Proof. By Hausdorff–Young inequality, $\{[(I - T_a)^n f]^\wedge\}_{n\in\mathbb{N}}$ converges in $L^q(\mathbb{R})$, $p^{-1} + q^{-1} = 1$. The proof is completed. □

3.3. *Convergence of* $\{(I - T_a)^n f\}_{n\in\mathbb{N}}$ *in* $L^s(\mathbb{R})$

It is well-known that the convergence of a sequence of functions in $L^p(\mathbb{R})$ $(1 \le p \le 2)$ implies the convergence of its sequence of Fourier transform in $L^q(\mathbb{R})$, $p^{-1}+q^{-1} = 1$, but it is not true conversely when $p \ne 2$. Therefore, although Theorem 3.3 gives a condition for the convergence of $\{[(I-T_a)^n f]^\wedge\}_{n\in\mathbb{N}}$ in $L^q(\mathbb{R})$, we cannot deduce the convergence of $\{(I - T_a)^n f\}_{n\in\mathbb{N}}$ in $L^p(\mathbb{R})$. However, under certain conditions, we can ensure the convergence of $\{(I - T_a)^n f\}_{n\in\mathbb{N}}$ in $L^s(\mathbb{R})$, $2 \le s \le \infty$. We denote $L^p(\mathbb{R}) + L^2(\mathbb{R}) := \{f + g : f \in L^p(\mathbb{R}), g \in L^2(\mathbb{R})\}$.

Theorem 3.5. *Let* $a \in L^1(\mathbb{R})$ *and* $f \in L^p(\mathbb{R}), 1 \le p \le 2$, *such that* $\hat{f}(\xi) = 0$ a.e. $\xi \in E_1$ *and* $\hat{f} \in L^r(E_{-1})$, $1 \le r \le 2$. *Then there exists an* $f_a \in L^p(\mathbb{R})+L^2(\mathbb{R})$ *such that* $\hat{f}_a = \hat{f}\chi_{E_0}$ *and*

$$\|(I - T_a)^n f - f_a\|_{L^s(\mathbb{R})} \to 0, \quad (n \to \infty),$$

where $s^{-1} + r^{-1} = 1$.

Proof. Since $f \in L^p(\mathbb{R})$, $1 \le p \le 2$, $\hat{f} \in L^q(\mathbb{R})$, where $2 \le q \le \infty$, $p^{-1} + q^{-1} = 1$. Also note that $\hat{f} \in L^r(E_{-1})$, so $\hat{f}\chi_{E_{-1}} \in L^r(\mathbb{R}) \cap L^q(\mathbb{R})$, which yields $\forall n \in \mathbb{N}$, $[1-\hat{a}]^n \hat{f}\chi_{E_{-1}} \in L^r(\mathbb{R})\cap L^q(\mathbb{R}) \subset L^2(\mathbb{R})$. Denote by f_n the inverse Fourier transform of $[1 - \hat{a}]^n \hat{f}\chi_{E_{-1}}$. Then $f_n \in L^2(\mathbb{R}) \cap L^s(\mathbb{R})$, where $s^{-1} + r^{-1} = 1$, and $\hat{f}_n = [1 - \hat{a}]^n \hat{f}\chi_{E_{-1}}$. Through a simple computation, we obtain

$$[(I - T_a)^n f]^\wedge(\xi) - \hat{f}_n(\xi) = \hat{f}(\xi)\chi_{E_0}(\xi), \qquad \text{a.e. } \xi \in \mathbb{R}.$$

Set $f_a := (I - T_a)f - f_1 \in L^p(\mathbb{R}) + L^2(\mathbb{R})$. Then $\hat{f}_a = \hat{f}\chi_{E_0}$. It follows that $[(I - T_a)^n f]^\wedge - \hat{f}_a = \hat{f}_n$, which yields

$$(I - T_a)^n f(x) - f_a(x) = f_n(x) = \frac{1}{2\pi}\hat{\hat{f}}_n(-x), \qquad \text{a.e. } x \in \mathbb{R}.$$

Applying the Hausdorff–Young inequality, we obtain

$$\|(I - T_a)^n f - f_a\|_{L^s(\mathbb{R})} \le \frac{1}{(2\pi)^{1/r}}\|\hat{f}_n\|_{L^r(\mathbb{R})} = \frac{1}{(2\pi)^{1/r}}\|[1 - \hat{a}]^n \hat{f}\|_{L^r(E_{-1})}.$$

By the Lebesgue dominated convergence theorem, there holds $\|(I - T_a)^n f - f_a\|_{L^s(\mathbb{R})} \to 0$ $(n \to \infty)$. □

Corollary 3.6. *Let* $a \in L^1(\mathbb{R})$ *such that* $\hat{a}$ *has compact support. Let* $f \in L^p(\mathbb{R}), 1 \le p \le 2$, *satisfy* $\hat{f}(\xi) = 0$ a.e. $\xi \in E_1$. *Then there exists an* $f_a \in L^p(\mathbb{R}) + L^2(\mathbb{R})$ *such that* $\hat{f}_a = \hat{f}\chi_{E_0}$, *and for all* $2 \le s \le \infty$, *there holds* $\|(I - T_a)^n f - f_a\|_{L^s(\mathbb{R})} \to 0$ $(n \to \infty)$.

