ON THE CONVERGENCE OF ITERATIVE FILTERING EMPIRICAL MODE DECOMPOSITION

YANG WANG AND ZHENGFANG ZHOU

Abstract. Empirical Mode Decomposition (EMD), an adaptive technique for data and signal decomposition, is a valuable tool for many applications in data and signal processing. One approach to EMD is the iterative filtering EMD, which iterates certain banded Toeplitz operators in $l^\infty(\mathbb{Z})$. The convergence of iterative filtering is a challenging mathematical problem. In this paper we study this problem, namely for a banded Toeplitz operator T and $\mathbf{x} \in l^\infty(\mathbb{Z})$ we study the convergence of $T^n(\mathbf{x})$. We also study some related spectral properties of these operators. Even though these operators don't have any eigenvalue in Hilbert space $l^2(\mathbb{Z})$, all eigenvalues and their associated eigenvectors are identified in $l^\infty(\mathbb{Z})$ by using the Fourier transform on tempered distributions. The convergence of $T^n(\mathbf{x})$ relies on a careful localization of the generating function for T around their maximal points and detailed estimates on the contribution from the tails of $\mathbf{x}$.

1. Introduction

Let $\mathbf{a} = (a_k) \in l^1(\mathbb{Z})$. We consider the operator $T_\mathbf{a} : l^\infty(\mathbb{Z}) \longrightarrow l^\infty(\mathbb{Z})$ associated with $\mathbf{a}$, given by

$$T_\mathbf{a}(\mathbf{x}) = \Big(\sum_{j\in\mathbb{Z}} a_j x_{k+j}\Big)_{k\in\mathbb{Z}}$$

where $\mathbf{x} = (x_k) \in l^\infty(\mathbb{Z})$. In the signal processing literature $T_\mathbf{a}$ is called a *filter*, and it is a *finite impulse response (FIR)* filter if $a_k \neq 0$ for only finitely many $k \in \mathbb{Z}$. Note that $T_\mathbf{a}$ is in fact a Toeplitz operator and a FIR filter simply means the Toeplitz operator $T_\mathbf{a}$ is *banded*. In this paper we shall use the terms filter and Toeplitz operator interchangeably, and only FIR filters and banded Toeplitz operators will be considered. Toeplitz operators are classical operators that have been studied extensively, see [3] and the references therein. There is an even larger literature on filters, which we shall not divulge into. In this paper our main focus is on the iteration of certain type of banded Toeplitz operators. More precisely, we consider the following question: Let $T_\mathbf{a}$ be banded and $\mathbf{x} \in l^\infty(\mathbb{Z})$. When will

Key words and phrases. Toeplitz operator, filter, spectrum, empirical mode decomposition (EMD), iterative filtering EMD, intrinsic mode function, Hilbert-Huang transform, sifting algorithm, Poisson summation formula.

The first author is supported in part by NSF grant DMS-0813750.

$T^n_{\mathbf{a}}(\mathbf{x})$ converge (in the sense that every entry converges) as $n\to\infty$? This question arises from signal and data processing using *empirical mode decomposition (EMD)*, which is an important tool for analyzing digital signals and data sets [10, 14]. Our study is motivated primarily by the desire to provide a mathematical framework for EMD.

Signal and data analysis is an important and necessary part in both research and practical applications. Understanding large data set is particularly important and challenging given the explosion of data and numerous ways they are being collected today. Often the challenge is to find hidden information and structures in data and signals. To do so one might encounter several difficulties with the data: The data represent a nonlinear process and is non-stationary; the essential information in the data is often mingled together with noise or other irrelevant information, and others. Historically, Fourier spectral analysis has provided a general method for analyzing signals and data. The term “spectrum” is synonymous with the Fourier transform of the data. Another popular technique is wavelet transform. These techniques are often effective, but are known to have their limitations. To begin with, none of these techniques is data adaptive. This can be a disadvantage in some applications. There are other limitations. For example, Fourier transform may not work well for non-stationary data or data from nonlinear systems. It also does not offer spatial and temporal localization to be useful for some applications in signal processing. Wavelet transform captures discontinuities very successfully. But it too has many limitations; see [10] for a more detailed discussion. These limitations have led Huang et al [10] to propose the *empirical mode decomposition (EMD)* as an highly adaptive technique for analyzing data. EMD has turned out to be a powerful complementary tool to Fourier and wavelet transforms. The goal of EMD is to decompose a signal into a finite number of *instrinsic mode functions (IMF)*, from which hidden information and structures can often be captured by analyzing their Hilbert transformations. We shall not discuss the details of IMF and EMD in this paper. They can be found in [4, 7, 10, 11, 14, 15, 18] and the references therein.

The original EMD is obtained through an algorithm called the *sifting algorithm*. The local maxima and minima of a function (signal) are respectively connected via cubic splines to form the so-called upper and lower envelopes. The average of the two envelopes is then subtracted from the original data. This process is iterated to obtain the first IMF in the EMD. The other IMF's are obtained by the same process on the residual signal. The sifting algorithm is highly adaptive. A small perturbation, however, can alter the envelopes dramatically, raising some questions about its stability. Another drawback there is no natural way to generalize EMD to higher dimensions, which severely limits the scope of its applications. As powerful as EMD is in many applications, a mathematical foundation is virtually nonexistent. Many fundamental mathematical issues such as the convergence of the sifting algorithm have never been established.

To address these concerns, a new approach, the *iterative filtering EMD*, is proposed in [14]. Instead of the average of the upper and lower envelopes, the iterative filtering EMD replaces them by certain FIR filters, usually low-pass filters that yield a "moving average" similar to the mean of the envelopes in the original sifting algorithm. It is shown in [14] that iterative filtering approach often leads to comparable EMD as the classical EMD, and in general it serves as a useful alternative or complement. Furthermore iterative filtering EMD has some advantages over the classic EMD, making it well suited for certain applications [Mao 2010, 17, 19].

The iterative filtering EMD proposed in [14] has the following set up: let $\mathbf{a} = (a_k)_{k\in\mathbb{Z}}$ be finitely supported, i.e. only finitely many $a_k \neq 0$, which we choose so that $T_{\mathbf{a}}(\mathbf{x})$ represents a "moving average" of $\mathbf{x}$. Now let

$$\mathcal{L}(\mathbf{x}) = \mathbf{x} - T_{\mathbf{a}}(\mathbf{x}). \tag{1.1}$$

The first IMF in the EMD is given by $\mathbf{I}_1 = \lim_{n\to\infty} \mathcal{L}^n(\mathbf{x})$, and subsequent IMF's are obtained recursively via $\mathbf{I}_k = \lim_{n\to\infty} \mathcal{L}_k^n(\mathbf{x} - \mathbf{I}_1 - \cdots - \mathbf{I}_{k-1})$. In practical applications the process stops when some stopping criterion is met. For a periodic $\mathbf{x}$ the convergence of $\mathcal{L}^n(\mathbf{x})$ is completely characterized in [14]. However, the convergence for $\mathbf{x} \in l^\infty(\mathbb{Z})$ in general is a much more difficult problem. The main purpose of this paper is to study this question.

The rest of the paper is organized as follows: In Section 2 we introduced the notations and state the main theorem. In Section 3 we prove a result on sum of Dirac measures, which is closely related to the Poisson Summation Formula as well as a classical result of Cordoba [6]. We use it to characterize all eigenvectors of banded Toeplitz operators on $l^p(\mathbb{Z})$ for $1 \leq p \leq \infty$. The proof of the main theorem, which is quite tedious, is given in Section 4.

