Sparse Time-Frequency Data Analysis: A Multi-scale Approach

Thesis by

Peyman Tavallali

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2014

(Defended May 9, 2014)

Acknowledgements

I need to express my deepest gratitude to my research adviser, Prof. Thomas Y. Hou, for his support, encouragement, and guidance throughout my Ph.D. research journey at Caltech. I have benefited greatly from his broad and deep insight into mathematics. I especially need to thank him for introducing me to the area of multiscale-adaptive data analysis, which led to the research in the topic of this thesis.

I am grateful to many other academic members and advisors. I would like to thank Dr. Mojtaba Mahzoon, my research and life adviser during my B.Sc. degree at Shiraz University, for his friendship, mentorship, and for introducing me to the interesting and charming field of applied mathematics. His logical advice helped me to overcome my uncertainty about continuing my studies outside of my home country. His words are always moral guidance for me. I am also grateful to Prof. Vladimir V. Kulish, my research adviser in Singapore, for his support during my stay in Nanyang Technological University. I would like to thank Prof. Morteza (Mory) Gharib, who gave me the opportunity to work on biological signals using our newly developed signal processing method, which resulted in some interesting interdisciplinary research results that could be used to save thousands of lives.

I want to thank Professors Thomas Y. Hou, Houman Owhadi, James L. Beck , and Mory Gharib whose comments as members of my dissertation committee were extremely valuable to my research.

It has always been a great pleasure to interact and collaborate with people at Caltech. First, I would like to thank Dr. Niema M. Pahlevan. Our collaboration and discussions on heart and cardiovascular signals have enabled a promising future for teletransport diagnosis. I would like to thank Dr. Derek Rinderknecht for helping

Niema and me develop our ideas. I would also like to thank Dr. Zuoqiang Shi for helping me understand many aspects of signal processing, giving me a deeper understanding of the topics in my dissertation. Furthermore, I am grateful to him for our satisfying collaboration which led to indepth publications on some of the topics from my dissertation. Finally, I thank my collegue Diane Guignard for helping me through some challenging problems while she worked on her Master's dissertation at Caltech.

I would also like to thank all the people who made my stay at Caltech a memorable experience. Special thanks to Carmen Nemer-Sirois, Sydney Garstang, Shiela Shull, Maria Lopez, and Dr. Guo Luo for helping me whenever I needed help. Thanks to my climbing friends who made my life more joyful: Fabian Schmidt, Erin Burkett, Thomas Clement, Sebastien Henz, Aurelien Hees, Victoria Meyer, Anna Kozlowska and Olgierd (Olek) Sulejman who is not only a climbing buddy and friend, but also a brother from other parents. Moreover, I would like to thank my friends Kaweh Hosseini, Ahmad Serjoei, Abouzar Kaboudian, Hamed Keramati, Alireza Akbarzadeh, Meisam Honarvar Nazari, Gita Mahmoudabadi, Arman MohsenNia, Arad Gharebaghi, Ali Rahim Taleqani, Amin Khajehnejad, Saeid Farivar, and Behrooz Abiri for creating a family-like environment for me.

Last but not least, I want to express my humble gratitude to my parents, Parvaneh Foolad and Javad Tavallali, my brothers, Hooman and Pooya, and my fiancé, Samira Rahim Taleghani. This journey would be impossible without their unconditional support, compassion and love. Certainly, Samira's warm heart helped me through ups and downs in this journey. Thank you for being there for me.

Abstract

In this work, we further extend the recently developed adaptive data analysis method, the Sparse Time-Frequency Representation (STFR) method. This method is based on the assumption that many physical signals inherently contain AM-FM representations. We propose a sparse optimization method to extract the AM-FM representations of such signals. We prove the convergence of the method for periodic signals under certain assumptions and provide practical algorithms specifically for the non-periodic STFR, which extends the method to tackle problems that former STFR methods could not handle, including stability to noise and non-periodic data analysis. This is a significant improvement since many adaptive and non-adaptive signal processing methods are not fully capable of handling non-periodic signals. Moreover, we propose a new STFR algorithm to study intrawave signals with strong frequency modulation and analyze the convergence of this new algorithm for periodic signals. Such signals have previously remained a bottleneck for all signal processing methods. Furthermore, we propose a modified version of STFR that facilitates the extraction of intrawaves that have overlaping frequency content. We show that the STFR methods can be applied to the realm of dynamical systems and cardiovascular signals. In particular, we present a simplified and modified version of the STFR algorithm that is potentially useful for the diagnosis of some cardiovascular diseases. We further explain some preliminary work on the nature of Intrinsic Mode Functions (IMFs) and how they can have different representations in different phase coordinates. This analysis shows that the uncertainty principle is fundamental to all oscillating signals.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1. Preamble

Accessing and utilizing the information hidden in such signals requires methods for processing and analyzing signals. Such methods must be able to denoise and analyze the signal in order to properly process the data. Mathematically, the easiest way to construct a signal processing method is to project the recorded signal on a predetermined algebraic basis or dictionary. A classical method of doing this task is the Fourier Transform (FT) method and a more recent one is the Wavelet Transform (WT) method [5, 41].

When a signal is periodic, the FT method is a powerful signal processing method. Using the FT method, one can project the data into the orthonormal basis of the Fourier domain (the frequency domain). This transition is a shift from the time domain to the frequency domain. The FT method can also perform as a robust method for denoising periodic signals. Fast and efficient implementation of FT, namely the Fast Fourier Transform (FFT) method, is popular among scientists and engineers. The shortcomings of the FT method can be categorized into two different groups. The first is the assumption of the periodicity (or stationarity) of the signal.

In fact, most of the signals that we encounter in practice are non-periodic. The second problem is the lack of time information in the frequency domain: since the frequencies extracted by the FT method are constants, one cannot perceive whether a certain event has occurred at a certain time simply by looking at the frequency content of the signal and the coefficients of the analysis. In order to overcome the time-frequency resolution of the FT method, the WT method was proposed as a

method that incorporates a time-frequency analysis of the signal by constructing a large dictionary of some orthonormal functions.

The FT and WT methods share one common property: decomposition is performed on a predefined basis, which is troublesome if the signal is not stationary. Recently, Norden Huang proposed Empirical Mode Decomposition (EMD), a new method of adaptive signal processing [30, 32, 33] in which the basis of the projection is adaptive. EMD which uses multiscale data-driven decompositions called Intrinsic Mode Functions (IMFs), is a step forward in data analysis. It has eliminated most of the issues present in the FT and WT methods.

In particular, EMD can produce a faithful extraction even if the signal is not periodic, and makes a sparser time-frequency analysis of the data. In fact, projection into a basis is not the ultimate goal in many recently developed signal processing methods. Researchers wish to have a projection that is as sparse as possible. In other words, it is important to have a representation of the signal in a basis by keeping only a few coefficients containing the pertinent information. In fact, in these methods, one should project the observed signal on a large overdetermined basis (dictionary [10, 12, 42]).

Since the IMFs are extracted adaptively from the data in EMD, the final decomposition is in general sparser than FT or WT method the extractions. If the data has a certain frequency scale-separation property, the extracted IMFs convey certain physical properties of the signal. Unfortunately, the empirical nature of the the EMD's decomposition makes it hard to analyze the results rigorously. In order to eliminate this problem, Hou and Shi have proposed a rigorous mathematical system as a counterpart of the EMD method [25, 24, 26]: the Sparse Time-Frequency Representation (STFR) method.

1.1.1. Scope of the Thesis. The main focus of this thesis is to further develop STFR methods. This development is two-fold. First, we introduce a non-periodic

STFR method that can analyze non-periodic signals and can robustly deal with signals polluted with noise. Second, we propose and prove the convergence of a new STFR method that can easily analyze intrawave signals, signals with intense Frequency Modulation (FM) that cannot currently be handled using Hou and Shi's STFR methods [25, 24, 26].

Technically speaking, this analysis can be done by relaxing the definition of the dictionaries that Hou and Shi have already used, see (1.2.3). To the best of our knowledge, this is the first time that intrawave signals can be extracted with high accuracy by an adaptive method.

This thesis also presents some real physical applications of our methods. We show that our methods can be applied in diverse fields like dynamical systems and cardiovascular signals. We demonstrate that our method is capable of extracting useful physical and biomedical information from the signals. Such information can possibly be used in system identification and medical diagnosis.

Finally, the thesis also clarifies theoretical understanding of the nature of IMFs. We show that an IMF can have infinitely many representations and that consequently, the uncertainty principle is an indispensable part of oscillating signals. Lacking a unique representation does not decrease the merit of any adaptive data analysis method, including the STFR method, but only shows the richness of the field and requires that an adaptive method can pick a certain representation based on a fixed preference.

1.2. A Brief Introduction to the STFR Method and its Applications

All STFR methods are based on the assumption that a relatively big subclass of the oscillatory signals are signals of the form

$$x(t) = a(t)\cos\theta(t), \tag{1.2.1}$$

with only one extrema between the zeros of the signal, in which the envelope is strictly positive, $a(t) > 0$, and the phase function $\theta(t)$ is a one-to-one, strictly increasing map between the time coordinate, t, and the phase coordinate, θ. The time derivative of this phase function is called the Instantaneous Frequency (IF). With some abuse of notation, we can say that the STFR methods deal with signals that have both Amplitude Modulation (AM) and Frequency Modulation (FM). The type of signal, in Equation (1.2.1) is called an Intrinsic Mode Function (IMF) in STFR and EMD terminology. A number of methods can extract each IMF from a combination of many IMFs, with different levels of accuracy. Methods that perform such extraction well include, but are not limited to, EMD [32], EEMD [56], Optimization based EMD [27], Wavelet [41], STFR [25, 24, 26], and Synchrosqueezed wavelet transforms [17]. However, when it comes to signals with strong frequency modulation, these methods have difficulties extracting a unique IMF specifically when the data is polluted with noise. Among these methods, the STFR method provides a better physical and mathematical understanding.

Physically speaking, $\theta(t)$ carries information about the rate of the change of the signal in time. The envelope of each IMF is set to be from a certain collection of functions. In mathematical terms we can express it as

$$a(t) \in V(\theta(t)) \ s.t. \ a(t) > 0, \tag{1.2.2}$$

where

$$V(\theta) = span\left\{1, \cos\left(\frac{\theta}{\lambda}\right), \sin\left(\frac{\theta}{\lambda}\right) \mid \lambda \geqslant 2\right\}. \tag{1.2.3}$$

A finite linear combination of a collection of the IMFs is called an Intrinsic Signal (IS),

$$s(t) = \sum_{i=1}^{M} a_i(t) \cos\theta_i(t). \tag{1.2.4}$$

The goal of the STFR method is to extract each IMF from an IS.

1.2.1. EMD Method. Most of the non-adaptive time-frequency signal analysis methods are based on the projection of a signal on to a predetermined basis or dictionary, such as a Fourier transform (FT) or Wavelet transform (WT). However, since the projection of the signal is on a predefined set of functions, the extraction is not adaptive and, possibly, not sparse. In order to get a better time-frequency resolution, the intuition is to project the signal on an adaptive basis that is found from the signal itself. Empirical Mode Decomposition and the more recent Ensemble Empirical Mode Decomposition were developed to do just that.

Empirical Mode Decomposition (EMD) methods were originally proposed to overcome the problem of predetermined bases which was hampering a proper sparse decomposition in the WT and FT methods. EMD [32] tries to find the basis of decomposition adaptively from the original signal in an empirical way. The EMD method uses Intrinsic Mode Function (IMF) [32] as constructing blocks of signals. An EMD definition of an IMF is:

Definition. An IMF is a function that satisfies the following two conditions:

- The number of extrema and the number of zero crossings are equal or differ at most by one on the whole data set
- The mean value of the envelopes defined by the local maxima and the local minima is zero at any point.

EMD decomposes a signal into an extracted IMF and a residue through a sifting process. The extracted IMF is then post-processed to find the Instantaneous Frequency (IF) either through Hilbert Transform (HT) or another sifting process [32, 33]. To summarize, for a signal $x(t)$, the decomposition looks like

$$x(t) = \sum_{k=1}^{n} c_k(t) + r_n(t),$$

where c_k's are the IMFs extracted from the signal and $r_n(t)$ is the residue.

1.2.2. Ensemble Empirical Mode Decomposition (EEMD) Method. EMD is highly sensitive to noise. In order to fix this shortcoming, Wu and Huang [56] proposed EEMD in 2009. EEMD minimizes noise effects by adding white noise to the signal. While the EMD and EEMD methods have attracted a lot of attention, they lack a theoretical basis. The STFR methods were introduced to fill the empirical gap.

1.2.3. STFR Methods. Although STFR methods were inspired by the EMD and EEMD methods [32, 56], they make use of techniques from compressive sensing theory [10, 12] and matching pursuit [42]. Compressive Sensing (CS) and Matching Pursuit (MP) methods try to find a sparse reconstruction of an observed signal from a predefined finite large dictionary. STFR, on the other hand tries to find a sparse reconstruction of an observed signal from a predefined *infinite* large dictionary. STFR methods use two major steps: they first construct a highly redundant dictionary of all IMFs, namely D, and then find the sparsest decomposition by solving a non-linear and non-convex optimization problem

(1.2.5)

$$\begin{aligned} &Minimize && M \\ &Subject\ to: && s(t) = \textstyle\sum_{i=1}^{M} a_i(t)\cos\theta_i(t), \quad a_i(t)\cos\theta_i(t) \in D, \quad i = 1, ..., M, \end{aligned}$$

which is an L_0 minimization problem. We explain possible approximations by which one can find a close solution of this problem. Hou and Shi have proposed a number of algorithms for approximating this problem [25, 24, 26] by decomposing the signal into two parts, a mean a_0 and a modulated oscillatory part, namely the IMF, $a_1 \cos\theta$:

$$s(t) = a_0(t) + a_1(t)\cos\theta(t),$$

where θ, a_1, and a_0 are the unknown phase function, envelope, and mean respectively. Upon finding one of the IMFs, $a_1 \cos\theta$, the residue (or the mean) a_0 is treated as a new signal and the same procedure is repeated until the residue is smaller than a preset value.

1.2.3.1. *Total Variation Method.* Hou and Shi [24] proposed the TV^3 minimization of the form

$$\begin{aligned} &Minimize && TV^3\left(a_0\right) + TV^3\left(a_1\right) \\ &Subject\ to: && s\left(t\right) = a_0\left(t\right) + a_1\left(t\right)\cos\theta_1\left(t\right), \quad \theta_1'\left(t\right) \geqslant 0, \end{aligned} \tag{1.2.6}$$

as an approximation to the original problem (1.2.5). This problem is solved in an iterative manner. First, an initial guess on θ_1 is set. Then based on that, the value of a_1 is approximated and the phase θ_1 is again updated. This procedure is repeated until there is no progress in updating the phase function. This approach is an exponential step compared to the EMD and EEMD methods; however, the algorithm is not stable to noise.