Proof. By hypothesis, E_{-1} is a bounded subset of $\mathbb{R}$, thus $\hat{f} \in L^q(E_{-1}) \subset L^r(E_{-1})$, where $p^{-1} + q^{-1} = 1$, $r^{-1} + s^{-1} = 1$. So the condition of Theorem 3.5 is satisfied, and the proof is completed. □

3.4. Convergence rate of $\{(I-T_a)^n f\}_{n\in\mathbb{N}}$

Theorem 3.7. *Let $a \in L^1(\mathbb{R})$ be an even real function such that either $|1-\hat{a}(\xi)| < 1$ or $\hat{a}(\xi) = 0$ for a.e. $\xi \in \mathbb{R}$. Let $f \in L^p(\mathbb{R}), 1 \le p \le 2$. If $\hat{f} \in L^r(E_{-1}),\ 1 \le r \le 2$, then there exists an $f_a \in L^p(\mathbb{R}) + L^2(\mathbb{R})$ such that $\hat{f}_a = \hat{f}\chi_{E_0}$ and*

$$\|(I-T_a)^n f - f_a\|^r_{L^s(\mathbb{R})} \le \frac{1}{2\pi}\Big[(1-\delta)^{rn}\|\hat{f}\|^r_{L^r(E_{-1})} + \int_{\{0<\hat{a}(\xi)<\delta\}} |\hat{f}(\xi)|^r\, d\xi\Big],$$

where $s^{-1}+r^{-1}=1,\ 0<\delta<1$ satisfying $\delta \le 2 - \sup_{\xi\in\mathbb{R}}|\hat{a}(\xi)|$.

Proof. Using the hypothesis it is easy to prove that $\hat{a}$ is a continuous function on $\mathbb{R}$ of real valued, and $0 \le \hat{a}(\xi) < 2$ holds for all $\xi \in \mathbb{R}$. Since $\lim_{|\xi|\to\infty}\hat{a}(\xi) = 0$, there exists $0<\delta<1$, such that $0 \le \hat{a}(\xi) \le 2-\delta$ ($\forall \xi \in \mathbb{R}$), so

$$\begin{aligned}\|(I-T_a)^n f - f_a\|^r_{L^s(\mathbb{R})} &\le \frac{1}{2\pi}\int_{E_{-1}} |[1-\hat{a}(\xi)]^n \hat{f}(\xi)|^r\, d\xi \\ &= \frac{1}{2\pi}\int_{\{\delta\le\hat{a}(\xi)\le 2-\delta\}} |[1-\hat{a}(\xi)]^n|^r \cdot |\hat{f}(\xi)|^r\, d\xi \\ &\quad + \frac{1}{2\pi}\int_{\{0<\hat{a}(\xi)<\delta\}} |[1-\hat{a}(\xi)]^n|^r \cdot |\hat{f}(\xi)|^r\, d\xi \\ &\le \frac{1}{2\pi}\Big[(1-\delta)^{rn}\|\hat{f}\|^r_{L^r(E_{-1})} + \int_{\{0<\hat{a}(\xi)<\delta\}} |\hat{f}(\xi)|^r\, d\xi\Big]. \quad \square\end{aligned}$$

This theorem shows that the less information f possesses in the frequency domain $\{\xi \in \mathbb{R} : 0 < \hat{a}(\xi) < \delta\}$, the faster convergence rate the iteration has. If the Lebesgue measure of the set $\{\xi \in \mathbb{R} : 0 < \hat{a}(\xi) < \delta\}$ converges to 0 as $\delta \to 0$, then for any given error $\epsilon > 0$, as long as $\delta > 0$ is sufficiently small, there holds $\int_{\{0<\hat{a}(\xi)<\delta\}} |\hat{f}(\xi)|^r\, d\xi < \epsilon^r/2$. The error of the other part $(1-\delta)^{rn}\|\hat{f}\|^r_{L^r(E_{-1})}$ decays exponentially. When $n > \ln(\epsilon/2^{1/r}\|\hat{f}\|_{L^r(E_{-1})})/\ln(1-\delta)$, we have $(1-\delta)^{rn}\|\hat{f}\|^r_{L^r(E_{-1})} < \epsilon^r/2$. Combining the inequalities obtained above, we conclude that to yield $\|(I-T)^n f - f_a\|_{L^s(\mathbb{R})} < \epsilon$.

Acknowledgment

This work is supported in part by NSFC (Nos. 10631080, 60873088, 60475042) and DMS-0813750.

References

1. Q. Chen, N. E. Huang, S. Riemenschneider and Y. Xu, A B-spline approach for empirical mode decompositions, *Adv. Comput. Math.* **24** (2006) 171–195.
2. G. B. Folland, *Real Analysis* (John Wiley and Sons Inc., 1984).
3. N. E. Huang, Z. Shen and S. R. Long, A new view of nonlinear water waves: The Hilbert spectrum, *Annu. Rev. Fluid Mech.* **31** (1999) 417–457.

4. N. E. Huang, Z. Shen, S. R. Long *et al.*, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. Roy. Soc. A Lond.* **454**(1971) (1998) 903–995.
5. L. Lin, Y. Wang and H. Zhou, Iterative filtering as an alternative algorithm for empirical mode decomposition, *Adv. Adapt. Data Anal.* **1**(4) (2009).
6. M. E. Montesinos, J. L. Munoz-Cobo and C. Perez, Hilbert–Huang analysis of BWR neutron detector signals: Application to DR calculation and to corrupted signal analysis, *Ann. Nuclear Energy* **30** (2003) 715–727.
7. S. C. Phillips, R. J. Gledhill, J. W. Essex and C. M. Edge, Application of the Hilbert–Huang transform to the analysis of molecular dynamics simulations, *J. Phys. Chem. A* **107** (2003) 4869–4876.
8. E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces* (Princeton University Press, 1971).
9. Y. Wang and Z. Zhou, Toeplitz operators in $l^\infty(\mathbb{Z})$ and their applications to empirical mode decompositions, preprint.