2. Main Result and Notations

For any $\mathbf{x} = (x_k)_{k\in\mathbb{Z}} \in l^\infty(\mathbb{Z})$ we shall use $\mathbf{x}_N$ to denote the cutoff of $\mathbf{x}$ from $k = -N$ to N, i.e. $\mathbf{x}_N = (y_k)$ such that $y_k = x_k$ for $-N \le x \le N$ and $y_k = 0$ otherwise. We shall often also view $\mathbf{x} = (x_k)_{k\in\mathbb{Z}} \in l^\infty(\mathbb{Z})$ as a function $\mathbf{x} : \mathbb{Z} \longrightarrow \mathbb{C}$ with $\mathbf{x}(k) = x_k$. We say $\mathbf{x} = (x_k) \in l^\infty(\mathbb{Z})$ is *symmetric* if $x_k = x_{-k}$ for all $k \in \mathbb{Z}$, and it is *finitely supported* if $\mathrm{supp}(\mathbf{x}) := \{k \in \mathbb{Z} : x_k \ne 0\}$ is a finite set.

Throughout this paper the Fourier transform of a function $f(x)$ is defined as

$$\mathcal{F}(f)(\xi) = \widehat{f}(\xi) := \int_{\mathbb{R}} f(x)e^{2\pi i x\xi}\, dx.$$

The inverse Fourier transform of $g(\xi)$ is

$$\mathcal{F}^{-1}(g)(x) := \int_{\mathbb{R}} g(\xi)e^{-2\pi i\xi x}\, d\xi.$$

For each $\mathbf{x} \in l^\infty(\mathbb{Z})$ there is an associated complex measure $\mu_{\mathbf{x}} := \sum_{k\in\mathbb{Z}} x_k\delta_k$, where δ_b is the Dirac measure supported at b for any $b \in \mathbb{R}$, i.e. $\delta_b(x) = \delta(x - b)$. It is well known that $\mu_{\mathbf{x}}$ is a tempered distribution. Thus $\widehat{\mu_{\mathbf{x}}}$ is also well defined as a tempered distribution. We shall often use $\widehat{\mathbf{x}}$ to denote $\widehat{\mu_{\mathbf{x}}}$ for simplicity, especially when $\mathbf{x}$ is finitely supported; in such case $\widehat{\mathbf{x}}(\xi)$ is a trigonometric polynomial.

Going back to Toeplitz operators, it is easy to check that for any $\mathbf{a} \in l^1(\mathbb{Z})$ we have

$$\widehat{T_{\mathbf{a}}(\mathbf{x})}(\xi) = \widehat{\mathbf{a}}(-\xi)\widehat{\mu_{\mathbf{x}}}(\xi) = \widehat{\mathbf{a}}(-\xi)\widehat{\mathbf{x}}(\xi).$$

For any finitely supported $\mathbf{a}$ the spectrum of $T_{\mathbf{a}}$ is precisely $\{\widehat{\mathbf{a}}(\xi) : \xi \in [0, 1)\}$. Let $Z_{\mathbf{a},\lambda} = \{\theta \in [0, 1) : \widehat{\mathbf{a}}(\theta) = \lambda\}$. This set will occur very frequently in this paper.

Before stating our main theorem we introduce a few more notations. For any $\theta \in \mathbb{R}$ let $\mathbf{v}_\theta := (e^{2\pi i k\theta})_{k\in\mathbb{Z}}$. If $\theta \in Z_{\mathbf{a},\lambda}$ then $T_{\mathbf{a}}(\mathbf{v}_\theta) = \lambda\mathbf{v}_\theta$. For any $\mathbf{x} = (x_k)$ and $\mathbf{y} = (y_k)$ in $l^\infty(\mathbb{Z})$ define

$$[\mathbf{x}, \mathbf{y}] = \lim_{n\to\infty} \frac{1}{2n+1} \sum_{k=-n}^{n} x_k\bar{y}_k$$

if it exists. One can view this as a form of "inner product."

One of the main objectives of the paper is to study the convergence of the new sifting algorithm from which we obtain the IMFs by $\mathbf{I}_k = \lim_{n\to\infty}(I - T_{\mathbf{a}_k})^n(\mathbf{x} - \mathbf{I}_1 - \cdots - \mathbf{I}_{k-1})$. Since $I - T_{\mathbf{a}}$ is simply the Toeplitz operator $T_{\boldsymbol{\delta}-\mathbf{a}}$ where $\boldsymbol{\delta} = (\delta_{k0})$ with $\delta_{00} = 1$ and $\delta_{k0} = 0$

for $k \neq 0$. So we shall focus on iterations of $T_{\mathbf{a}}$ for general finitely supported $\mathbf{a}$. Our main theorem of the paper is:

Theorem 2.1. *Let $\mathbf{a} = (a_k)$ be finitely supported and symmetric such that $-1 < \widehat{\mathbf{a}}(\xi) \leq 1$ and $\widehat{\mathbf{a}}(\xi) \not\equiv 1$. For any $\mathbf{x} \in l^\infty(\mathbb{Z})$, if $[\mathbf{x}, \mathbf{v}_\theta]$ exists for all $\theta \in Z_{\mathbf{a},1}$ then*

$$\lim_{n\to\infty} T_{\mathbf{a}}^n(\mathbf{x}) = \sum_{\theta \in Z_{\mathbf{a},1}} [\mathbf{x}, \mathbf{v}_\theta]\mathbf{v}_\theta \qquad \textit{pointwise.} \tag{2.1}$$

Here pointwise convergence means the k-th entry $T_{\mathbf{a}}^n(\mathbf{x})(k)$ of $T_{\mathbf{a}}^n(\mathbf{x})$ converges for each $k \in \mathbb{Z}$. Informally speaking, $T_{\mathbf{a}}^n(\mathbf{x})$ converges pointwise to the "projection" of $\mathbf{x}$ onto the 1-eigenspace of $T_{\mathbf{a}}$. Note that the eigenvalues of $T_{\mathbf{a}}$ are precisely $\{\widehat{\mathbf{a}}(\xi) : \xi \in [0, 1)\}$ (see Section 3), so the condition $-1 < \widehat{\mathbf{a}}(\xi) \leq 1$ is natural. It is not clear whether the condition $[\mathbf{x}, \mathbf{v}_\theta]$ exists for each $\theta \in Z_{\mathbf{a},1}$ is a necessary condition. The following example shows that $\lim_{n\to\infty} T_{\mathbf{a}}^n(\mathbf{x})$ does not exist for a $x \in l^\infty(\mathbb{Z})$.

Example 2.1. Let $\mathbf{a} = (a_k)$ with $a_0 = \frac{1}{2}$, $a_1 = a_{-1} = \frac{1}{4}$ and $a_k = 0$ for all other k. $\widehat{\mathbf{a}}(\xi) = \sin^2 \frac{\xi}{2}$ satisfies the hypothesis of Theorem 2.1. Let $\mathbf{x} = (x_k)$ where $x_k = 0$ for all $k \leq 0$ and $x_k = (-1)^{n-1}$ for $2^{n!} \leq K < 2^{(n+1)!}$. Then it is easy to show that $T_{\mathbf{a}}^n(\mathbf{x})(0)$ does not converge. In fact every point in $[-\frac{1}{2}, \frac{1}{2}]$ is a limit point of the sequence.

3. Eigenvectors of Banded Toeplitz Operators

To study the iterations of banded Toeplitz operators it is natural to ask about their eigenvalues and eigenvectors in $l^\infty(\mathbb{Z})$. We state some results here. While these results may not be new (although we have not found them in the literature), our approach appears to be.