1.2.3.2. *Periodic Fourier-based Sparse Time-Frequency Method.* To address the noise instability problem, Hou and Shi, proposed the Periodic Fourier-based STFR method [25], which performs well on periodic data even in the presence of noise. This method implicitly solves

$$\begin{aligned} &Minimize && \left\|s\left(t\right) - a\left(t\right)\cos\theta\left(t\right)\right\|_2^2 \\ &Subject\ to: && a\left(t\right)\cos\theta\left(t\right) \in D \end{aligned}$$

using a Fast Fourier Transformation (FFT) iterative scheme. The Periodic STFR algorithm works in the following way: an initial guess is proposed on the phase function θ_0. In the first step, the whole signal is mapped to the phase function space and the Fourier transform of the signal is then found in that space. At this step, the possible candidates for the envelope functions are found. Later, the new phase function is updated and the algorithm begins a new iteration.

While this algorithm works well for periodic data, the algorithm does not extract the IMFs properly for non-periodic data.

1.2.3.3. *Non-Periodic Sparse Time-Frequency Method.* TV STFR is not accurate for noisy signals and Periodic STFR is not suitable for non-periodic signals. Non-periodic STFR has a slower speed of convergence than Periodic STFR since it uses a

L_1-norm regularized with L_2-norm optimization rather than FFT for each iteration. Nevertheless, it can successfully be applied to non-periodic signals and also to signals polluted with noise.

In the non-periodic STFR algorithm, we assume that the envelopes of the IMFs have a sparse structure in their respective dictionary. This assumption is not far from reality since we are using an infinitely large dictionary when we try to extract IMFs. This assumption can be formulated in the following way:

$$\begin{array}{ll} \underset{a,b,\theta}{Minimize} & \delta\left(\|a\left(\lambda\right)\|_1+\|b\left(\lambda\right)\|_1\right)+\|s\left(t\right)-\mathcal{A}\left(t\right)\cos\theta\left(t\right)\|_2^2 \\ Subject\ to: & \mathcal{A}\left(t\right)=\int_2^\infty\left(a\left(\lambda\right)\cos\frac{\theta(t)}{\lambda}+b\left(\lambda\right)\sin\frac{\theta(t)}{\lambda}\right)d\lambda, \\ & \frac{d\theta(t)}{dt}>0. \end{array} \tag{1.2.7}$$

In this formulation, the envelope $\mathcal{A}\left(t\right)$ of the IMF $\mathcal{A}\left(t\right)\cos\theta\left(t\right)$ is assumed to have a sparse structure that is captured by $\|a\left(t\right)\|_1+\|b\left(t\right)\|_1$. In this formulation, the dictionary is used explicitly by $\mathcal{A}\left(t\right)=\int_2^\infty\left(a\left(\lambda\right)\cos\frac{\theta(t)}{\lambda}+b\left(\lambda\right)\sin\frac{\theta(t)}{\lambda}\right)d\lambda$. The following example shows the performance of this type of STFR.

Example. We tested the algorithm on a signal with a linear trend and two constant envelope and frequency IMFs, $f\left(t\right)=6t+\cos\left(8\pi t\right)+0.5\cos\left(40\pi t\right)$ (see Figure 1.1). The algorithm extracted IMFs (see red lines in Figures 1.2-1.4). Besides some tiny boundary misalignment, the extractions were accurate, suggesting that the non-periodic STFR method is accurate away from the boundaries.

When compared with other STFR methods, the only shortcoming of the Non-Periodic STFR method is the speed of the algorithm.

1.2.4. Extraction of Intrawave, Sharp, and Rare Event Signals using Sparse Time-Frequency Method. In general, intrawave signals are oscillatory signals that have intense frequency modulation in at least one θ-coordinate. By intense modulation we mean that the IF has oscillations that is comparable with the oscillation of the IMF itself. The EMD method can extract one intrawave IMF in the

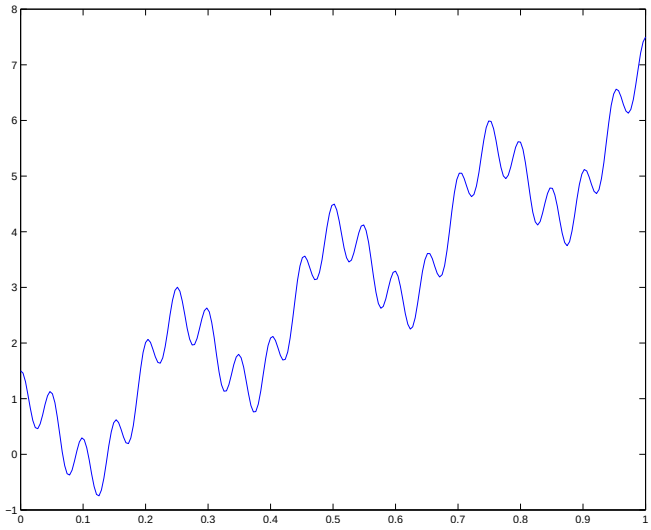

FIGURE 1.1. Signal with a Linear Trend: The horizontal axis is the time variable and the vertical one is the signal itself.

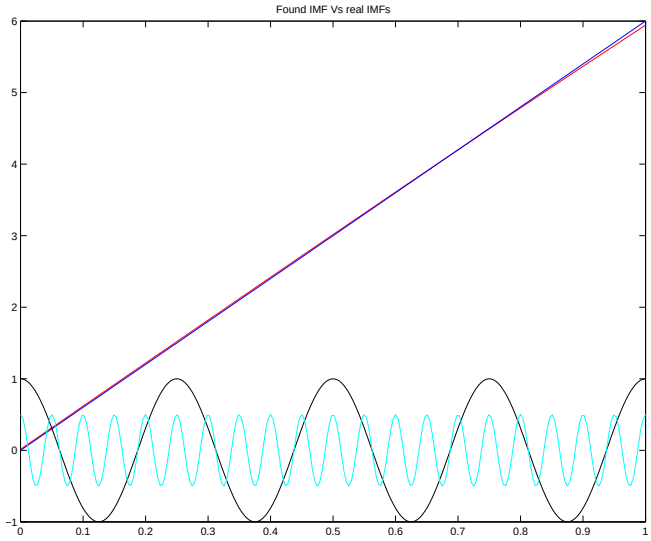

FIGURE 1.2. Extracted Trend: The linear trend is in blue and the extracted trend is in red. Except the right boundary, the error is small in the extracted trend.

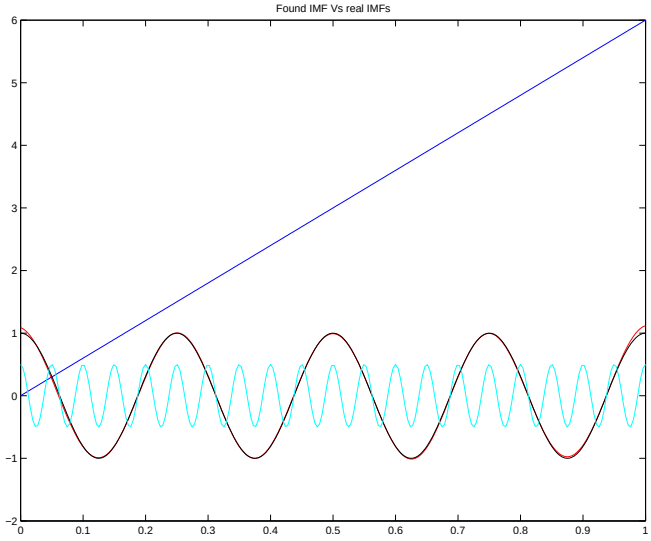

FIGURE 1.3. Extraction of the first IMF: The extracted IMF is in red. As can be seen, except from the boundaries the extraction is faithful.

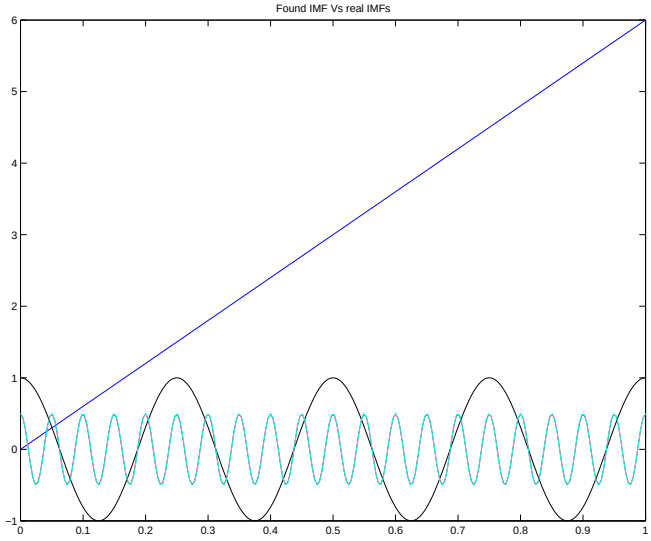

FIGURE 1.4. Extraction of the second IMF: The extracted second IMF is in red. It is almost indistinguishable from the high-frequency original IMF.

absence of noise. However, neither EMD nor EEMD can extract even one intrawave IMF in the presence of noise. If the frequency modulation becomes even more intense, the resulting signal is called a sharp signal. Analyzing these IMFs has so far been a challenging problem in signal processing. Rare events are IMFs with compact support in the time domain. A rare event is essentially like a spike. Adaptive methods like EMD/EEMD are not able to process such signals with acceptable accuracy.

Since the main difficulty in dealing with intrawave signals comes from their wide band representation in the frequency domain, which cannot be properly analyzed using methods with explicit or implicit narrow band filters, we propose a method that modifies the normal envelope dictionary in an STFR framework in order to extract intrawave signals with high accuracy. This small modification, which is enlarging the STFR filter, allows us to treat intrawave signals without major changes to the original STFR Algorithms. Furthermore, we show that although enlarging the filter requires that the IMF components of the IS must have enough separate time-frequency representations, the method is not problematic when extracting non-separable time-frequency IMFs from a signal provided they are extracted simultaneously.

Not only is the algorithm that we use to extract the IMFs stable, then, it is also stable to noise perturbation. The EMD/EEMD methods fail to extract the two IMFs properly. In fact all other adaptive methods fail to extract one IMF with intrawave frequency modulation in the presence of noise, let alone two IMFs with intrawave characteristics that have mode mixture.

1.2.5. Analysis of Convergence of Sparse Time-Frequency Method. We prove that for any signal, whether intrawave or not, increasing the filter span reduces extraction error. We show that STFR converges to an IMF that is close to one of the IMF representations, but with an error associated with the width (span) of the filter. We assume that an IS can be represented in the following format:

$$f(t) = f_0(t) + f_1(t)\cos\theta(t), \tag{1.2.8}$$

for $f_1(t) > 0$, $\frac{d\theta}{dt} = \theta' > 0$ and $t \in [0,1]$. We assume that the signal is periodic with mean zero. The main convergence theorem states that:

Theorem. *Assume that the instantaneous frequency in equation (1.2.8) is M_0-sparse; i.e. $\theta' \in V_{M_0} = span\left\{e^{i2\pi kt}, |k| \leq M_0, k \in \mathbb{Z}\right\}$. Furthermore, assume that*

$$\left|\hat{f}_{0,\bar{\theta}}(k)\right| \leq \frac{C_0}{|k|^p}, \left|\hat{f}_{1,\bar{\theta}}(k)\right| \leq \frac{C_0}{|k|^p}$$

for $C_0 > 0$ and $p \geq 4$. If the initial guess satisfies $\frac{\left\|\mathcal{F}\left(\triangle\theta^0\right)'\right\|_1}{2\pi M_0} \leq \frac{1}{4}$, then there exists an $\eta_0 > 0$ such that for $L > \eta_0$ we have

$$\left\|\mathcal{F}\left(\theta - \theta^{m+1}\right)'\right\|_1 \leq \Gamma_1 \lambda^{2p-2} L^{-p+2} + \tfrac{1}{2}\left\|\mathcal{F}\left(\theta - \theta^m\right)'\right\|_1$$

for $\lambda > 1$ and $\Gamma_1 > 0$.

Here, $\hat{f}_{\bar{\theta}}(k) = \int_0^1 f\left(\bar{\theta}\right) e^{-i2\pi k\bar{\theta}} d\bar{\theta}$ and $\left(\mathcal{F}(g)\right)_k = \int_0^1 g(t) e^{-i2\pi kt} dt$. This theorem can be generalized to the case where the signal is polluted with noise. Take

$$f(t) = f_0(t) + f_1(t)\cos\theta(t) + \Im, \tag{1.2.9}$$

where $\Im$ is a periodic perturbation to the original signal.

Theorem. *Assume that the instantaneous frequency in equation (1.2.9) is M_0-sparse; i.e. $\theta' \in V_{M_0}$. Furthermore, assume that*

$$\left|\hat{f}_{0,\bar{\theta}}(k)\right| \leq \frac{C_0}{|k|^p}, \left|\hat{f}_{1,\bar{\theta}}(k)\right| \leq \frac{C_0}{|k|^p}$$

for $C_0 > 0$ and $p \geq 4$. If the initial guess satisfies $\frac{\left\|\mathcal{F}\left(\triangle\theta^0\right)'\right\|_1}{2\pi M_0} \leq \frac{1}{4}$, then there exists an $\eta_0 > 0$ such that for $L > \eta_0$, and $\|\Im\|_\infty \leq \epsilon_0$ (ϵ_0 sufficiently small) we have

$$\left\|\mathcal{F}\left(\theta - \theta^{m+1}\right)'\right\|_1 \leq \Upsilon_0(L,\lambda)\|\Im(t)\|_\infty + \Gamma_1\lambda^{2p-2}L^{-p+2} + \tfrac{1}{2}\left\|\mathcal{F}\left(\theta - \theta^m\right)'\right\|_1$$

for $\lambda > 1$, $\Gamma_1 > 0$ and $\Upsilon_0(L,\lambda)$.

These theorems explain only that we have a convergent algorithm, not that the algorithm's extraction is unique. Uniqueness is a difficult theoretical problem that requires further research for more clarification.

1.2.6. Applications of Sparse Time-Frequency Method. In this thesis, we study the application of STFR in two different fields: dynamical systems and cardiovascular disease diagnosis.

1.2.6.1. *Dynamical Systems:* In many scientific applications that requires signal analysis, such as biological investigations, the complexity of the underlying physical problem is perplexing and the appropriate governing equation that describes its dynamics is unknown. Researchers would like to be able to determine whether the underlying dynamical system is linear or nonlinear, including quantifying the degree of any nonlinearity.

We present a method for quantifying the nonlinearity of IMFs given by the STFR method. The main idea is to establish a connection between the IMFs and classical second order differential equations. We show that each IMF can be associated with a solution of a second order ordinary differential equation of the form $\ddot{x} + p(x,t)\,\dot{x} + q(x,t) = f(t)$. This method provides a new way to interpret the hidden intrinsic information contained in the extracted IMF of a signal.

1.2.6.2. *Cardiovascular Disease Diagnosis:* Cardiovascular Diseases (CVD) are one of the main causes of death in the United States every year [39]. With an increasing number of deaths every year, there is a need to develop new CVD diagnosis methods.

Using the Periodic STFR method, we observed that the IF changes its trend before and after the closure of heart aortic valve (i.e. dicrotic notch). Since we observed the the IF trend shift, we proposed a modified version of STFR, the Intrinsic Frequency (InF) method, to better address the special type of the signals that we were analyzing. The InF method assumes that there are two constant dominant frequencies before and after the dicrotic notch. These frequencies are called intrinsic frequencies.