For any $\mathbf{x} = (x_k) \in l^\infty(\mathbb{Z})$ the associated measure $\mu_{\mathbf{x}}$ is a tempered distribution [9, 13]. Hence its Fourier transform, given by

$$\langle \widehat{\mu_{\mathbf{x}}}, \phi \rangle := \langle \mu_{\mathbf{x}}, \widehat{\phi} \rangle = \sum_{k \in \mathbb{Z}} x_k \widehat{\phi}(k) \tag{3.1}$$

for any ϕ in the Schwartz class, is also a tempered distribution.

Lemma 3.1. *Let $\mathbf{x} \in l^\infty(\mathbb{Z})$ such that $\operatorname{supp}(\widehat{\mu_{\mathbf{x}}}) = \Lambda$ is a uniformly discrete set in $\mathbb{R}$. Then*

$$\widehat{\mu_{\mathbf{x}}} = \sum_{\beta \in \Lambda} c_\beta \delta_\beta \tag{3.2}$$

for some bounded sequence $(c_\beta)_{\beta\in\Lambda}$ *in* $\mathbb{C}$.

Proof. Since Λ is uniformly discrete we may find $\psi_\beta \in C_0^\infty(\mathbb{R})$ for each $\beta \in \Lambda$ such that $\sum_{\beta\in\Lambda} \psi_\beta = 1$ and $\mathrm{supp}(\psi_\beta) \cap \Lambda = \{\beta\}$. Now for any $\beta \in \Lambda$, $\psi_\beta\, \widehat{\mu_{\mathbf{x}}}$ is a tempered distribution supported on a single point $\{\beta\}$. It follows that

$$\psi_\beta\, \widehat{\mu_{\mathbf{x}}} = \sum_{j=0}^{N} a_j \delta_\beta^{(j)},$$

where $\delta_\beta^{(j)}$ denotes the j-th derivative of δ_β (see e.g. Folland [8]). We only need to show that $a_j = 0$ for $j > 0$ for all $\beta \in \Lambda$. If not there exists some $\alpha^* \in \Lambda$ such that $\psi_{\alpha^*}\, \widehat{\mu_{\mathbf{x}}} = \sum_{j=0}^{N} a_j \delta_{\alpha^*}^{(j)}$ with $a_N \neq 0$, $N > 0$.

Without loss of generality we assume that $\alpha^* = 0$. Now take a test function $\phi \in C_0^\infty(\mathbb{R})$ such that $\mathrm{supp}(\phi) = [-\varepsilon, \varepsilon]$, $\mathrm{supp}(\phi) \cap \mathrm{supp}(\psi_\beta) = \emptyset$ for all $\beta \neq 0$, and $\phi^{(j)}(0) = 0$ for $j < N$ but $\phi^{(N)}(0) \neq 0$. Set $\phi_\lambda(x) = \phi(\lambda x)$. Then

$$\langle \widehat{\mu_{\mathbf{x}}}, \phi_\lambda \rangle = \Big\langle \sum_{\beta\in\Lambda} \psi_\beta\, \widehat{\mu_{\mathbf{x}}}, \phi_\lambda \Big\rangle = \langle \psi_0\, \widehat{\mu_{\mathbf{x}}}, \phi_\lambda \rangle = \Big\langle \sum_{j=0}^{N} a_j \delta_0^{(j)}, \phi_\lambda \Big\rangle = a_N \lambda^N \phi^{(N)}(0),$$

which goes to ∞ as $\lambda \to \infty$. On the other hand,

$$|\langle \widehat{\mu_{\mathbf{x}}}, \phi_\lambda \rangle| \;=\; |\langle \mu_{\mathbf{x}}, \widehat{\phi_\lambda} \rangle| \;=\; \Big|\, \frac{1}{\lambda} \sum_{k\in\mathbb{Z}} x_k \widehat{\phi}\Big(\frac{k}{\lambda}\Big) \,\Big| \;\leq\; \frac{C}{\lambda} \sum_{k\in\mathbb{Z}} \Big|\, \widehat{\phi}\Big(\frac{k}{\lambda}\Big) \,\Big|,$$

which goes to $C \int_{\mathbb{R}} |\widehat{\phi}(\xi)|\, d\xi$ as $\lambda \to \infty$. This is a contradiction. Thus $\widehat{\mu_{\mathbf{x}}} = \sum_{\beta\in\Lambda} c_\beta \delta_\beta$.

It remains to show c_β are bounded. Take a test function $\phi \in C^\infty(\mathbb{R})$ such that $\phi(0) = 1$ and $\mathrm{supp}(\phi) = [-\varepsilon, \varepsilon]$, where $\varepsilon < \inf\{|\alpha_1 - \alpha_2| : \alpha_1, \alpha_2 \in \Lambda, \alpha_1 \neq \alpha_2\}$. Applying $\langle \widehat{\mu_{\mathbf{x}}}, \varphi \rangle = \langle \mu_{\mathbf{x}}, \widehat{\varphi} \rangle$ to $\varphi(x) = \phi(x - \beta)$ for each $\beta \in \Lambda$ we obtain

$$|c_\beta| = |\langle \mu_{\mathbf{x}}, \widehat{\varphi} \rangle| \;=\; \Big|\, \sum_{k\in\mathbb{Z}} x_k e^{-2\pi i k\beta} \widehat{\phi}(k) \,\Big| \;\leq\; \|\mathbf{x}\|_\infty \sum_{k\in\mathbb{Z}} \Big|\, \widehat{\phi}(k) \,\Big|.$$

This proves the lemma. ∎

The following theorem is closely related to a well known result of Cordoba [6], which is a classic result in the study of quasicrystals.

Theorem 3.2. *Let* Λ *be a uniformly discrete set in* $\mathbb{R}$ *and* $\mu = \sum_{\beta\in\Lambda} x_\beta \delta_\beta$ *where* (x_β) *is bounded. Assume that* $\mathrm{supp}(\widehat{\mu}) \subset \mathbb{Z}$. *Then*

(A) *There exist* $\alpha_1, \ldots, \alpha_m \in [0, 1)$ *such that* $\Lambda = \bigcup_{j=1}^{m} (\alpha_j + \mathbb{Z})$.

(B) *There exist* $c_1, \ldots, c_m$ *such that* $x_\beta = c_j$ *for all* $\beta \in \alpha_j + \mathbb{Z}$. *Thus*

$$\mu = \sum_{j=1}^{m} c_j \sum_{k\in\mathbb{Z}} \delta_{\alpha_j + k}.$$

(C) $\widehat{\mu} = \sum_{k\in\mathbb{Z}} \Big(\sum_{j=1}^m c_j e^{2\pi i \alpha_j k} \Big) \delta_k$.

Proof. By Lemma 3.1 we have $\widehat{\mu} = \sum_{k\in\mathbb{Z}} p_k \delta_k$. For $\phi \in C_0^\infty(\mathbb{R})$ denote $\phi_{\lambda,t}(x) := \phi(\lambda x) e^{2\pi i t x}$. Then $\widehat{\phi_{\lambda,t}}(\xi) = \lambda^{-1}\phi\big(\lambda^{-1}(\xi - t)\big)$.

It follows from $\langle \mu, \widehat{\phi_{\lambda,t}} \rangle = \langle \widehat{\mu}, \phi_{\lambda,t} \rangle$ that

$$\sum_{k\in\mathbb{Z}} p_k \phi_{\lambda,t}(k) = \sum_{\alpha\in\Lambda} x_\alpha \widehat{\phi_{\lambda,t}}(\alpha).$$

This yields

$$\sum_{k\in\mathbb{Z}} p_k \phi(\lambda k) e^{2\pi i k t} = \frac{1}{\lambda} \sum_{\alpha\in\Lambda} x_\alpha \widehat{\phi}\Big(\frac{\alpha - t}{\lambda}\Big). \tag{3.3}$$

Substituting $1/\lambda$ for λ we can rewrite the equation as

$$\lambda^{-1} \sum_{k\in\mathbb{Z}} p_k \phi(\lambda^{-1} k) e^{2\pi i k t} = \sum_{\alpha\in\mathbb{R}} x_\alpha \widehat{\phi}\left(\lambda(\alpha - t)\right), \tag{3.4}$$

where $x_\alpha = 0$ for $\alpha \notin \Lambda$. Observe that because all x_α are bounded and Λ is uniformly discrete we have

$$\lim_{\lambda\to\infty} \sum_{\alpha\in\mathbb{R}} x_\alpha \widehat{\phi}\left(\lambda(\alpha - t)\right) = x_t \widehat{\phi}(0).$$

However the right hand side of (3.4) has

$$\lambda^{-1} \sum_{k\in\mathbb{Z}} p_k \phi(\lambda^{-1} k) e^{2\pi i k t_1} = \lambda^{-1} \sum_{k\in\mathbb{Z}} p_k \phi(\lambda^{-1} k) e^{2\pi i k t_2}$$

for any t_1, t_2 with $t_1 - t_2 \in \mathbb{Z}$. By choosing ϕ such that $\widehat{\phi}(0) = \int_{\mathbb{R}} \phi \neq 0$ it follows that $x_{t_1} = x_{t_2}$ whenever $t_1 - t_2 \in \mathbb{Z}$. Thus Λ must be the union of equivalent classes modulo $\mathbb{Z}$, i.e. cosets of $\mathbb{Z}$. Being uniformly discrete Λ can only be a finitely union of cosets of $\mathbb{Z}$. Hence there exist $\alpha_1, \ldots, \alpha_m \in [0, 1)$ such that $\Lambda = \bigcup_{j=1}^m (\alpha_j + \mathbb{Z})$. Furthermore, $x_\beta = c_j$ for all $\beta \in \alpha_j + \mathbb{Z}$. Finally (C) follows directly from taking the Fourier transform of μ and the Poisson Summation Formula

$$\widehat{\sum_{k\in\mathbb{Z}} \delta_{k+\alpha}} = \sum_{k\in\mathbb{Z}} e^{2\pi i \alpha k} \delta_k.$$

∎

Remark 1. The condition $\mathrm{supp}(\widehat{\mu}) \subset \mathbb{Z}$ in the theorm can be replaced with $\mathrm{supp}(\widehat{\mu}) \subset \Gamma$ for some lattice Γ. In this setting the theorem still holds if the set $\mathbb{Z}$ in (A) and (B) is replaced by the dual lattice Γ^* of Γ, and the $\mathbb{Z}$ in (C) is replaced by Γ.

Remark 2. A theorem of Cordoba [6] draws the same conclusions under the hypotheses that $\mathrm{supp}(\widehat{\mu})$ is a uniformly discrete set but requires that the set $\{x_\beta : \ \beta \in \Lambda\}$ is finite.

Theorem 3.3. *Let* $\mathbf{a} = (a_k)$ *be finitely supported. Suppose* $T_{\mathbf{a}} \neq c\, I$ *where* I *is the identity map. Then* λ *is an eigenvalue of* $T_{\mathbf{a}}$ *if and only if* $\lambda \in \{\widehat{\mathbf{a}}(\xi) : \xi \in [0,1)\}$*. Furthermore* $\mathbf{x} \in l^\infty(\mathbb{Z})$ *is an eigenvector of* $T_{\mathbf{a}}$ *for the eigenvalue* λ *if and only if*

$$\mathbf{x} = \sum_{\theta \in Z_{\mathbf{a},\lambda}} c_\theta \mathbf{v}_\theta \tag{3.5}$$

for some constants c_θ*, where* $Z_{\mathbf{a},\lambda} := \{t \in [0,1) : \widehat{\mathbf{a}}(t) = \lambda\}$*.*

Proof. For any $\lambda = \widehat{\mathbf{a}}(t)$ it is easy to check that $\mathbf{v}_t$ is an eigenvector of $T_{\mathbf{a}}$. Let λ be an eigenvalue of $T_{\mathbf{a}}$ with $T_{\mathbf{a}}(\mathbf{x}) = \lambda \mathbf{x}$ for some nonzero $\mathbf{x} \in l^\infty(\mathbb{Z})$. Observe that $\mathcal{F}^{-1}(\mu_{T_{\mathbf{a}}(\mathbf{x})}) = \widehat{\mathbf{a}}\mathcal{F}^{-1}(\mu_{\mathbf{x}})$. Thus

$$\widehat{\mathbf{a}}\mathcal{F}^{-1}(\mu_{\mathbf{x}}) = \lambda \mathcal{F}^{-1}(\mu_{\mathbf{x}}), \qquad \text{and} \qquad (\widehat{\mathbf{a}} - \lambda)\mathcal{F}^{-1}(\mu_{\mathbf{x}}) = 0.$$

It follows that $\mathrm{supp}(\mathcal{F}^{-1}(\mu_{\mathbf{x}})) \subseteq Z_{\mathbf{a},\lambda} + \mathbb{Z}$. Thus $\lambda \in \{\widehat{\mathbf{a}}(\xi) : \xi \in [0,1)\}$, and because $T_{\mathbf{a}} \neq c\, I$ the set $Z_{\mathbf{a},\lambda}$ is finite. Hence $Z_{\mathbf{a},\lambda} + \mathbb{Z}$ is uniformly discrete. Lemma 3.1 implies that

$$\mathcal{F}^{-1}(\mu_{\mathbf{x}}) = \sum_{\alpha \in Z_{\mathbf{a},\lambda} + \mathbb{Z}} b_\alpha \delta_\alpha$$

for some bounded sequence(b_α). Theorem 3.2 now applies to $\mathcal{F}^{-1}(\mu_{\mathbf{x}})$ to show that

$$\mathcal{F}^{-1}(\mu_{\mathbf{x}}) = \sum_{\theta \in Z_{\mathbf{a},\lambda}} c_\theta \sum_{k \in \mathbb{Z}} \delta_{\theta + k}.$$

The structure of $\mu_{\mathbf{x}}$ now follows from part (C) of Theorem 3.2, which yields

$$\mathbf{x} = \sum_{\theta \in Z_{\mathbf{a},\lambda}} c_\theta \mathbf{v}_\theta.$$

∎

Corollary 3.4. *For any finitely supported* $\mathbf{a} = (a_k)$ *the operator* $T_{\mathbf{a}}$ *has no point spectrum in* $l^p(\mathbb{Z})$ *for any* $1 \le p < \infty$ *unless* $T_{\mathbf{a}} = c\, I$*.*

Proof. Clearly any eigenvector for $T_{\mathbf{a}}$ in $l^p(\mathbb{Z})$ is also an eigenvector in $l^\infty(\mathbb{Z})$ for the same eigenvalue. If $T_{\mathbf{a}} \neq c\,I$ then by Theorem 3.3, all eigenvectors of $T_{\mathbf{a}}$ in $l^\infty(\mathbb{Z})$ are of the form (3.5), which do not belong to $l^p(\mathbb{Z})$ for an $1 \leq p < \infty$. This is easily seen from the fact that such $\mathbf{x}$ are almost periodic so the entries do not tend to 0. Thus $T_{\mathbf{a}}$ has no point spectrum in $l^p(\mathbb{Z})$ for $1 \leq p < \infty$. ∎

4. Proof of Main Theorem

In this section we assume the hypotheses of Theorem 2.1 and prove the theorm by breaking it down into a series of lemmas and estimates. Without loss of generality we assume that $\mathbf{a} = (a_k)$ is symmetric and $a_k = 0$ for $k > q$ or $k < -q$, i.e. $\mathrm{supp}(\mathbf{a}) \subset [-q, q]$. To prove the theorem it suffices to prove that $\lim_{n\to\infty} T_{\mathbf{a}}^n(\mathbf{x})(0) = \sum_{\theta \in Z_{\mathbf{a},1}} [\mathbf{x}, \mathbf{v}_\theta]$.