It is assumed that before and after the dicrotic notch, we have the following simple waveforms for the general IMF of the aortic pressure wave at time t:

$$s_i = a_i \cos \omega_i t + b_i \sin \omega_i t + \bar{p}, \quad i = 1, 2.$$

This assumption has shown its credibility as an index to characterize the heart and cardiovascular diseases [44]. In this formula, $i = 1$ corresponds to the behavior of the IMF before the valve closure, and $i = 2$ to the behavior of the IMF after that. Here, a_i, b_i are constants and correspond to the envelopes of the IMF. The constants ω_1, ω_2 correspond to the InFs of the IMF. $\bar{p}$ is the mean pressure during the heartbeat period.

Take $[0, T]$ to be the whole period of the pressure wave and T_0 as the location of the dicrotic notch: $0 < T_0 < T$. Also, define the *indicator function* as

$$\mathbf{1}_{[x,y)}(t) = \begin{cases} 1, & x \leq t < y, \\ 0, & else. \end{cases}$$

Now, the main IMF of the pressure waveform can be expressed as

$$\begin{aligned} S(a_i, b_i, \bar{p}, \omega_i; t) &= (a_1 \cos \omega_1 t + b_1 \sin \omega_1 t + \bar{p}) \mathbf{1}_{[0,T_0)}(t) + \\ &\quad (a_2 \cos \omega_2 t + b_2 \sin \omega_2 t + \bar{p}) \mathbf{1}_{[T_0,T)}(t). \end{aligned}$$

The goal is to extract the IMF carrying most of the energy and consequently the InFs, ω_1, ω_2, from the observed aortic pressure waveform $f(t)$. Taking t as a continuous variable, one can use the least squares minimization to find the unknowns.

$$\underset{a_i, b_i, \omega_i, \bar{p}}{minimize} \qquad \|f(t) - S(a_i, b_i, \bar{p}, \omega_i; t)\|_2^2$$

$$\begin{aligned} subject\ to \quad a_1 \cos \omega_1 T_0 + b_1 \sin \omega_1 T_0 &= a_2 \cos \omega_2 T_0 + b_2 \sin \omega_2 T_0, \\ a_1 &= a_2 \cos \omega_2 T + b_2 \sin \omega_2 T. \end{aligned}$$

Using this modified version of the STFR method, we can identify the optimum heart rate from the aortic pressure wave alone.

We employed the InF algorithm on pressure wave signals collected from human beings (both invasively using a catheter and non-invasively using an iPhone camera) and dogs (invasively using a catheter). We found that heart performance (the Ejection Fraction[1] (EF)) can be predicted using the normalized values of the InFs and the normalized value of $\bar{p}$. The performance of the InF algorithm clearly shows potential for use in health care systems.

[1]Ejection Fraction is essentially a measure of the percentage of blood leaving the heart in each contraction. Ejection Fraction is a good measure of the performance of the heart. A very low Ejection Fraction corresponds to some cases of Heart Failure. A very traditional way of measuring the Ejection Fraction is through Echocardiography.

CHAPTER 2

Literature Review

While there are different methods for signal processing, this chapter provides an introductory overview of the more prevalent methods of signal decomposition and processing, focusing on Fourier methods, Wavelet methods, and Empirical Mode Decomposition methods, as well as a brief review of the Total Variation (TV) STFR method[24], a first generation STFR method. Further details on these topics can be found in the references cited in this chapter. This overview follows the steps of the work by Guignard [23].

2.1. Fourier Methods

The classical approach to signal decomposition, Fourier analysis or Fourier Transform, was first proposed by French mathematician Joseph Fourier, who initially used it to solve heat transfer problems [7]. The most famous because of its simplicity and accuracy, Fourier analysis decomposes a signal over a frequency space. Here, we briefly introduce the method in mathematical terms.

2.1.1. Fourier Series.

Definition 1. Take the 2π-periodic real function $f \in L^1(0, 2\pi)$, and $N \in \mathbb{N}$. The *Partial Fourier Series* of order N of the function f is defined by

$$\mathcal{F}_N(f(t)) = a_0 + \sum_{n=1}^{N} (a_n \cos(nt) + b_n \sin(nt)), \tag{2.1.1}$$

having

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(t)\, dt,$$

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(nt)\, dt,$$

$$b_n = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin(nt)\,dt.$$

This real definition can be extended into the complex form.

$$\mathcal{F}_N(f(t)) = \sum_{n=-N}^{N} c_n e^{int}, \tag{2.1.2}$$

with

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} f(t)\,e^{-int}dt. \tag{2.1.3}$$

The Fourier Series of a function, if exists, is defined as

$$\mathcal{F}(f(t)) = \lim_{N\to\infty} \mathcal{F}_N(f(t)).$$

In order to extend the Fourier Series from $L^1(0,2\pi)$ functions into $L^2(0,2\pi)$, we first need to show that the set

$$\left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos(mt)}{\sqrt{\pi}}, \frac{\sin(nt)}{\sqrt{\pi}}\right\}, \tag{2.1.4}$$

is an orthonormal set for $m, n \in \mathbb{N}$. This can be shown using the relations

$$\int_0^{2\pi} \cos(mt)\sin(nt)\,dt = 0,$$

$$\int_0^{2\pi} \sin(mt)\sin(nt)\,dt = \begin{cases} \pi, & m = n \\ 0, & m \neq n \end{cases},$$

and

$$\int_0^{2\pi} \cos(mt)\cos(nt)\,dt = \begin{cases} \pi, & m = n \\ 0, & m \neq n \end{cases}.$$

The following theorem shows that the set (2.1.4) is an orthonormal basis for $L^2(0,2\pi)$ [5].

Theorem 1. *Let $f \in L^2(0,2\pi)$. Then*

$$\lim_{N\to\infty} \|\mathcal{F}_N(f(t)) - f(t)\|_2 = 0,$$

where $\|\mathcal{F}_N(f(t)) - f(t)\|_2 = \left(\int_0^{2\pi} |\mathcal{F}_N(f(t)) - f(t)|^2 dt\right)^{\frac{1}{2}}$. *Furthermore, the Parseval's identity is satisfied*

$$\int_0^{2\pi} |f(t)|^2 dt = 2\pi \sum_{n=-N}^{N} |c_n|^2 .$$

2.1.2. Fourier Transform (FT). The Fourier transform is the continuous version of the Fourier Series. Here the function is not necessarily periodic and it is defined on the whole real axis.

Definition 2. Let $f \in L^1(\mathbb{R})$. Then, the Fourier Transform of f is defined as

$$\mathcal{F}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-ikt} dt.$$

If $\mathcal{F} \in L^1(\mathbb{R})$, the function f can be recovered by

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{F}(k) e^{ikt} dk.$$

To properly define the Fourier Transform if $f \in L^2(\mathbb{R})$, but is not in $L^1(\mathbb{R})$, one must find a sequence of functions $f_N \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ which converges to f in the L^2 sense. Then the above definition can be used for f_N as N goes to infinity.

2.1.3. Fast Fourier Transform (FFT). In practice, the recorded signal is always in a discrete format, $\mathbf{f} = (f_0, \ldots, f_{N-1})^T$. Without loss of generality, one can assume that the data is recorded on a uniform mesh, $t_n = \frac{2\pi n}{N}$, for $n = 0, \ldots, N-1$. Using the trapezoidal rule, one can evaluate (2.1.3) to find the Fourier Series $\mathbf{F} = (F_0, \ldots, F_{N-1})^T$ in a discrete way:

$$F_k = \sum_{n=0}^{N-1} f_n e^{-\frac{i2\pi kn}{N}},$$

for $k = 0, \ldots, N-1$. The original signal can be recovered using

$$f_n = \frac{1}{N} \sum_{k=0}^{N-1} F_k e^{\frac{i2\pi kn}{N}}.$$

Now, the discrete Fourier transform can be be written in a matrix multiplication form. Take the following Vandermonde matrix by letting $\omega = e^{\frac{2\pi i}{N}}$.

$$\mathbf{M} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)^2} \end{pmatrix}$$

Hence, $\mathbf{F} = \mathbf{Mf}$. This matrix multiplication needs $O(N^2)$ operations. However, by rearranging the terms in odd and even members in $\mathbf{f}$, the algorithm can speed up to $O(N \log N)$ [15]. The latter is called the FFT method. In fact, the FFT method is a descendant of the divide and conquer algorithms.

2.1.4. Windowed Fourier Transform (WFT). Since Fourier Series suffers from the fact that there is no information about temporal instances in the Fourier domain, WFT was proposed by Gabor [41]. WFT analyzes a small part of the signal by multiplying it by a window function that is real, symmetric and non-zero:

$$\mathcal{F}_g(k, s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\, g(t - s)\, e^{-ikt} dt.$$

where g is the window function. While this provides information regarding the time-frequency nature of the signal, it does not completely find the time-frequency content of the signal. Depending on the length of the support of the window function, one can get different time-frequency resolutions due to the Heisenberg Uncertainty Principle. Furthermore, it is almost impossible to distinguish between the different time-frequency contents of the different components of a signal that is made up of many AM-FM signals.

2.1.5. Heisenberg Uncertainty Principle. A time signal f and its Fourier transform $\mathcal{F}$ cannot be simultaneously localized in a small domain of the time and frequency sheet [5]. This phenomenon tells us that no variation of the FT method,

including the WFT, can exactly state whether a certain frequency has happened at a certain time.

To enable better time-frequency resolution, the Wavelet Transform (WT) was introduced [16]. WT does not defy the Heisenberg Uncertainty Principle; however, it tries to present a decomposition that eliminates the principle's side effects as much as possible.

2.2. Wavelet Methods

Wavelet Transform was introduced to analyze a signal at different scales more accurately than WFT [16]. WT extracts the details of a signal at different time and frequency scales.

2.2.1. The Wavelet Transform. A wavelet is a complex function $\psi : \mathbb{R} \to \mathbb{C}$ satisfying the following conditions:

(1) $\psi \in L^2(\mathbb{R})$

(2) $C_\psi := 2\pi \int_{\mathbb{R}^*} \frac{|\mathcal{F}(\psi)(a)|^2}{|a|} da < \infty$,

where $\mathcal{F}(\psi)$ is the FT of ψ, and $\mathbb{R}^* = \mathbb{R} - \{0\}$. A wavelet dictionary $\mathcal{D}$ is the dilation and translations of the wavelet ψ:

$$\mathcal{D} = \{\psi_{a,b}(t), (a,b) \in \mathbb{R}^* \times \mathbb{R}\}, \tag{2.2.1}$$

where $\psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right)$. The parameters a, b are the scaling and translational parameters respectively. The wavelet transform of a signal is defined as the inner product of the signal f and the wavelet function:

$$\mathcal{W}f(a,b) := \langle f, \psi_{a,b} \rangle = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt. \tag{2.2.2}$$

One can recover the original function from the transform by

$$f = \frac{1}{C_\psi} \int_{\mathbb{R}^* \times \mathbb{R}} a^{-2} (\mathcal{W}f(a,b)) (\psi_{a,b}(t)) \, dadb. \tag{2.2.3}$$

The parameter b is a translational parameter that can spot the occurrence, in the time domain of a certain frequency, which is characterized by the scaling parameter a. Hence, the wavelet transform $\mathcal{W}f\,(a,b)$ conveys the information regarding the time-frequency content of a signal. Numerical approximation of Wavelet Analysis is called Multiresolution Approximations.

2.2.2. Multiresolution Approximations. A fast and numerically feasible approximation of a signal at different scales requires an orthonormal basis of $L^2\,(\mathbb{R})$ at different scales, in other words, an orthogonal projection on $\{V_j\}_{j\in\mathbb{Z}}$, as defined in Definition 3, so that the behavior of the signal at different scales can be captured. Here we present the properties of multiresolution spaces [41]:

Definition 3. The sequence of closed subspaces of $L^2\,(\mathbb{R})$, namely $\{V_j\}_{j\in\mathbb{Z}}$, is a multiresolution approximation, if the following properties are satisfied:

1- Translation Invariance: $f\,(t)\in V_j \Leftrightarrow f\,(t-2^j n)\in V_j$ for all $(j,n)\in\mathbb{Z}^2$.

2- Nesting: $V_{j+1}\subset V_j$ for all $j\in\mathbb{Z}$.

3- Scaling: $f\,(t)\in V_j \Leftrightarrow f\left(\frac{t}{2}\right)\in V_{j+1}$ for all $j\in\mathbb{Z}$.

4- Separation: $\bigcap\limits_{j=-\infty}^{\infty} V_j=\{\mathbf{0}\}$.

5- Density: $\overline{\bigcup\limits_{j=-\infty}^{\infty} V_j}=L^2\,(\mathbb{R})$.

6- Existence of Riesz Basis: There exists θ such that $\{\theta\,(t-n)\}_{n\in\mathbb{Z}}$ is a Riesz basis of V_0.

Using this definition, the following theorem constructs a family of orthonormal basis for V_j [41].

Theorem 2. *Let $\{V_j\}_{j\in\mathbb{Z}}$ be a multiresolution approximation and ϕ be the scaling function satisfying*

$$\mathcal{F}\,(\phi)\,(\xi)=\frac{\mathcal{F}\,(\theta)\,(\xi)}{\left(\sum_{k=-\infty}^{\infty}\left|\mathcal{F}\,(\theta)\,(\xi+2k\pi)\right|^2\right)^{\frac{1}{2}}}.$$

Then the family $\{\phi_{j,n}\}_{n\in\mathbb{Z}}$ *is an orthonormal basis of* $\{V_j\}_{j\in\mathbb{Z}}$. *Here,* $\phi_{j,n}$ *is defined as*

$$\phi_{j,n} = \frac{1}{\sqrt{2j}}\phi\left(\frac{t-2^j n}{2^j}\right).$$

To find the orthonormal basis of $L^2(\mathbb{R})$, one must first find the orthogonal complement of V_j, namely W_j, in V_{j-1},

$$V_{j-1} = V_j \oplus W_j.$$

This can be achieved using Theorems 3 and 4 [5, 41].

Theorem 3. *Let* ϕ *be a scaling function and let* ψ *be the function that satisfies*

$$\mathcal{F}(\psi)(\xi) = e^{\frac{i\xi}{2}}\overline{H\left(\frac{\xi}{2}+\pi\right)}\mathcal{F}(\phi)(\xi).$$

Here $H(\xi) = \frac{1}{\sqrt{2}}\sum_{n=-\infty}^{\infty} h_n e^{-in\xi}$, *with* $h_n = \langle \phi(t), \phi_{-1,n}(t)\rangle$. *Then the family* $\{\psi_{j,n}\}_{n\in\mathbb{Z}}$ *is an orthonormal basis of* $\{W_j\}_{j\in\mathbb{Z}}$, *with*

$$\psi_{j,n} = \frac{1}{\sqrt{2j}}\psi\left(\frac{t-2^j n}{2^j}\right).$$

Theorem 4. *The family*$\{\psi_{j,n}\}_{(j,n)\in\mathbb{Z}^2}$ *, is an orthonormal basis of* $L^2(\mathbb{R})$.