Lemma 4.1. *Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Then*

$$\lim_{n\to\infty} T_{\mathbf{a}}^n(\mathbf{x})(0) = \int_{\mathbb{T}} \widehat{\mathbf{a}}^n(\xi)\widehat{\mathbf{x}_{qn}}(\xi)\,d\xi \tag{4.1}$$

and

$$\lim_{n\to\infty} \Big(T_{\mathbf{a}}^n(\mathbf{x})(0) - \sum_{\theta\in Z_{\mathbf{a},1}} \int_{|\xi-\theta|<\delta} \widehat{\mathbf{a}}^n(\xi)\widehat{\mathbf{x}_{qn}}(\xi)\,d\xi\Big) = 0 \tag{4.2}$$

for any $\delta > 0$ such that the intervals $\{(\theta-\delta, \theta+\delta) : \theta \in Z_{\mathbf{a},1}\}$ in $\mathbb{T}$ are disjoint.

Proof. Note that $\widehat{\mathbf{a}}^n$ is a trigonometric polynomial of degree qn, $T_{\mathbf{a}}^n(\mathbf{x})(0)$ is the constant term of $\widehat{\mathbf{a}}^n(-\xi)\widehat{\mathbf{x}}_{qn}(\xi)$. Integrating it over $\mathbb{T}$ yields $T_{\mathbf{a}}^n(\mathbf{x})(0)$. (4.1) follows from the fact that $\widehat{\mathbf{a}}^n(-\xi) = \widehat{\mathbf{a}}^n(\xi)$.

To prove (4.2) we observe that

$$(T_{\mathbf{a}}^n(\mathbf{x})(0) - \sum_{\theta\in Z_{\mathbf{a},1}} \int_{|\xi-\theta|<\delta} \widehat{\mathbf{a}}^n(\xi)\widehat{\mathbf{x}_{qn}}(\xi)\,d\xi = \int_E \widehat{\mathbf{a}}^n(\xi)\widehat{\mathbf{x}_{qn}}(\xi)\,d\xi,$$

where $|\xi - \theta| \geq \delta$ on E for any $\theta \in Z_{\mathbf{a},1}$. Thus there exists an $\varepsilon > 0$ such that $|\widehat{\mathbf{a}}(\xi)| \leq 1 - \varepsilon$ on E. Also $|\widehat{\mathbf{x}}_{qn}(\xi)| \leq \|\mathbf{x}\|_\infty qn$, so

$$\lim_{n\to\infty}\Big|\int_E \widehat{\mathbf{a}}^n(\xi)\widehat{\mathbf{x}_{qn}}(\xi)\,d\xi\Big| \leq \lim_{n\to\infty} (1-\varepsilon)^n \|\mathbf{x}\|_\infty qn = 0. \tag{4.3}$$

∎

Throughout this section we shall assume that $\delta > 0$ is small enough so that $\{(\theta-\delta, \theta+\delta) : \theta \in Z_{\mathbf{a},1}\}$ in $\mathbb{T}$ are disjoint. Our next step shows that with small enough $\delta > 0$, for any

$\varepsilon > 0$ the estimate

$$\Big|\int_{|\xi-\theta|<\delta} \widehat{\mathbf{a}}^n(\xi)\widehat{\mathbf{x}_{qn}}(\xi)\,d\xi - [\mathbf{x},\mathbf{v}_\theta]\Big| < \varepsilon \tag{4.4}$$

holds for sufficiently large n. This is achieved by performing a series of delicate estimates. Obviously Theorem 2.1 follows readily from (4.4).

We now fix any $\theta \in Z_{\mathbf{a},1}$. Note that $\widehat{\mathbf{a}} \leq 1$ so $\widehat{\mathbf{a}}'(\theta) = 0$ and

$$\widehat{\mathbf{a}}(\xi) = 1 - c_\theta(\xi-\theta)^{2m} + O((\xi-\theta)^{2m+1})$$

near θ, where $c_\theta > 0$. It follows that $\widehat{\mathbf{a}}(\theta+t) = 1 - c_\theta t^{2m} + h_\theta(t)$ where $h_\theta(t) = O(t^{2m+1})$ is bounded and

$$\int_{|\xi-\theta|<\delta} \widehat{\mathbf{a}}^n(\xi)\widehat{\mathbf{x}_{qn}}(\xi)\,d\xi = \int_{-\delta}^{\delta} \widehat{\mathbf{a}}^n(\theta+t)\widehat{\mathbf{x}_{qn}}(\theta+t)\,dt = A(n,\theta,\delta) + B(n,\theta,\delta),$$

where

$$A(n,\delta,\theta) = \int_{-\delta}^{\delta} (1-c_\theta t^{2m})^n \widehat{\mathbf{x}_{qn}}(\theta+t)\,dt, \tag{4.5}$$

$$B(n,\delta,\theta) = \int_{-\delta}^{\delta} \big(\widehat{\mathbf{a}}(\theta+t)^n - (1-c_\theta t^{2m})^n\big)\,\widehat{\mathbf{x}_{qn}}(\theta+t)\,dt. \tag{4.6}$$

We first prove that $\lim_{n\to\infty} B(n,\delta,\theta) = 0$. To do so we study the term $\widehat{\mathbf{a}}(\theta+t)^n - (1 - c_\theta t^{2m})^n$ on $[0,\delta]$.

Lemma 4.2. *Let $F_n(t) = (1-t^k+h(t))^n - (1-t^k)^n$ where $k \geq 2$ and $h(t) = o(t^k)$ is analytic and nonzero. Let $\delta > 0$ be sufficiently small. Then for sufficiently large n the function F_n has only one critical point $t_n \in (0,\delta]$, at which $|F(t_n)| = \max_{t\in[0,\delta]} |F_n(t)| \leq Cn^{-\frac{1}{k}}$.*

Proof. Let $f(t) = 1 - t^k + h(t)$ and $g(t) = 1 - t^k$. Since h is analytic and nonzero, we have $h(t) = Kt^m + O(t^{m+1})$ where $m > k$ and $K \neq 0$. Note that $F_n(0) = 0$ and $F_n(\delta) \to 0$ exponentially. But $F_n(n^{-\frac{1}{k}})$ does not go to 0 exponentially. Hence F_n has at least one critical point in $(0,\delta]$ for sufficiently large n. Let t_n be a critical point, i.e. $F_n'(t_n) = 0$. It follows that $f^{n-1}(t_n)f'(t_n) - g^{n-1}(t_n)g'(t_n) = 0$, and thus

$$\frac{g^{n-1}(t_n)}{f^{n-1}(t_n)} = \frac{f'(t_n)}{g'(t_n)} = 1 - \frac{Km}{k} t_n^{m-k} + O(t_n^{m-k+1}). \tag{4.7}$$

Claim. $\lim_{n\to\infty}(n-1)t_n^k = m/k$.

It is clear that if $t_{n_j}\to\varepsilon$ where $\varepsilon > 0$ then $\frac{g^{n_j-1}(t_{n_j})}{f^{n_j-1}(t_{n_j})}\to 0$ when $K > 0$, the limit is $+\infty$ when $K < 0$. This is a contradiction. Hence $t_n\to 0$. Now observe that $g(t_n)/f(t_n) = 1 - Kt_n^m + O(t_n^{m+1})$. By taking logarithm on both sides of (4.7) we obtain

$$-(n-1)Kt_n^m + O((n-1)t_n^{m+1}) = -\frac{Km}{k}t_n^{m-k} + O(t_n^{m-k+1}).$$

The claim $\lim_{n\to\infty}(n-1)t_n^k = m/k$ now follows. □

We next show that this t_n is unique for sufficiently large n. This is done by the sign of $F''(t_n)$.

$$F_n'' = n\left(f''f^{n-1} - g''g^{n-1}\right) + n(n-1)\left(f'^2 f^{n-2} - g'^2 g^{n-2}\right).$$

Thus

$$\frac{F_n''}{nf^{n-2}} = \Big(f''f - g''\frac{g^{n-1}}{f^{n-2}}\Big) + (n-1)\Big(f'^2 - g'^2\frac{g^{n-2}}{f^{n-2}}\Big).$$

At $t = t_n$ we have $\frac{g^{n-1}}{f^{n-1}} = \frac{f'}{g'}$. It is not hard to verify that this yields

$$\frac{F_n''}{nf^{n-2}} = fg'\Big(\frac{f'}{g'}\Big)' + (n-1)f'g\Big(\frac{f}{g}\Big)' \qquad \text{at } t = t_n.$$

One can also check easily that at $t = t_n$,

$$fg'(f'/g')' = Km(m-k)t_n^{m-2} + O(t_n^{m-1}),$$

$$f'g(f/g)' = -Km(m-k)t_n^{m+k-2} + O(t_n^{m+k-1}).$$

Thus by the Claim,

$$\lim_{n\to\infty}\frac{F_n''(t_n)}{n\,t_n^{l-2}g^{n-2}(t_n)} = Km(m-k) - \lim_{n\to\infty}Kmk(n-1)t_n^k = -Kmk.$$

If $K > 0$, we have $F_n(t) \geq 0$ on $[0,\delta]$ and $F_n''(t_n) < 0$, which implies that any critical point t_n is a local maximum for sufficiently large n. But any two local maximum must sandwich a local minimum. Thus there can only be one critical point, at which F_n must achieve its maximum. If $K < 0$, $F_n(t) \leq 0$ on $[0,\delta]$ and $F_n''(t_n) > 0$. By the same reason t_n is the only critical point of F_n and it is the minimum. Thus in either case we must have $|F(t_n)| = \max_{t\in[0,\delta]}|F_n(t)|$. Finally,

$$|F_n(t)| \leq (1-t^k)^n\left((1+K_1t^m)^n - 1\right)$$

for some $K_1 > 0$ on $[0,\delta]$. It follows from $\lim_{n\to\infty}nt_n^k = m/k$ that $(1-t_n^m)^n\to e^{\frac{m}{k}}$ and $(1+K_1t_n^m)^n - 1 = O(t_n^{m-k}) = O(n^{-\frac{m-k}{k}})$. This proves the lemma. ■

Lemma 4.3. $\lim_{n\to\infty}B(n,\delta,\theta) = 0$.

Proof. Let $F_n(t) = \widehat{\mathbf{a}}(\theta+t)^n - (1-c_\theta t^{2m})^n$. Without loss of generality we may assume that $c_\theta = 1$. Then F_n satisfies the hypothesis of Lemma 4.2. Let t_n be the unique critical point of F_n on $(0,\delta]$. So F_n is monotone on $[0,t_n]$ and $[t_n,\delta]$. A special Mean Value Theorem for integration (see e.g. Bartle [1], Theorem 30.11) now implies that for some $\eta \in [0,\delta]$,

$$\int_0^{t_n} F_n(t)\widehat{\mathbf{x}_{qn}}(\theta+t)\,dt = F_n(0)\int_0^{\eta} \widehat{\mathbf{x}_{qn}}(\theta+t)\,dt + F_n(t_n)\int_\eta^{t_n} \widehat{\mathbf{x}_{qn}}(\theta+t)\,dt.$$

Note that $F_n(0) = 0$ and

$$\Big|\int_\eta^{\delta} \widehat{\mathbf{x}_{qn}}(\theta+t)\,dt\Big| \le C_1\sum_{j=1}^{qn}\frac{1}{j} = O(\ln n).$$

It follows that

$$\Big|\int_0^{t_n} F_n(t)\widehat{\mathbf{x}_{qn}}(\theta+t)\,dt\Big| = O(n^{-\frac{1}{2m}}\ln n).$$

By the same token,

$$\Big|\int_{t_n}^{\delta} F_n(t)\widehat{\mathbf{x}_{qn}}(\theta+t)\,dt\Big| = O(n^{-\frac{1}{2m}}\ln n).$$

Thus $\int_0^\delta F_n(t)\widehat{\mathbf{x}_{qn}}(\theta+t)\,dt \longrightarrow 0$. Similarly $\int_{-\delta}^0 F_n(t)\widehat{\mathbf{x}_{qn}}(\theta+t)\,dt \longrightarrow 0$. These combine to yield $\lim_{n\to\infty} B(n,\delta,\theta) = 0$. ∎

Lemma 4.4. *Assume that $b,k>0$ and $p\ge 0$. For any $\epsilon>0$ such that $b\epsilon^k<1$ we have*

$$\int_0^{\epsilon}(1-bt^k)^n t^p\,dt \le \min\Big\{\frac{1}{p+1}\epsilon^{p+1}, C\,n^{-\frac{p+1}{k}}\Big\},$$

where $C = \int_0^\infty e^{-bs^k}s^p\,ds$.

Proof. Using the fact that $1-x\le e^{-x}$ for all $x\ge 0$, we have $(1-bt^k)^n \le e^{-nbt^k}$. Make the substitution $s = \sqrt[k]{n}t$ we have

$$\int_0^{\epsilon}(1-bt^k)^n t^p\,dt \le n^{-\frac{p+1}{k}}\int_0^{\sqrt[k]{n}\epsilon} e^{-bs^k}s^p\,ds.$$

The lemma follows from two estimates. First, the integral $\int_0^{\sqrt[k]{n}\epsilon} e^{-bs^k}s^p\,ds$ is bounded by $C_2 = \int_0^\infty e^{-bs^k}s^p\,ds$. Second, it is also bounded by $\frac{1}{p+1}(\sqrt[k]{n}\epsilon)^{p+1}$. ∎

Next we concentrate on estimating $A(n,\delta,\theta)$. To achieve this, for each $\varepsilon>0$ we break $\widehat{\mathbf{x}_{qn}}(\theta+t)$ up into three parts:

$$\widehat{\mathbf{x}_{qn}}(\theta+t) = \Big(\sum_{|k|\le\varepsilon n^\sigma} + \sum_{\varepsilon n^\sigma<|k|\le\varepsilon^{-1}n^\sigma} + \sum_{\varepsilon^{-1}n^\sigma<|k|\le qn}\Big)x_k e^{2\pi ik(\theta+t)} = J_1+J_2+J_3,$$

where $\sigma = \frac{1}{2m}$. As a result we write $A(n,\delta,\theta) = A_1(n,\delta,\theta,\varepsilon) + A_2(n,\delta,\theta,\varepsilon) + A_3(n,\delta,\theta,\varepsilon)$, where

$$A_j(n,\delta,\theta,\varepsilon) = \int_{-\delta}^{\delta} (1 - c_\theta t^{2m})^n J_j(\theta + t)\, dt, \quad j = 1,2,3. \tag{4.8}$$

Note that here all $J_j(\theta + t)$ depend on n, ε, but for simplicity of notations we keep the dependence in the background.