These theorems enabled the development of the Fast Wavelet Transform (FWT) algorithm. The FWT algorithm decomposes an approximation $f_j \in V_j$ into a coarser approximation $f_{j+1} \in V_{j+1}$ and the details $w_{j+1} \in W_{j+1}$ and is then applied to $f_{j+1} \in V_{j+1}$. In other words, for a $K-1$ step algorithm we have

$$f_j = f_{j+K} + \sum_{k=j+1}^{j+K} w_k.$$

For further details of the algorithm, see [41].

FWT is essentially the wavelet counterpart of FFT. It is a numerical implementation of the WT method. FWT gives a better time-frequency resolution than WFT. However, since the projection of the signal is on a predefined set of wavelet functions,

the extraction is not adaptive. In order to get a better time-frequency resolution, the intuition is to project the signal on an adaptive basis that is found from the signal itself. Empirical Mode Decomposition and the more recent Ensemble Empirical Mode Decomposition were developed to do just that.

2.3. Empirical Mode Decomposition (EMD) Methods

Empirical Mode Decomposition (EMD) methods were originally proposed to overcome the problem of predetermined bases which was hampering a proper sparse decomposition in the WT and FT methods. EMD [32], which tries to find the basis of decomposition adaptively from the original signal in an empirical way, has gained much acceptance in the engineering and scientific community [30, 31].

2.3.1. EMD Algorithm. The EMD method uses Intrinsic Mode Function (IMF) [32] as constructing blocks of signals. IMFs are defined slightly differently in different methods. Here, we present an EMD definition of an IMF.

Definition 4. An IMF is a function that satisfies the following two conditions:

- The number of extrema and the number of zero crossings are equal or differ at most by one on the whole data set
- The mean value of the envelopes defined by the local maxima and the local minima is zero at any point.

EMD decomposes a signal into an extracted IMF and a residue through a sifting process. The extracted IMF is then post-processed to find the Instantaneous Frequency (IF) either through Hilbert transform or another sifting process [32, 33]. To summarize, for a signal $x(t)$, the decomposition looks like

$$x(t) = \sum_{k=1}^{n} c_k(t) + r_n(t),$$

where c_k's are the IMFs extracted from the signal and $r_n(t)$ is the residue. The residue is not itself an IMF but is more like a trend that is either completely increasing, or

Algorithm 1 EMD Sifting Algorithm

- 1: $k = 1$
- 2: Find all local extrema of $x(t)$
- 3: Interpolate the minima (using cubic spline) to get the lower envelope $e_{min}(t)$
- 4: Interpolate the maxima (using cubic spline) to get the upper envelope $e_{max}(t)$
- 5: Compute the mean $m(t) = \frac{e_{max}(t)+e_{min}(t)}{2}$
- 6: Extract the detail $d_k(t) = x(t) - m(t)$
- 7: if $d_k(t)$ satisfies the definition of an IMF then
- 8: return $c(t) = d_k(t)$
- 9: else
- 10: $x(t) = d_k(t)$
- 11: $k = k + 1$
- 12: go to 2.
- 13: end if

decreasing, or has only one local minimum or maximum. The original EMD sifting algorithm, which is explained in Algorithm 1 [23], extracts only one IMF from the signal. Hence, when one IMF $c(t)$ is extracted from the signal $x(t)$, then $x(t) - c(t)$ is considered to be a new signal (if it is not already a residual) and the sifting process is now performed on this new signal. This procedure is done repeatedly until all IMFs are extracted.

There is a problem with the sifting algorithm: the sifting process can inadvertently eliminate physical information about the extractions by over-smoothing. In order to prevent such an error, a stopping criterion is needed [32]. Taking a parameter like

$$SD = \sum_{i=1}^{N} \frac{|d_k(t_i) - d_{k+1}(t_i)|^2}{(d_k(t_i))^2},$$

it is possible to define an empirical stopping criterion. Here, it is assumed that the signal is sampled on discrete points $t_i, i = 1, \ldots, N$. Usually, if SD is less than 0.2 or 0.3, the sifting process must stop. Unfortunately, this is another empirical nature of the algorithm.

Once the first IMF is extracted, it is subtracted from the whole signal and the result is treated as a new signal and the sifting algorithm is then again applied to the

residue. The extractions finally stop when the residual is too small or it is monotonic. Once all the IMFs are extracted, the IF can be found by the Hilbert transform or another sifting process on the IMF [33]:

First, the Hilbert transform of the j-th IMF c_j is defined as

$$h_j(t) = \frac{1}{\pi} \fint_{-\infty}^{\infty} \frac{c_j(\tau)}{t-\tau} d\tau.$$

Next, the *analytic signal* of the Hilbert transform is defined as

$$z_j(t) = c_j(t) + ih_j(t) = a_j(t)\, e^{i\theta_j(t)}.$$

At the end, the IF is defined as the time derivative of $\theta_j(t)$. In other words, $\omega_j(t) = \frac{d\theta_j(t)}{dt}$. The problem with this approach is that the found IF is based on a certain definition that is derived from an analytic signal.

2.3.2. Ensemble Empirical Mode Decomposition (EEMD) Algorithm. EMD is highly sensitive to noise. In order to fix this shortcoming, Wu and Huang [56] proposed EEMD in 2009. EEMD minimizes noise effects by adding white noise to the signal. The method then follows with finding the IMF, repeating the same procedure with many other realizations of the white noise, and finally take the average of these IMFs as the corresponding IMF. The number of times that the white noise is added is called the ensemble number E. The EEMD algorithm is detailed in Algorithm 2, [23].

The total number of IMF extractions should remain the same throughout each time that the noise is added to the signal $x_i(t)$, which is a shortcoming of the EEMD. The proposed number is usually $\lfloor \log_2 N \rfloor - 1$ for a signal of length N [56]. This is again another empirical value beside the empirical nature of the algorithm. Furthermore, the extracted functions are not necessarily IMFs [56]. While the EMD and EEMD methods have attracted a lot of attention, they lack a theoretical basis. The STFR methods were introduced to fill the empirical gap. In the next section we start with the first rigorous STFR method.

Algorithm 2 EEMD Algorithm

- 1: for $i = 1 \longmapsto E$ do
- 2: Let $w_i(t)$ be a white noise series
- 3: Compute $x_i(t) = x(t) + w_i(t)$
- 4: Perform the decomposition of $x_i(t)$ using the EMD method:

$$x_i(t) = \sum_{j=1}^{n} c_{j,i}(t) + r_{n,i}(t),$$

- 5: end for
- 6: Compute the (ensemble) mean IMFs

$$c_j(t) = \frac{1}{E} \sum_{i=1}^{E} c_{j,i}(t),$$

for $j = 1, \ldots, n$.

2.4. STFR Methods

Although STFR methods were inspired by the EMD and EEMD methods [32, 56], they make use of techniques from compressive sensing theory [10, 12] and matching pursuit [42]. Compressive Sensing (CS) and Matching Pursuit (MP) methods try to find a sparse reconstruction of an observed signal from a predefined finite large dictionary. STFR, on the other hand tries to find a sparse reconstruction of an observed signal from a predefined *infinite* large dictionary. STFR methods use two major steps: they first construct a highly redundant dictionary of all IMFs, namely D, then find the sparsest decomposition by solving a non-linear and non-convex optimization problem (2.4.2), an L_0 minimization problem. Solving it is fundamentally difficult. In the coming parts, we explain possible approximations by which one can find a close solution of problem (2.4.2).

In order to explain the STFR methods, we need some preliminary definitions. The set of the collection of IMFs constitute a dictionary. However, before defining the dictionary of IMFs, we need another definition.

Definition 5. The envelope functions set $V(\theta)$ is

$$V(\theta) = span\left\{1, \cos\frac{\theta}{\lambda}, \sin\frac{\theta}{\lambda} \mid \lambda \geq \lambda_0\right\},$$

where λ_0 is called the level of smoothness.

A common value for λ_0 is 2 so that the frequencies of the envelope terms are always less than or equal to $\frac{1}{2}$, in θ-coordinate. This states that the envelope is not as oscillatory as $\cos\theta$.

Definition 6. The IMF dictionary D is defined as

$$D = \left\{ a(t)\cos\theta(t) \mid a(t) \in V(\theta(t)), a(t) > 0, \frac{d\theta(t)}{dt} > 0, \frac{d\theta(t)}{dt} \in C \right\}. \tag{2.4.1}$$

The goal of the STFR methods is to find the sparsest decomposition among all possible IMFs in a dictionary. In mathematical terms, the sparsest decomposition of the signal is to be found by solving a nonlinear optimization problem

(2.4.2)

$$\begin{aligned} &Minimize && M \\ &Subject\ to: && s(t) = \sum_{i=1}^{M} a_i(t)\cos\theta_i(t), \quad a_i(t)\cos\theta_i(t) \in D, \quad i = 1, ..., M. \end{aligned}$$

The assumption behind this optimization problem is that the nature of the signal is nothing but an IS. This problem is an L_0 minimization. Like other optimization problems, the constraint $s(t) = \sum_{i=1}^{M} a_i(t)\cos\theta_i(t)$ can be relaxed into an inequality in L_2-norm if noise (or approximation error) is present in observation.

The optimization problem (2.4.2) is NP hard. Hou and Shi have proposed a number of algorithms for approximating this problem [25, 24, 26] by decomposing the signal into two parts, a mean a_0 and a modulated oscillatory part, namely the IMF, $a_1\cos\theta$:

$$s(t) = a_0(t) + a_1(t)\cos\theta(t),$$

where θ, a_1, and a_0 are the unknown phase function, envelope, and mean respectively. This sequential decomposition is nothing but a matching pursuit [42] to find one of the IMFs, $a_1\cos\theta$. After this has been extracted, the residue (or the mean) a_0 is treated as a new signal and the same procedure is repeated until the residue is smaller than

a preset value. This part of the STFR methods is essentially the same as the one we observed in the EMD and EEMD methods.

2.4.1. Total Variation (TV) Method. To find a good acceptable approximation, Hou and Shi [24] proposed an adaptive STFR method based on total variation minimization, the TV STFR method. They used TV to impose a smoothness condition on the mean $a_0(t)$ and envelope $a_1(t)$ functions. It is well known that minimizing the total variation $TV(g) = \int_a^b |g'(x)|dx$ would generate the "stair case." As a result, Hou and Shi proposed the TV^3 minimization[1] of the form

$$\begin{aligned} &Minimize && TV^3(a_0) + TV^3(a_1) \\ &Subject\ to: && s(t) = a_0(t) + a_1(t)\cos\theta_1(t), \quad \theta_1'(t) \geqslant 0, \end{aligned} \tag{2.4.3}$$

as an approximation to the original problem (2.4.2). This enforces a higher order regularity of the local median and envelope. Minimizing the third-order total variation of any function tends to produce a piecewise constant approximation to the third order derivative of the function. Thus, a TV^3 minimization tends to produce a cubic spline approximation for $a_0(t)$ and $a_1(t)$. In this sense, this method is similar to that of the EMD method.

This problem is solved in an iterative manner. First, an initial guess on θ_1 is set. Then based on that, the value of a_1 is approximated and the phase θ_1 is again updated. This procedure is repeated until there is no progress in updating the phase function. Here, the definition of the dictionary is not used explicitly. The smoothness requirements of the mean a_0 and the envelope a_1 are enforced through a TV^3 minimization. This approach is an exponential step compared to the EMD method and fixed-basis methods like FT and WT methods. Unfortunately, the algorithm is not stable to noise.

[1] $TV^n(g) = \int_a^b |g^{(n+1)}(x)|dx = \left\|g^{(n+1)}\right\|_1$

2.4.2. Periodic Fourier-based Sparse Time-Frequency Method. While TV STFR performs well for signals with no noise perturbation, noise is an indispensable part of almost all signals. To address this problem, Hou and Shi, proposed the Periodic Fourier-based STFR method [25], which performs well on periodic data even in the presence of noise.

2.4.2.1. *Theory and Algorithm.* Without loss of generality, assume that the signal s is sampled over a uniform grid $t_j = \frac{j}{N}$ for $j = 0, 1, \ldots, N-1$ and even N. The Periodic STFR algorithm is explained in Algorithm 3. The Periodic STFR algorithm works in the following way: an initial guess is proposed on the phase function θ_0. In the first step, the whole signal is mapped to the phase function space and the Fourier transform of the signal is then found in that space. At this step, the possible candidates for the envelope functions $a^k_{\theta^n_k}$, $b^k_{\theta^n_k}$ are found by (2.4.7) and (2.4.8). Later, in Step 5, the new phase function is updated and the algorithm begins a new iteration.

The cutoff function in the algorithm is

$$\chi(\omega) = \begin{cases} 1, & -\frac{1}{2} < \omega < \frac{1}{2}, \\ 0, & otherwise. \end{cases} \tag{2.4.12}$$

2.4.2.2. *Technical details of the algorithm's updater, envelope extraction, and underlying theory.* The algorithm's updater (2.4.9) assumes θ^n_k is the n^{th} guess on the phase function. This guess needs to be updated by a perturbation $\triangle\theta$. In other words, $\theta^{n+1}_k = \theta^n_k + \triangle\theta$. Having this, one can state

$$\begin{aligned} A^n_k \cos(\theta^n_k + \triangle\theta) &= A^n_k \cos(\triangle\theta)\cos(\theta^n_k) - A^n_k \sin(\triangle\theta)\sin(\theta^n_k) \\ &= a^k_n \cos(\theta^n_k) + b^k_n \sin(\theta^n_k), \end{aligned}$$

with

$$a^k_n = A^n_k \cos(\triangle\theta),$$

$$b^k_n = -A^n_k \sin(\triangle\theta).$$

Algorithm 3 Periodic STFR Algorithm

- $k=1$, $r^{k-1}=s$
- Step 0: $n=0$, $\theta_k^n=\theta_0$
- Step 1: Interpolate r^{k-1} from the t-coordinate into a uniform grid in θ-coordinate. In other words, for $j=0,\ldots,N-1$:

$$\left(r_{\theta_k^n}^{k-1}\right)_j = Interpolate\left(t, r^{k-1}, \left(\theta_k^n\right)_j\right) \tag{2.4.4}$$

- Step 2: Apply the Fourier transform to $r_{\theta_k^n}^{k-1}$ for $\left(\theta_k^n\right)_j = 2\pi L_{\theta_k^n}\frac{j}{N}$:

$$\mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_\omega = \frac{1}{N}\sum_{j=1}^{N}\left(r_{\theta_k^n}^{k-1}\right)_j e^{-i\omega\frac{\left(\theta_k^n\right)_j}{L_{\theta_k^n}}}, \omega = -\frac{N}{2}+1,\ldots,\frac{N}{2} \tag{2.4.5}$$

- Step 3: For the Fourier inverse defined as

$$\left(\mathcal{F}_{\theta_k^n}^{-1}\left(\mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)\right)\right)_j = \frac{1}{N}\sum_{\omega=-\frac{N}{2}+1}^{\frac{N}{2}}\mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_\omega e^{i\omega\frac{\left(\theta_k^n\right)_j}{L_{\theta_k^n}}}, j=0,\ldots,N-1, \tag{2.4.6}$$

apply a cutoff function χ to $\mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)$ to calculate the envelopes

$$a_{\theta_k^n}^k = \mathcal{F}_{\theta_k^n}^{-1}\left[\left(\mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_{\omega+L_{\theta_k^n}} + \mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_{\omega-L_{\theta_k^n}}\right)\chi\left(\frac{\omega}{L_{\theta_k^n}}\right)\right], \tag{2.4.7}$$

$$b_{\theta_k^n}^k = \mathcal{F}_{\theta_k^n}^{-1}\left[-i\left(\mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_{\omega+L_{\theta^n}} - \mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_{\omega-L_{\theta_k^n}}\right)\chi\left(\frac{\omega}{L_{\theta_k^n}}\right)\right]. \tag{2.4.8}$$

- Step 4: Interpolate the calculated envelopes background into the t-coordinate for $j=0,\ldots,N-1$, by $\left(a_n^k\right)_j = Interpolate\left(\theta_k^n, a_{\theta_k^n}^k, t_j\right)$ and $\left(b_n^k\right)_j = Interpolate\left(\theta_k^n, b_{\theta_k^n}^k, t_j\right)$.
- Step 5: Update the phase function

$$\triangle\theta' = P_{V_{M_0}}\left(\frac{d}{dt}\arctan\left(\frac{b_n^k}{a_n^k}\right)\right), \triangle\theta\left(t\right) = \int_0^t \triangle\theta'\left(s\right)ds, \tag{2.4.9}$$

$$\theta_k^{n+1} = \theta_k^n + \beta\triangle\theta. \tag{2.4.10}$$

Choose $\beta\in[0,1]$ such that the phase function is always strictly increasing:

$$\beta = \max\left\{\alpha\in[0,1] \mid \frac{d}{dt}\left(\theta_k^n+\alpha\triangle\theta\right)>0\right\}. \tag{2.4.11}$$

The projection, $P_{V_{M_0}}$, into $V_{M_0} = span\left\{e^{i2\pi kt}, |k|\le M_0, k\in\mathbb{Z}\right\}$ is a smoothing step. M_0 is fixed.

- Step 6: If $\left\|\theta_k^{n+1}-\theta_k^n\right\| \le \epsilon_I$, then $IMF_k = a_n^k\cos\theta_k^n + b_n^k\sin\theta_k^n$, $n\leftarrow 0$, $r^k = r^{k-1}-IMF_k$,$k\leftarrow k+1$. Else, $n\leftarrow n+1$, goto Step 1.
- Step 7: If $\left\|r^k\right\|\le\epsilon_{II}$, stop. Else, goto Step 0.

Hence, the perturbation $\triangle\theta$ is

$$\triangle\theta = -\arctan\left(\frac{b_n^k}{a_n^k}\right).$$

Since arctan (.) is not good function to handle numerically due to jumps, the derivative of this function is used. Also, in order to denoise the updater (2.4.9), a projection on $P_{V_{M_0}}$ is used.

In this algorithm, the envelopes $a_{\theta_k^n}^k$, $b_{\theta_k^n}^k$ are extracted as shown in (2.4.7) and (2.4.8). We need the following lemma to be able to depict the justification.

Lemma 1. *For a function $f \in L^1(\mathbb{R})$, we have*

$$\mathcal{F}[f(t)\cos(t)](k) = \frac{1}{2}\left(\mathcal{F}[f](k-1) + \mathcal{F}[f](k+1)\right),$$

$$\mathcal{F}[f(t)\sin(t)](k) = \frac{1}{2i}\left(\mathcal{F}[f](k-1) - \mathcal{F}[f](k+1)\right).$$

PROOF. The proof is just direct calculations with $\cos(t) = \frac{e^{it}+e^{-it}}{2}$ and $\sin(t) = \frac{e^{it}-e^{-it}}{2i}$. □

An IMF is concentrated around frequency 1 in its corresponding θ-coordinate, if there is no mode mixture. In other words, if the signal is s, then one of the IMFs c in one of the θ-coordinates approximately looks like

$$\mathcal{F}_\theta[c](k) \approx \mathcal{F}_\theta[s](k)\,\chi(k-1). \tag{2.4.13}$$

Here, $\chi(k)$ is a low-pass filter centered at 0. We know that the IMF has a representation like

$$c(\theta) = a(\theta)\cos(\theta) + b(\theta)\sin(\theta)$$

in the sought θ-coordinate. Using Lemma 1 we have

$$\mathcal{F}_\theta[c](k) = \frac{1}{2}\left(\mathcal{F}_\theta[a](k-1) + \mathcal{F}_\theta[a](k+1)\right) + \frac{1}{2i}\left(\mathcal{F}_\theta[b](k-1) - \mathcal{F}_\theta[b](k+1)\right). \tag{2.4.14}$$

Since we have assumed that $a(\theta), b(\theta)$ are in $V(\theta)$, if we multiply (2.4.14) by $\chi(k-1)$ we have

$$\mathcal{F}_\theta[c](k)\chi(k-1) = \frac{1}{2}\left(\mathcal{F}_\theta[a](k-1) - i\mathcal{F}_\theta[b](k-1)\right).$$

Using this and approximation (2.4.13) we have

$$\mathcal{F}_\theta[s](k)\chi(k-1) \approx \frac{1}{2}\left(\mathcal{F}_\theta[a](k-1) - i\mathcal{F}_\theta[b](k-1)\right),$$

which is equivalent to

$$\mathcal{F}_\theta[s](k+1)\chi(k) \approx \frac{1}{2}\left(\mathcal{F}_\theta[a](k) - i\mathcal{F}_\theta[b](k)\right). \tag{2.4.15}$$

Now, if we multiply (2.4.14) by $\chi(k+1)$, and follow the same steps as we just took, we have

$$\mathcal{F}_\theta[s](k-1)\chi(k) \approx \frac{1}{2}\left(\mathcal{F}_\theta[a](k) + i\mathcal{F}_\theta[b](k)\right). \tag{2.4.16}$$

Combining (2.4.15) and (2.4.16) we have

$$\mathcal{F}_\theta[a](k) \approx \left(\mathcal{F}_\theta[s](k+1) + \mathcal{F}_\theta[s](k-1)\right)\chi(k),$$

$$\mathcal{F}_\theta[b](k) \approx -i\left(\mathcal{F}_\theta[s](k+1) - \mathcal{F}_\theta[s](k-1)\right)\chi(k).$$

Hence, the Step 3 of the algorithm and equations (2.4.7) and (2.4.8) are clearly explained.

2.4.2.3. *Numerical Examples.* In this part, we show a few examples on the capabilities of this algorithm. We will show the algorithm is able to extract the IMFs from an IS even in the presence of heavy noise perturbation.

Example 1. For a tested signal (Figure 2.1) with constituent IMFs that are constant envelope IMFs with IFs shown in Figure 2.2, we used the periodic algorithm introduced in this section and extracted IMFs and IFs (shown in red in Figures 2.3-2.6). The method was accurate with almost no error in the extraction of the IMFs and their corresponding IFs. In other words, in the absence of noise, the extractions are

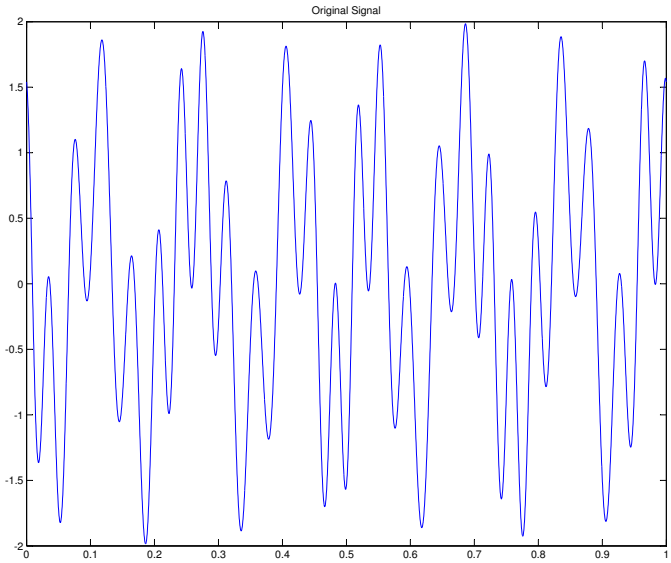

FIGURE 2.1. The Original Signal: The horizontal axis is the time variable and the vertical axis shows the signal itself.

almost accurate since the IFs are completely separate in the frequency domain, see Figure 2.2.

Example 2. We added white noise with amplitude one to the same signal that was used in Example 1 (original signal is shown in Figure 2.1 and the noisy signal in Figure 2.7). Using the periodic algorithm introduced in this section, we extracted IMFs and IFs (shown in red in Figures 2.8-2.11). Even in the presence of noise, the extractions remain accurate. While the second IF is not captured properly (Figure 2.11) the IMF corresponding to the second IMF is recovered from the original noisy signal with good accuracy.

Examples 1 and 2 show that the Periodic STFR method has good accuracy, even when noise is present. While this algorithm works well for periodic data, for non-periodic data, the algorithm would not extract the IMFs properly as the Gibbs phenomenon pollutes the extraction. In the next chapter, we propose another algorithm

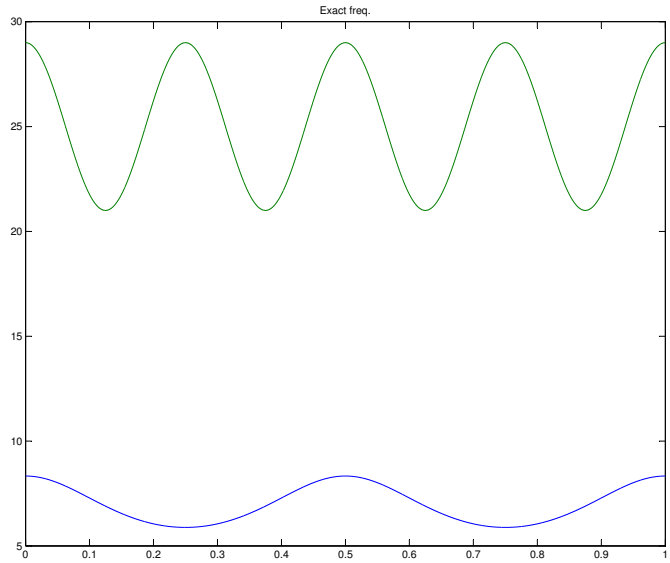

FIGURE 2.2. The Instantaneous Frequencies of the Signal Shown in Figure (2.1): The horizontal axis is the time variable and the vertical axis shows two different IFs of the constituent IMFs of the signal.

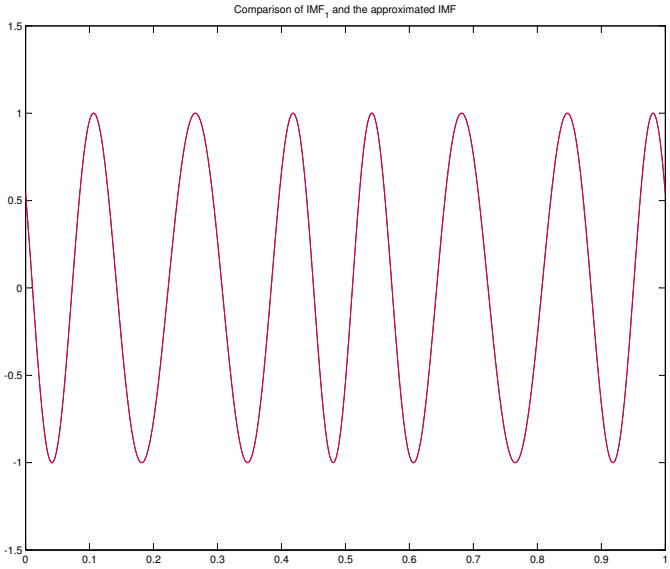

FIGURE 2.3. First IMF Extraction: The horizontal axis is the time variable. The extracted IMF is in red and the original IMF is in blue. The extraction is almost with no error.

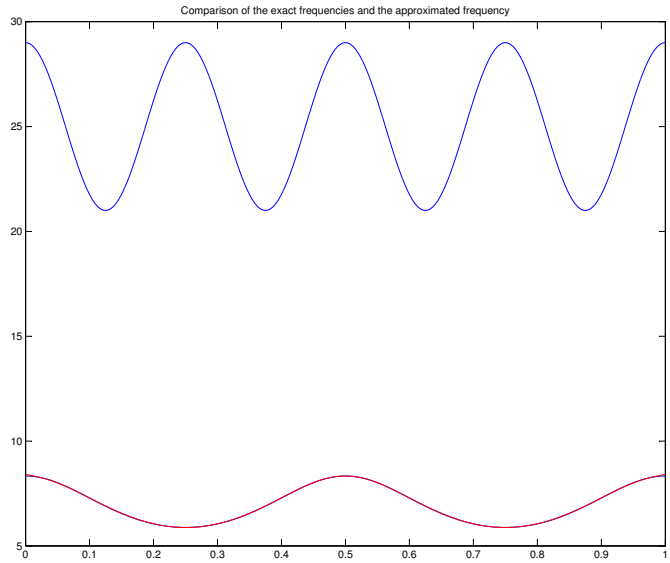

FIGURE 2.4. First IMF IF: The horizontal axis is the time variable. The extracted IF of the first IMF is in red. It completely overlaps the lowest IF content of the signal.

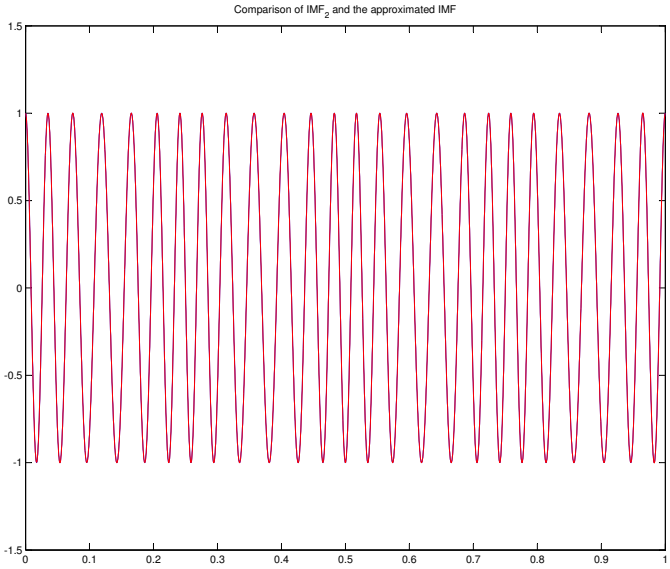

FIGURE 2.5. Second IMF Extraction: The horizontal axis is the time variable. The extracted IMF is in red and the original IMF is in blue. The extraction is almost with no error.