Lemma 4.5. *Let $\varepsilon > 0$. Then $|A_1(n,\delta,\theta,\varepsilon)| \le C_1\sqrt{\varepsilon}$ for sufficiently large n, where $C_1 > 0$ is independent of n.*

Proof. By the Cauchy-Schwartz inequality we have

$$|A_1(n,\delta,\theta,\varepsilon)|^2 \le \int_{-\delta}^{\delta} (1 - c_\theta t^{2m})^{2n}\, dt \int_{-\delta}^{\delta} |J_1(\theta + t)|^2\, dt.$$

Using the orthogonality of $e^{2\pi ikt}$ on $\mathbb{T}$, we have

$$\int_{-\delta}^{\delta} |J_1(\theta + t)|^2\, dt \le \int_{-\frac{1}{2}}^{\frac{1}{2}} |J_1(\theta + t)|^2\, dt \le 4\varepsilon n^\sigma \|\mathbf{x}\|_\infty^2.$$

Also by Lemma 4.4, $\int_{-\delta}^{\delta} (1 - c_\theta t^{2m})^{2n}\, dt \le Cn^{-\sigma}$. The lemma now follows. ∎

Lemma 4.6. *Let $\varepsilon > 0$. Then $|A_3(n,\delta,\theta,\varepsilon)| \le C_3\varepsilon$ for sufficiently large n, where $C_3 > 0$ is independent of n.*

Proof. We first establish the inequality

$$\Big| \int_{-\delta}^{\delta} (1 - c_\theta t^{2m})^n e^{2\pi ikt}\, dt \Big| = 2 \Big| \int_{0}^{\delta} (1 - c_\theta t^{2m})^n \cos(2\pi kt)\, dt \Big| \le \frac{C_3' n^\sigma}{k^2} \tag{4.9}$$

for all $|k| > \varepsilon^{-1} n^\sigma$, where again $\sigma = \frac{1}{2m}$. The substitution $s = n^\sigma t$ yields

$$\int_{0}^{\delta} (1 - c_\theta t^{2m})^n \cos(2\pi kt)\, dt = \frac{1}{n^\sigma} \int_{0}^{\delta n^\sigma} g_n(s) \cos(Ls)\, ds \tag{4.10}$$

where $L = \frac{2\pi k}{n^\sigma}$ and $g_n(s) = (1 - \frac{c_\theta s^{2m}}{n})^n$. Again, $1 - \frac{c_\theta s^{2m}}{n} \le e^{-\frac{c_\theta s^{2m}}{n}}$ so $g_n(s) \le e^{-c_\theta s^{2m}}$. Observe that we have $g_n(\delta n^\sigma) = O(e^{-c_\theta \delta^{2m} n})$, $g_n'(0) = 0$ and $g_n'(\delta n^\sigma) = O(ne^{-c_\theta \delta^{2m} n})$. Combining these with integration by parts twice on (4.10) we obtain

$$\int_{0}^{\delta n^\sigma} g_n(s) \cos(Ls)\, ds = O(ne^{-c_\theta \delta^2 n}) - \frac{1}{L^2} \int_{0}^{\delta n^\sigma} g_n''(s) \cos(Ls)\, ds.$$

It is easy to check that $|g_n''(s)\cos(Ls)| \le (a_1 s^{2m-2} + a_2 s^{4m-2})e^{-c_\theta s^{2m}}$ for some constants $a_1, a_2 > 0$. Thus $\int_0^{\delta n^\sigma} g_n''(s)\cos(Ls)\,ds$ is bounded by $\int_0^\infty (a_1 s^{2m-2} + a_2 s^{2m-2})e^{-c_\theta s^{2m}}\,ds$, which is finite. Hence there exists $C_3'' > 0$ such that

$$\Big| \int_0^{\delta n^\sigma} g_n(s)\cos(Ls)\,ds \Big| \le \frac{C_3''}{L^2} = \frac{C_3'' n^{2\sigma}}{4\pi^2 k^2},$$

which yields (4.9). Finally by (4.9),

$$|A_3(n,\delta,\theta,\varepsilon)| \le C_3' \|\mathbf{x}\|_\infty \sum_{\varepsilon^{-1} n^\sigma < |k| \le qn} \frac{n^\sigma}{k^2} \le C_3 \varepsilon.$$

■

Lemma 4.7. *Assume that* $[\mathbf{x}, \mathbf{v}_{-\theta}] = 0$. *Let* $\varepsilon > 0$. *Then* $|A_2(n,\delta,\theta,\varepsilon)| \le C_2\varepsilon$ *for sufficiently large* n, *where* $C_2 > 0$ *is independent of* n.

Proof. Set $\mathbf{y} = (y_k) := (x_k e^{2\pi i k\theta})_{k\in\mathbb{Z}}$. Then $\widehat{\mathbf{y}_{qn}}(t) = \widehat{\mathbf{x}_{qn}}(\theta + t)$. By the fact that $\int_{-\delta}^{\delta} (1 - c_\theta t^{2m})^n \sin(2\pi k t)\,dt = 0$,

$$\begin{aligned} A_2(n,\delta,\theta,\varepsilon) &= \sum_{\varepsilon n^\sigma < |k| \le \varepsilon^{-1} n^\sigma} y_k \int_{-\delta}^{\delta} (1 - c_\theta t^{2m})^n e^{2\pi i k t}\,dt \\ &= 2 \sum_{\varepsilon n^\sigma < k \le \varepsilon^{-1} n^\sigma} (y_k + y_{-k}) \int_0^\delta (1 - c_\theta t^{2m})^n \cos(2\pi k t)\,dt. \end{aligned}$$

Now denote $S_k := \sum_{j=-k}^{k} y_j$ and $U_k = \int_0^\delta (1 - c_\theta t^{2m})^n \cos(2\pi k t)\,dt$. Then $y_k + y_{-k} = S_k - S_{k-1}$. Using summation by parts

$$A_2(n,\delta,\theta,\varepsilon) = \sum_{k=M}^{N} (S_k - S_{k-1})U_k = \sum_{k=M}^{N-1} S_k\,(U_k - U_{k+1}) - S_{M-1}U_M + S_N U_N$$

where $M = \lfloor \varepsilon n^\sigma \rfloor + 1$ and $N = \lfloor \varepsilon^{-1} n^\sigma \rfloor$. Using the fact $|S_k| \le a_1 k$ for some constant a_1 and Lemma 4.4 we have

$$|S_{M-1}U_M| \le a_1 \varepsilon n^\sigma \int_0^\delta (1 - c_\theta t^{2m})^n\,dt \le a_1' \varepsilon.$$

By (4.9) there exists some $a_2 > 0$ such that

$$|S_N U_N| \le a_1 C_3' N \frac{n^\sigma}{N^2} \le a_2 \varepsilon^{-1} n^\sigma \frac{n^\sigma}{(\varepsilon^{-1} n^\sigma)^2} = a_2 \varepsilon.$$

It remains to estimate $T := \sum_{k=M}^{N-1} S_k\,(U_k - U_{k+1})$. The hypothesis $[\mathbf{x}, \mathbf{v}_{-\theta}] = 0$ implies that $\lim_{k\to\infty} S_k/k = 0$. Thus for $n > N_0$ we have $\sup_{k \ge \varepsilon n^\sigma} |S_k/k| \le \varepsilon^3$. It follows from the

Cauchy-Schwartz inequality that

$$|T|^2 = \Big| \sum_{k=M}^{N-1} \frac{S_k}{k}\, k\,(U_k - U_{k+1}) \Big|^2 \le \varepsilon^6 \Big(\sum_{k=M}^{N-1} k^2 \Big) \Big(\sum_{k=M}^{N-1} (U_k - U_{k+1})^2 \Big).$$

Now $U_k - U_{k+1} = \int_0^\delta (1 - c_\theta t^{2m})^n \sin(\pi t) \sin(\pi(2k+1)t)\, dt$. Observe that the functions $\{\sqrt{2}\sin(\pi(2k+1)t)\}$ are orthonormal on $[0,1]$. Parseval's inequality yields

$$\sum_{k=M}^{N-1} (U_k - U_{k+1})^2 \le \frac{1}{2} \int_0^\delta (1 - c_\theta t^{2m})^{2n} \sin^2(\pi t)\, dt \le \frac{\pi^2}{2} \int_0^\delta (1 - c_\theta t^{2m})^{2n} t^2\, dt. \tag{4.11}$$