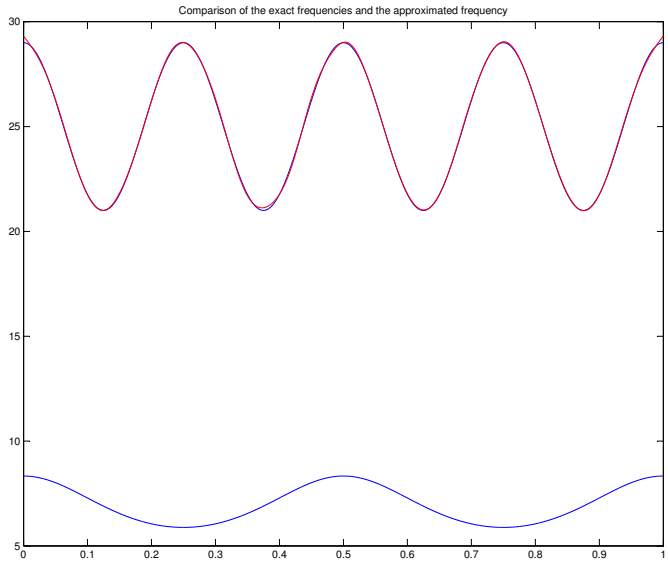

FIGURE 2.6. Second IMF IF: The horizontal axis is the time variable. The extracted IF of the second IMF is in red. It completely overlaps the highest IF content of the signal except near the boundaries.

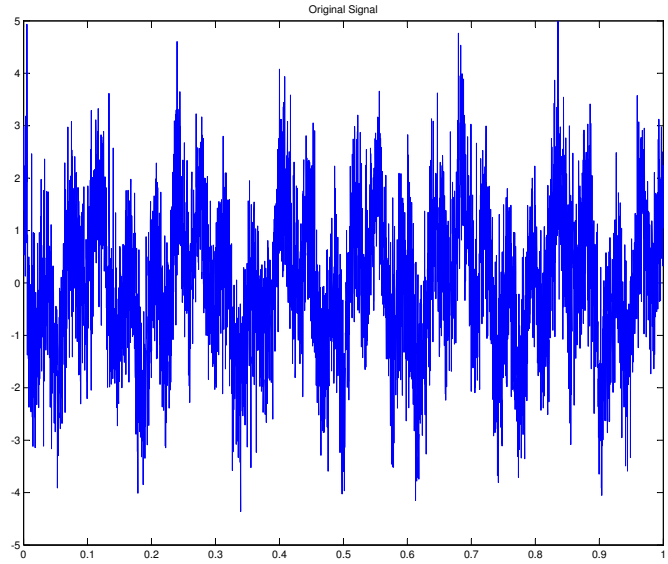

FIGURE 2.7. Original Noisy Signal: The horizontal axis is the time variable and the vertical axis shows the signal itself.

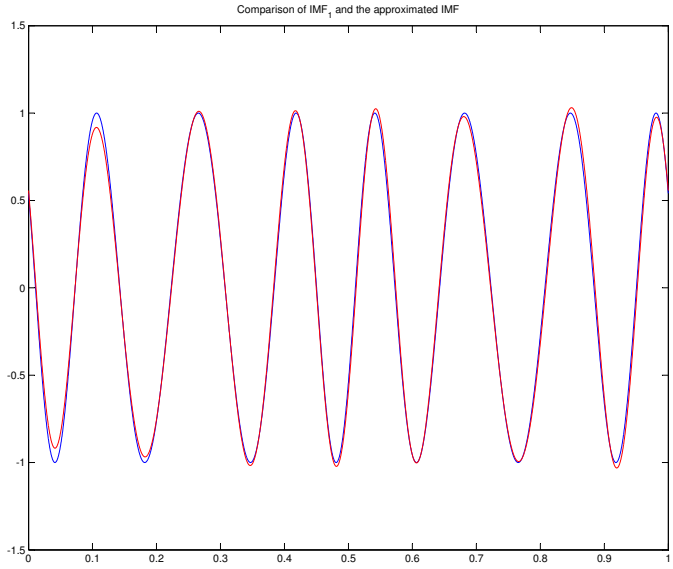

FIGURE 2.8. First IMF Extraction: The horizontal axis is the time variable. The extracted IMF is in red and the original IMF is in blue. The extraction has some minor error.

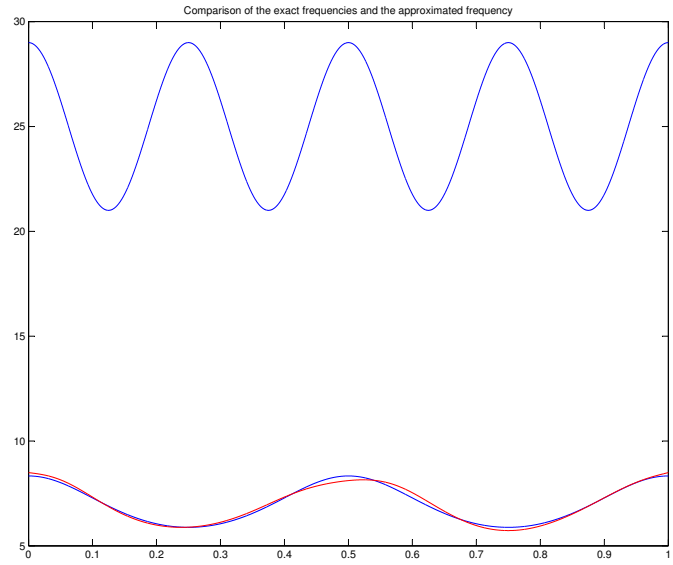

FIGURE 2.9. First IMF IF: The horizontal axis is the time variable. The extracted IF of the first IMF is in red. It does not overlap the original IF of the signal. However, it is capturing its trend properly.

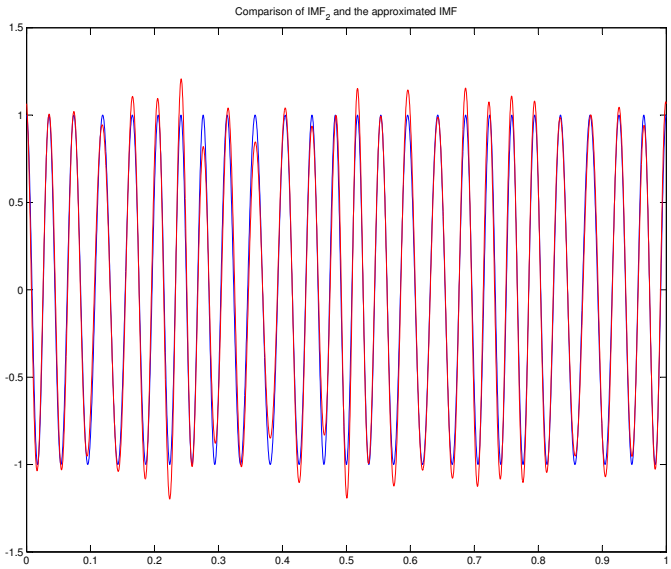

FIGURE 2.10. Second IMF Extraction: The horizontal axis is the time variable. The extracted IMF is in red and the original IMF is in blue. The extraction is acceptable even in the presence of noise.

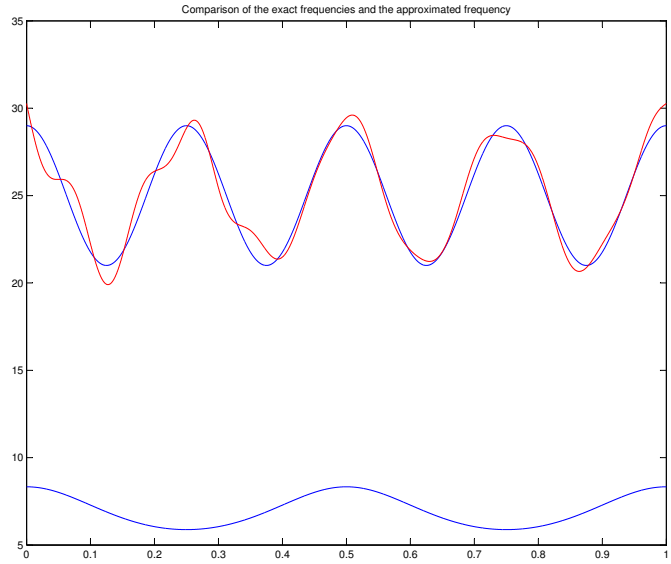

FIGURE 2.11. Second IMF IF: The horizontal axis is the time variable. The extracted IF of the second IMF is in red. It does not completely cover the whole trend of the IF due to the presence of noise perturbation.

where this problem is addressed properly. In the next chapter, we introduce the non-periodic STFR method by which one can even handle non-periodic IS signals in the presence of noise perturbation.

CHAPTER 3

Non-Periodic Sparse Time-Frequency Method

TV STFR is not accurate for noisy signals and Periodic STFR is not suitable for non-periodic signals (in particular it cannot accurately extract the trend of a non-periodic signal). In this chapter, we propose a non-periodic STFR method that overcomes the limitations of previous STFR methods. Non-periodic STFR has a slower speed of convergence than Periodic STFR since it uses a L_1-norm regularized with L_2-norm optimization rather than FFT for each iteration. Nevertheless, it can successfully be applied to non-periodic signals and also to signals polluted with noise.

3.1. Theory and Algorithm

Non-periodic STFR uses the same approach as Periodic STFR: we try to extract the sparsest representation of IMFs from an IS. Essentially, we are looking for an appropriate approximation to the problem,

(3.1.1)

$$\begin{aligned} &Minimize && M \\ &Subject\ to: && s(t) = \textstyle\sum_{i=1}^{M} a_i(t)\cos\theta_i(t), \quad a_i(t)\cos\theta_i(t) \in D, \quad i = 1, ..., M. \end{aligned}$$

In the non-periodic STFR algorithm, we assume that the envelopes of the IMFs have a sparse structure in their respective dictionary. This assumption is not far from reality since we are using an infinitely large dictionary when we try to extract IMFs. This assumption can be formulated in the following way:

(3.1.2)

$$\begin{aligned} &\underset{a,b,\theta}{Minimize} && \delta\left(\|a(\lambda)\|_1 + \|b(\lambda)\|_1\right) + \|s(t) - \mathcal{A}(t)\cos\theta(t)\|_2^2 \\ &Subject\ to: && \mathcal{A}(t) = \int_2^{\infty}\left(a(\lambda)\cos\tfrac{\theta(t)}{\lambda} + b(\lambda)\sin\tfrac{\theta(t)}{\lambda}\right)d\lambda, \\ & && \tfrac{d\theta(t)}{dt} > 0. \end{aligned}$$

In this formulation, the envelope $\mathcal{A}(t)$ of the IMF $\mathcal{A}(t)\cos\theta(t)$ is assumed to have a sparse structure that is captured by $\|a(t)\|_1 + \|b(t)\|_1$. Here, δ is the regularization parameter to enforce the sparsity. In this formulation, the dictionary is defined explicitly by

$$\mathcal{A}(t) = \int_2^\infty \left(a(\lambda)\cos\frac{\theta(t)}{\lambda} + b(\lambda)\sin\frac{\theta(t)}{\lambda} \right) d\lambda,$$

such that $a, b \in C^1$. One can understand this similarity by referring to the definition of the dictionary:

$$D = \left\{ a(t)\cos\theta(t) \mid a(t) \in V(\theta(t)), a(t) > 0, \frac{d\theta(t)}{dt} > 0 \right\},$$

where

$$V(\theta) = span\left\{ 1, \cos\frac{\theta}{\lambda}, \sin\frac{\theta}{\lambda} \mid \lambda \geq \lambda_0 \right\}.$$

The optimization in (3.1.2) is in a continuous format. In order to solve such an optimization problem using a computer algorithm, we need a discrete format of the problem.

3.1.1. Discrete Formulation. Here, for the sake of simplicity in implementation, we set $\theta(0) = 0$. The following vectors and matrices can help us to convert it into a discrete format suitable for implementation. We start with the semi-discrete version of the envelop function $A(\theta)$:

$$A(\theta) = \bar{a}^c + \sum_l a^c_{k_l}\cos\frac{\theta}{k_l} + b^c_{k_l}\sin\frac{\theta}{k_l}, \tag{3.1.3}$$

$$\{k_l\}_{l=1}^m \subset [2, \infty). \tag{3.1.4}$$

This formulation states that the envelope of a to-be-extracted IMF has a representation derived from dictionary D. The same formulation stands for $B(\theta)$:

$$B(\theta) = \bar{a}^s + \sum_l a^s_{k_l}\cos\frac{\theta}{k_l} + b^s_{k_l}\sin\frac{\theta}{k_l}. \tag{3.1.5}$$

In fact $A(\theta)$ and $B(\theta)$ are numerical counterparts of the envelope of the IMF $\mathcal{A}(t)$. In other words, we assume that the IMF looks like $A(\theta)\cos\theta + B(\theta)\sin\theta$, when $\theta(0) = 0$ is enforced. The definition of the IMF dictionary is used explicitly when having $\{k_l\}_{l=1}^m \subset [2,\infty)$ side by side with equation (3.1.3). In practice, we pick $O(2^{10})$ uniformly distributed points between 2 and 10, as instances for the values of k_l. So far, the semi-discrete version of (3.1.2), using the numerical counterparts of $\mathcal{A}(t)$ (see (3.1.3)), is

$$\begin{aligned}
&\delta\left(|\bar{a}^c| + \sum_{l=1}^m \left(\left|a_{k_l}^c\right| + \left|b_{k_l}^c\right|\right)\right) \\
&+\delta\left(|\bar{a}^s| + \sum_{l=1}^m \left(\left|a_{k_l}^s\right| + \left|b_{k_l}^s\right|\right)\right) \\
&\qquad + ||s(\theta) \quad - \left(\bar{a}^c + \sum_{l=1}^m \left(a_{k_l}^c \cos\tfrac{\theta}{k_l} + b_{k_l}^c \sin\tfrac{\theta}{k_l}\right)\right)\cos\theta \\
&\qquad\qquad - \left(\bar{a}^s + \sum_{l=1}^m \left(a_{k_l}^s \cos\tfrac{\theta}{k_l} + b_{k_l}^s \sin\tfrac{\theta}{k_l}\right)\right)\sin\theta||_2^2.
\end{aligned}$$

In this formula $\bar{a}^c, a_{k_l}^c, b_{k_l}^c, \bar{a}^s, a_{k_l}^s, b_{k_l}^s \in \mathbb{R}$. Now, in the discrete domain, for $i = 1,\ldots,N$, define

$$\begin{aligned}
\boldsymbol{\theta} &= \{\theta_i\} \in \mathbb{R}^N, \\
\mathbf{s} &= \{s(t_i)\} \in \mathbb{R}^N, \\
\mathbf{C} &= \left\{\cos\tfrac{\theta_i}{k_l}\right\} = \{C_{il}\} \in \mathbb{R}^{N\times m}, \\
\mathbf{S} &= \left\{\sin\tfrac{\theta_i}{k_l}\right\} = \{S_{il}\} \in \mathbb{R}^{N\times m},
\end{aligned} \tag{3.1.6}$$

where $\theta_i = \theta(t_i)$. Next, define the envelope vectors $\mathbf{A}, \mathbf{B}$ as

$$\mathbf{A} = [\mathbf{C},\mathbf{S},\mathbf{1}]\begin{bmatrix}\mathbf{a}^c\\ \mathbf{b}^c\\ \bar{\mathbf{a}}^c\end{bmatrix}, \quad \mathbf{B} = [\mathbf{C},\mathbf{S},\mathbf{1}]\begin{bmatrix}\mathbf{a}^s\\ \mathbf{b}^s\\ \bar{\mathbf{a}}^s\end{bmatrix}. \tag{3.1.7}$$

Other matrices needed to describe the algorithm are

$$\begin{aligned}
\Psi &= [\mathbf{C},\mathbf{S},\mathbf{1}], & \boldsymbol{Sin\theta} &= diag(\sin\boldsymbol{\theta}), \\
\mathbf{x} &= \begin{bmatrix}\mathbf{a}^c & \mathbf{b}^c & \bar{\mathbf{a}}^c & \mathbf{a}^s & \mathbf{b}^s & \bar{\mathbf{a}}^s\end{bmatrix}^T, & \boldsymbol{Cos\theta} &= diag(\cos\boldsymbol{\theta}),
\end{aligned} \tag{3.1.8}$$

Algorithm 4 Non-periodic STFR Algorithm

- $k = 1$, $\mathbf{r}^{k-1} = \mathbf{s}$
- Step 0: $n = 0$, $\boldsymbol{\theta}_k^{(n)} = \boldsymbol{\theta}_0$
- Step 1: Find $\mathbf{x}_k^{(n)} = \left(\left[\begin{array}{cccccc} \mathbf{a}^c & \mathbf{b}^c & \bar{\mathbf{a}}^c & \mathbf{a}^s & \mathbf{b}^s & \bar{\mathbf{a}}^s \end{array} \right]^T \right)_k^{(n)}$ by solving the convex minimization problem:

$$Minimize \;\; \delta \|\mathbf{x}_k^{(n)}\|_1 + \|\mathbf{s} - \mathbf{H}_k^{(n)} \mathbf{x}_k^{(n)}\|_2^2 \tag{3.1.10}$$

- Step 2: Calculate

$$\mathbf{A}_k^n = [\mathbf{C}, \mathbf{S}, \mathbf{1}]_k^{(n)} \left(\left[\begin{array}{c} \mathbf{a}^c \\ \mathbf{b}^c \\ \bar{\mathbf{a}}^c \end{array} \right] \right)_k^{(n)}, \mathbf{B}_k^n = [\mathbf{C}, \mathbf{S}, \mathbf{1}]_k^{(n)} \left(\left[\begin{array}{c} \mathbf{a}^s \\ \mathbf{b}^s \\ \bar{\mathbf{a}}^s \end{array} \right] \right)_k^{(n)} \tag{3.1.11}$$

- Step 3: Update the phase vector

$$\boldsymbol{\theta}_k^{(n+1)} = \boldsymbol{\theta}_k^{(n)} - \arctan\left(\frac{\mathbf{B}_k^n}{\mathbf{A}_k^n} \right) \tag{3.1.12}$$

- Step 4: If $\left\| \boldsymbol{\theta}^{(n)} - \boldsymbol{\theta}^{(n-1)} \right\| \leq \epsilon_I$, then $\mathbf{IMF}_k = \mathbf{H}_k^{(n)} \mathbf{x}_k^{(n)}$, $n \leftarrow 0$, $\mathbf{r}^k = \mathbf{r}^{k-1} - \mathbf{IMF}_k$,$k \leftarrow k+1$. Else, $n \leftarrow n+1$, goto Step 1.
- Step 5: If $\left\| \mathbf{r}^k \right\| \leq \epsilon_{II}$, stop. Else, goto Step 0.

$$\mathbf{H} = \mathbf{H}(\boldsymbol{\theta}) = [\boldsymbol{Cos\theta}\Psi, \boldsymbol{Sin\theta}\Psi] \in \mathbb{R}^{N \times 2(2m+1)}. \tag{3.1.9}$$

Here, the vector $\mathbf{x}$ is the vector of all unknowns and is assumed to have a sparse structure in the whole dictionary. All in all, the problem that we want to solve is

$$\underset{\mathbf{x},\boldsymbol{\theta}}{minimize} \quad \delta \|\mathbf{x}\|_1 + \|\mathbf{s} - \mathbf{H}\mathbf{x}\|_2^2 \; .$$

3.1.2. Algorithm. Algorithm 4 explains each step of the non-periodic STFR method. Take the sampled signal to be like a column vector $\mathbf{s}$. Please note that the compact form of the updater in the third step of this algorithm is not used explicitly. Instead, we use the discrete version of equation (2.4.9) (see step 5 of Algorithm 3):

$$\triangle\theta' = \frac{d}{dt} \arctan\left(\frac{\mathbf{B}_k^n}{\mathbf{A}_k^n} \right), \triangle\theta(t) = \int_0^t \triangle\theta'(\xi) \, d\xi.$$

Least-squares l_1-regularized minimization (3.1.10) is used in this algorithm for two reasons: first, we assume that the envelope can be approximated by a sparse number of terms and second, it acts as an stabilizer of the l_2 norm. Furthermore, unlike TV

STFR and periodic STFR, non-periodic STFR uses the dictionary explicitly in the algorithm. We use the method introduced in [34] to solve (3.1.10) numerically.

3.2. Numerical Examples

We ran a number of tests on this method, which show that it has all the strong attributes of other STFR methods. The regularization parameter δ was taken to be $O(1)$ for the cases that we do not have noise and $O(50)$ for the cases that we have noise. One can even pick different values of δ when extracting different IMFs. However, this makes the method slightly empirical than completely mathematical. The initial guess in all the cases were picked as constant-frequency random numbers. For example, if we intended to extract a trend, we used to take a low constant-frequency number for the initial guess.

Example 3. We first tested the algorithm's performance on an example signal with a linear trend and two constant envelope and frequency IMFs, $f(t) = 6t + \cos(8\pi t) + 0.5\cos(40\pi t)$ (see Figure 3.1). The algorithm extracted IMFs (see red lines in Figures 3.2-3.4). Besides some tiny boundary misalignment, the extractions were accurate, suggesting that the non-periodic STFR method is accurate away from the boundaries.

Example 4. In order to further test the abilities of the method, we tested a more complex signal (see Figure 3.5)

$$f(t) = 6t^2 + \cos\left(10\pi t + 10\pi t^2\right) + \cos\begin{cases} 60\pi t & 0 \le t \le 0.5 \\ 80\pi t - 10\pi & 0.5 \le t \le 1 \end{cases}$$

in which one of the IMFs has a jump in its IF. The trend in this case is quadratic. The IMFs were again extracted with high accuracy except for some boundary effects (see Figures 3.6-3.8) suggesting the method remains accurate even when the IF of one of the IMFs has some irregularities.

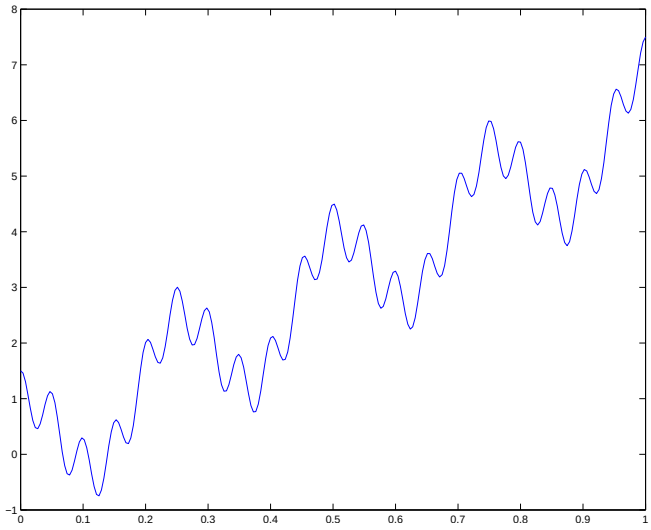

FIGURE 3.1. Signal with a Linear Trend: The horizontal axis is the time variable and the vertical one is the signal itself.

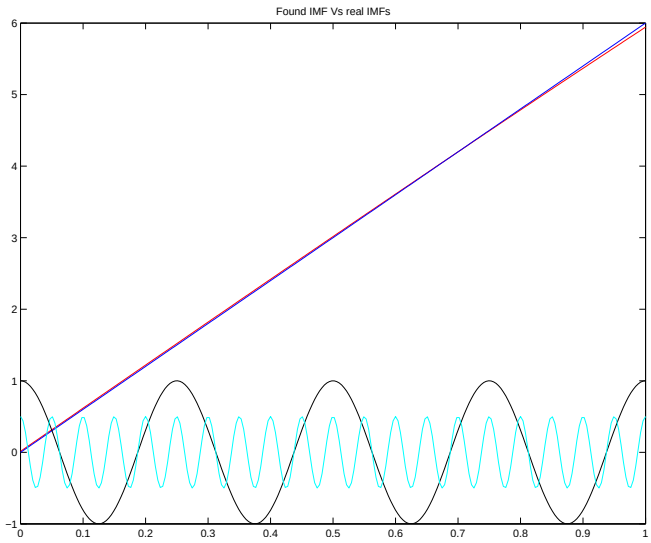

FIGURE 3.2. Extracted Trend: The linear trend is in blue and the extracted trend is in red. Except the right boundary, the error is small in the extracted trend.

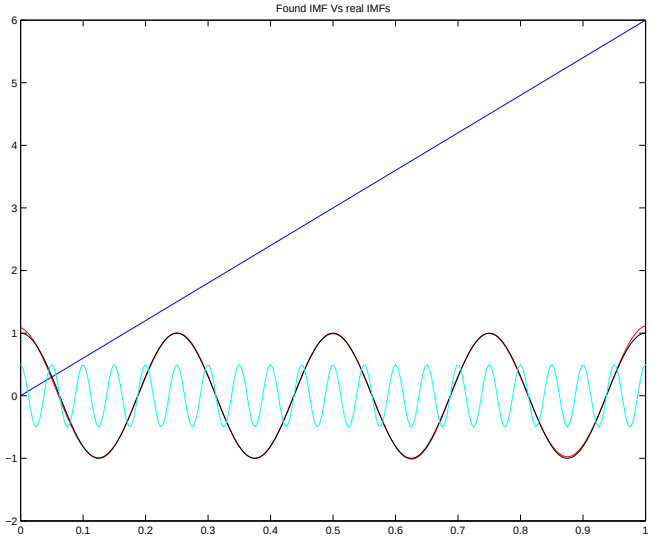

FIGURE 3.3. Extraction of the first IMF: The extracted IMF is in red. As can be seen, except from the boundaries the extraction is faithful.

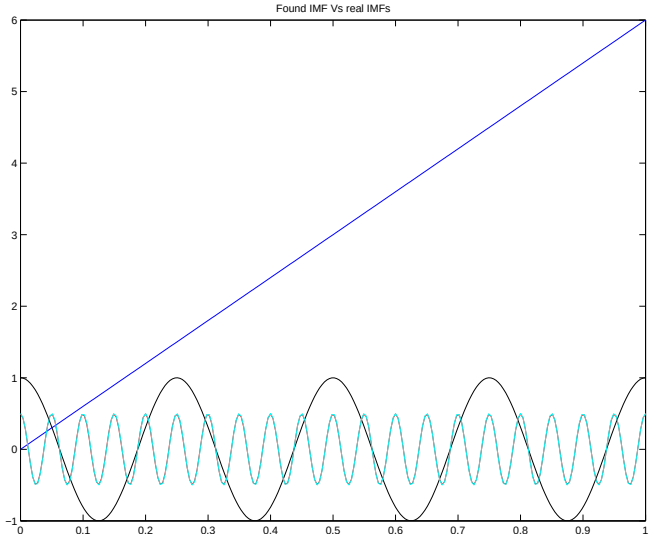

FIGURE 3.4. Extraction of the second IMF: The extracted second IMF is in red. It is almost indistinguishable from the high-frequency original IMF.

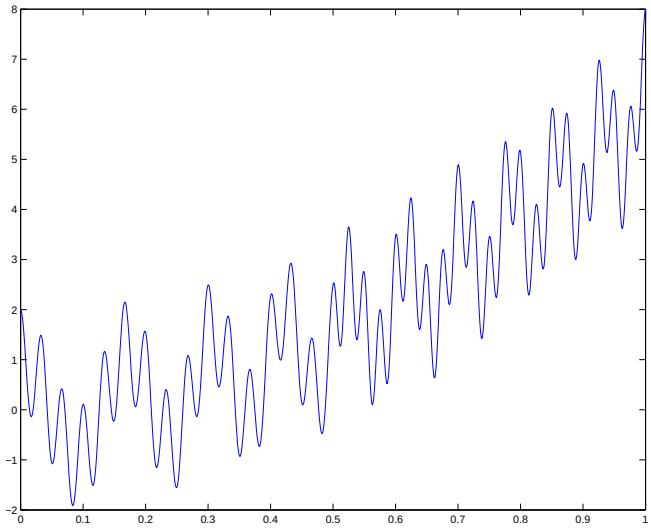

FIGURE 3.5. Signal with a Quadratic Trend: The horizontal axis is the time variable and the vertical one is the signal itself.

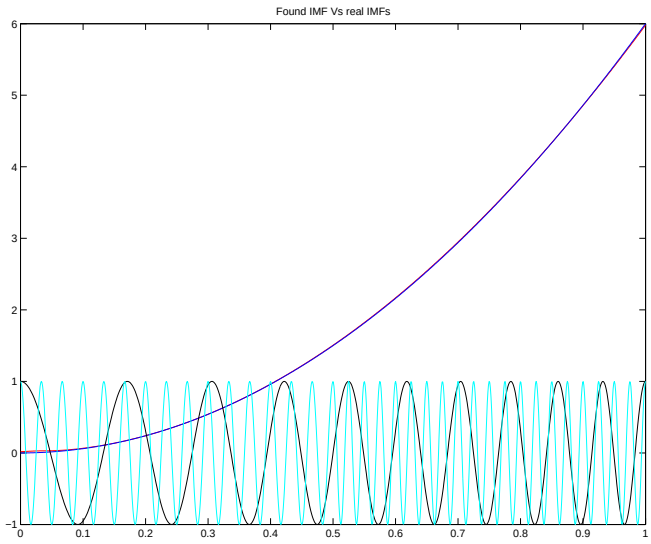

FIGURE 3.6. Extracted Trend: The quadratic trend is in blue and the extracted trend is in red. There is almost no error in the extraction.

Example 5. To further test the method, we ran the algorithm on a signal that contained some intrawave modulation in one of the IMFs (intrawave signals are addressed in greater detail in the next chapter). The signal used (see Figure 3.9) is described

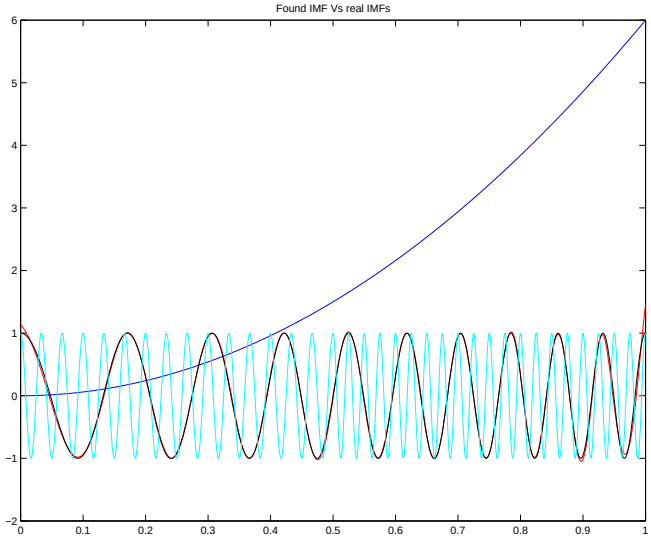

FIGURE 3.7. Extraction of the first IMF: The extracted IMF is in red. As can be seen, except from the boundaries the extraction is faithful.