By Lemma 4.4

$$\int_0^\delta (1 - c_\theta t^{2m})^{2n} t^2\, dt \le C n^{-\frac{3}{2m}} = C n^{-3\sigma}.$$

Thus

$$|T|^2 \le a_3 \varepsilon^6 (\varepsilon^{-1} n^\sigma)^3 n^{-3\sigma} = a_3 \varepsilon^3.$$

These estimates show that for sufficiently large n we have

$$|A_2(n, \delta, \theta, \varepsilon)| \le C_2 \varepsilon.$$

∎

We can now complete the proof of Theorem 2.1. Let $\tilde{\mathbf{x}} := \mathbf{x} - \sum_{\theta \in Z_{\mathbf{a},1}} [\mathbf{x}, \mathbf{v}_\theta] \mathbf{v}_\theta$. Then $T_{\mathbf{a}}^n(\mathbf{x}) = T_{\mathbf{a}}^n(\tilde{\mathbf{x}}) + \sum_{\theta \in Z_{\mathbf{a},1}} [\mathbf{x}, \mathbf{v}_\theta] \mathbf{v}_\theta$. Note that $\tilde{\mathbf{x}}$ satisfies the hypothesis of Lemma 4.7. Combining Lemma 4.3 and Lemmas 4.5-4.7 yields $T_{\mathbf{a}}^n(\tilde{\mathbf{x}})(0) \to 0$. Thus

$$\lim_{n \to 0} T_{\mathbf{a}}^n(\mathbf{x})(0) = \sum_{\theta \in Z_{\mathbf{a},1}} [\mathbf{x}, \mathbf{v}_\theta] \mathbf{v}_\theta(0) = \sum_{\theta \in Z_{\mathbf{a},1}} [\mathbf{x}, \mathbf{v}_\theta].$$

Finally, let τ be the left shift operator on $l^\infty(\mathbb{Z})$, i.e. $\tau((x_k)) = (x_{k+1})$. Then $T_{\mathbf{a}} \circ \tau = \tau \circ T_{\mathbf{a}}$. It follows that

$$T_{\mathbf{a}}^n(\mathbf{x})(k) = \tau^k \circ T_{\mathbf{a}}^n(\mathbf{x})(0) = T_{\mathbf{a}}^n \left(\tau^k(\mathbf{x}) \right)(0).$$

But $[\tau^k(\mathbf{x}), \mathbf{v}_\theta] = [\mathbf{x}, \mathbf{v}_\theta] e^{2\pi i k \theta}$ for $\theta \in Z_{\mathbf{a},1}$. Thus

$$T_{\mathbf{a}}^n(\mathbf{x})(k) = T_{\mathbf{a}}^n \left(\tau^k(\mathbf{x}) \right)(0) = \sum_{\theta \in Z_{\mathbf{a},1}} [\mathbf{x}, \mathbf{v}_\theta] e^{2\pi i k \theta}.$$

This completes the proof of Theorem 2.1.

Remark: Lemma 4.7 is the only place where the condition $[\mathbf{x}, \mathbf{v}_\theta]$ exists for all $\theta \in Z_{\mathbf{a},1}$ is being used. With this condition we may apply summation by parts and the convergence of S_k/k to obtain the necessary final estimates. It is also clear from the proof that we can apply summation by parts again to show the following: Let $S_k(\theta) := \sum_{j=-k}^{k} x_j e^{-2\pi j\theta}$ and $S'_k(\theta) := \sum_{1\leq |j|\leq k} S_j(\theta)/j$. Assume that $\lim_{k\to\infty} S'_k(\theta)/k$ exists for every $\theta \in Z_{\mathbf{a},1}$ then the conclusion of the theorem still holds. Unfortunately the convergence of $S'_k(\theta)/k$ is equivalent to the convergence of $S_k(\theta)/k$. We shall omit the proof here.

References

[1] R. Bartle, *The Elements of Real Analysis*, 2nd Edition, John Wiley and Sons, 1976.

[2] B. Boashash, Estimating and interpreting the instantaneous frequency of a signal, Proc. IEEE 80 (1992), 417-430.

[3] A. Böttcher and S. Grudsky, *Toeplitz Matrices, Asymptotic Linear Algebra, and Functional Analysis*, Birkhäuser, 2000.

[4] Q. Chen, N. Huang, S. Riemenschneider and Y. Xu, B-spline approach for empirical mode decomposition, preprint.

[5] L. Cohen, Time-Frequency Analysis. Englewood Cliffs, NJ, Prentice Hall (1995).

[6] A. Cordoba, Dirac combs, Lett. Math. Phys. 17 (1989), 191-196.

[7] J. C. Echeverria, J. A. Crowe, M. S. Woolfson and B. R. Hayes-Gill, Application of empirical mode decomposition to heart rate variability analysis, Medical and Biological Engineering and Computing 39 (2001), 471-479.

[8] G. Folland, *Real Analysis and Its Applications*, Birkhauser.

[9] L. Hörmander, *The Analysis of Linear Partial Differential Operators I*, Springer-Verlag, 1983.

[10] N. Huang *et al*, The empirical mode decomposition and the Hilbert spectrum for nonlinear nonstationary time series analysis, Proceedings of Royal Society of London A 454 (1998), 903-995.

[11] N. Huang, Z. Shen and S. Long, A new view of nonlinear water waves: the Hilbert spectrum, Annu. Rev. Fluid Mech. 31 (1999), 417-457.

[12] C. Huang, Y. Wang and L. Yang, Convergence of a convolution-filtering-based algorithm for empirical mode decomposition, *Advances in Adaptive Data Analysis* 1 No. 4 (2009), 561–571.

[13] J. Lagarias, Mathematical quasicrystals and the problem of diffraction, in: *Directions in Mathematical Quasicrystals, M. Baake and R. V. Moody, Eds.*, CRM Monograph Series, Vol. 13, Amer. Math. Soc., Providence, RI, 2000, pp. 61–93 .

[14] L. Lin, Y. Wang, and H. Zhou, Iterative filtering as an alternative algorithm for empirical mode decomposition, *Advances in Adaptive Data Analysis* Vol. 1, No. 4 (2009) 543–560.

[15] B. Liu, S. Riemenschneider and Y. Xu, Gearbox fault diagnosis using empirical mode decomposition and hilbert spectrum, preprint.

[16] S. Mallat, A Wavelet Tour of Signal Processing. London, Academic Press (1998).

[Mao 2010] D. Mao, D. Rockmore, Y. Wang, and Q. Wu, EMD analysis of visual stylometry, preprint.

[17] D. Mao, Y. Wang, and Q. Wu, A new approach for analyzing physiological time series, preprint.

[18] D. Pines and L. Salvino, Health monitoring of one dimensional structures using empirical mode decomposition and the Hilbert-Huang Transform, Proceedings of SPIE 4701(2002), 127-143.

[19] Z.-G. Yu, V. Anh, Y. Wang and D. Mao, Modeling and simulation of the horizontal component of the magnetic field by fractional stochastic differential equation in conjunction with epirical mode decomposition, *Journal of Geophysical Research - Space Physics*, to appear.

Department of Mathematics, Michigan State University, East Lansing, MI 48824-1027, USA.

E-mail address: `ywang@math.msu.edu`

Department of Mathematics, Michigan State University, East Lansing, MI 48824-1027, USA.

E-mail address: `zfzhou@math.msu.edu`