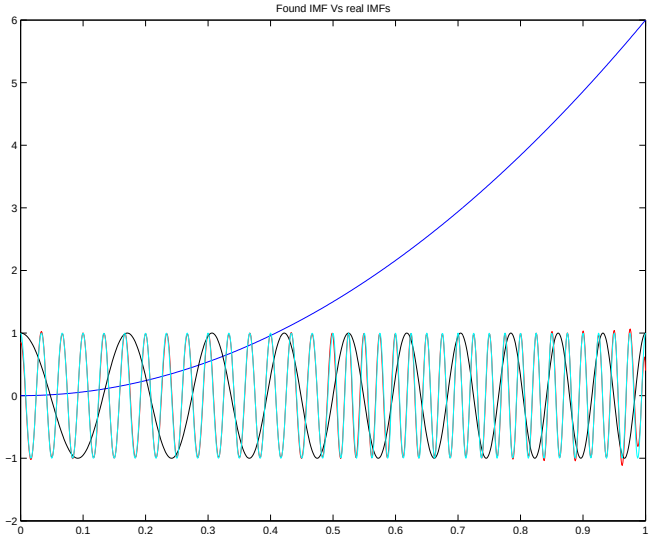

FIGURE 3.8. Extraction of the second IMF: The extracted IMF is in red. Like the previous example, the extraction is accurate, except near the right boundary.

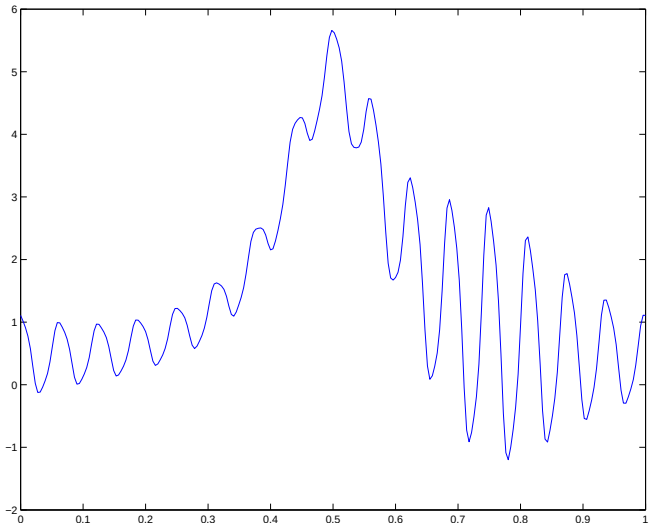

FIGURE 3.9. Signal with a Hump-like Trend: The horizontal axis is the time variable and the vertical one is the signal itself.

in mathematical terms as

$$f(t) = \frac{1}{1.2 + \cos(2\pi t)} + \frac{1}{1.5 + \sin(2\pi t)} \cos(32\pi t + 0.2\cos(64\pi t)).$$

The trend is like a hump. The IMFs were again extracted with high accuracy except near the boundaries (see Figures 3.10, 3.11).

Example 6. To check the stability of the algorithm with noise perturbation, we ran a series of tests that added white noise (represented by $\chi(t)$) to an IMF. In this example, the signal used (see Figure 3.12) was:

$$f(t) = \cos(60\pi t + 10\sin(2\pi t)) + \chi(t).$$

The extracted IMF was compared to the original one (see Figure 3.13). Even in the presence of heavy noise, the extraction is still acceptable, showing the stability of the non-periodic STFR method in the presence of noise.

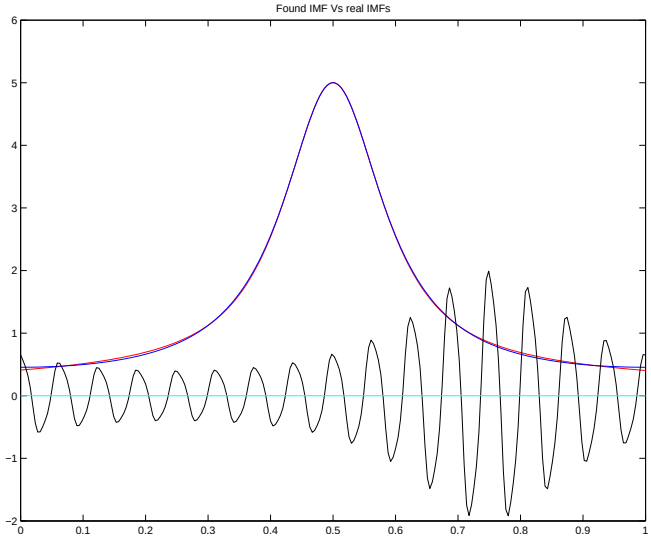

FIGURE 3.10. Extracted Trend: The trend is in blue and the extracted trend is in red. There is almost no error in the extraction, except near the boundaries.

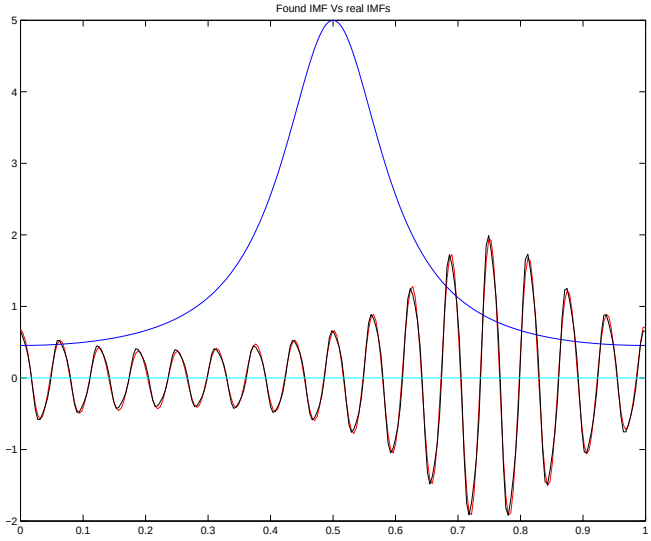

FIGURE 3.11. Extraction of the IMF: The extracted IMF is in red. Since the signal has intrawave modulation, the extraction has slight phase lags seen near the peaks and troughs. Still, the extraction is faithful.

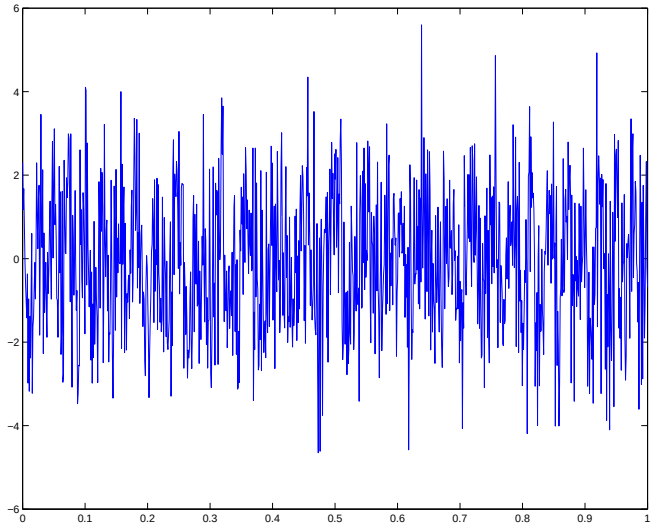

FIGURE 3.12. Signal with Noise Perturbation: The horizontal axis is the time variable and the vertical one is the signal itself.

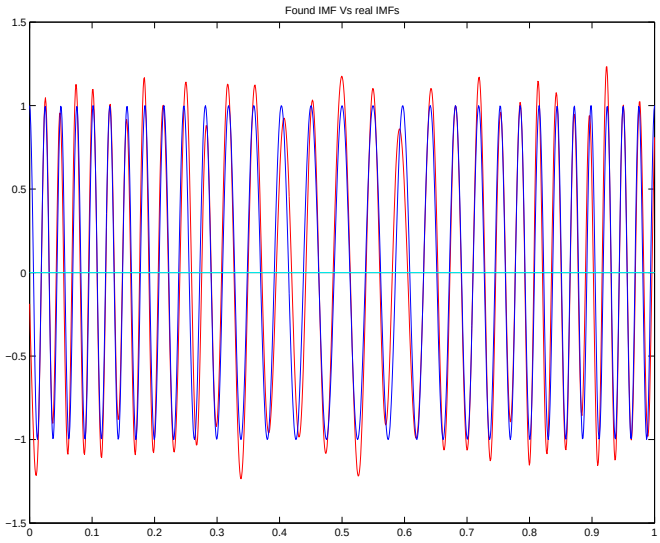

FIGURE 3.13. Extraction of the IMF: The extracted IMF is in red. Even in the presence of noise perturbation, the generality of the extraction is acceptable.

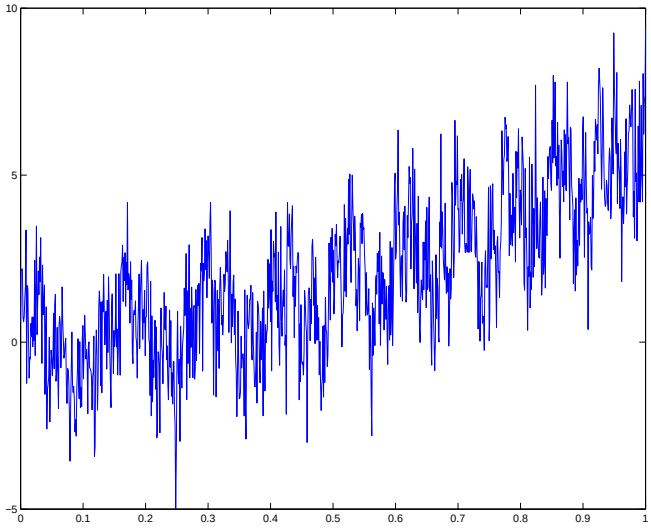

FIGURE 3.14. Signal with a Quadratic Trend Polluted with Noise Perturbation: The horizontal axis is the time variable and the vertical one is the signal itself.

Example 7. To further test the method's stability, we tested it on the signal from Example 4 plus white noise (see Figure 3.14)

$$f(t) = 6t^2 + \cos\left(10\pi t + 10\pi t^2\right) + \cos\begin{cases} 60\pi t & 0 \le t \le 0.5 \\ 80\pi t - 10\pi & 0.5 \le t \le 1 \end{cases} + \chi(t).$$

The trend is extracted fairly well, (see Figure 3.15). The IMFs are also acceptable (see Figures 3.16, 3.17). In fact, there is no observable phase error in extraction. Although there are jumps near the peaks and troughs of the IMFs, the trends of the IMFs are extracted properly, except for some boundary error. This further demonstrates the stability of the non-periodic STFR method in the presence of noise.

When compared with other STFR methods, the only shortcoming of the Non-Periodic STFR method is the speed of the algorithm (see Table 3.1). In particular, non-periodic STFR can be seen as a strong trend detector compared to even the EMD\EEMD [58].

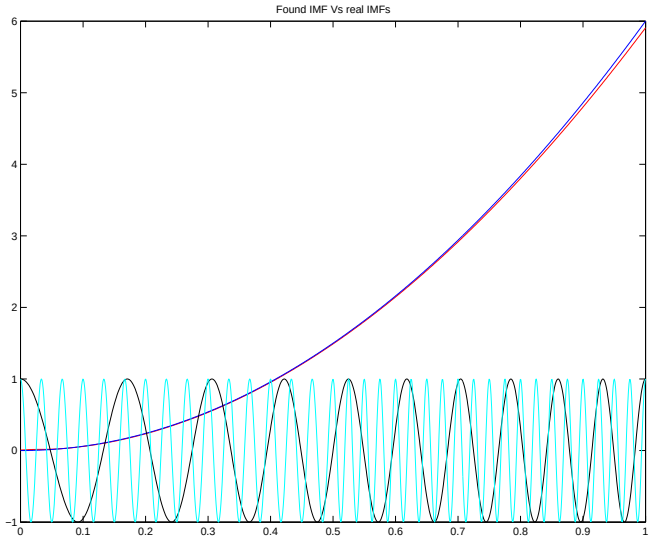

FIGURE 3.15. Extracted Trend: The linear trend is in blue and the extracted trend is in red. Due to the presence of noise, the extracted trend deviates from the original trend slightly.

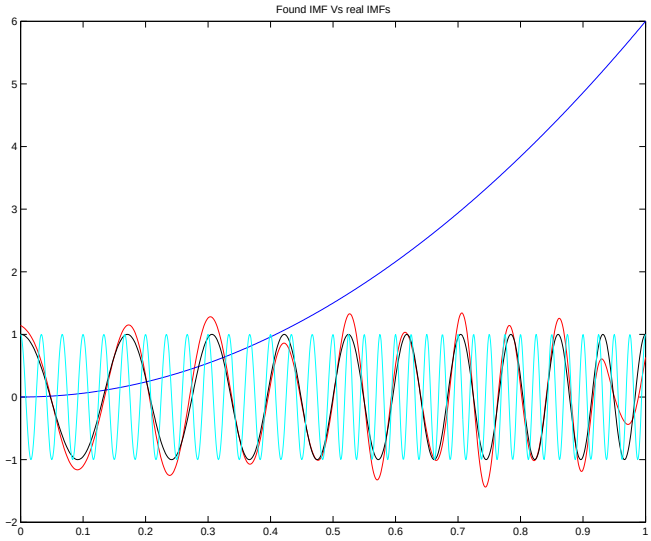

FIGURE 3.16. Extraction of the first IMF: The extracted IMF is in red. The only part of the extraction that is not completely acceptable is the right boundary of the extraction.

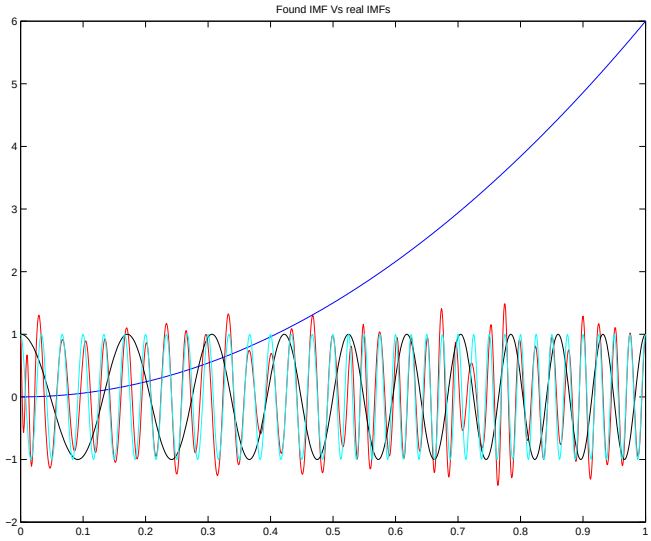

FIGURE 3.17. Extraction of the second IMF: The extracted IMF is in red. Here, the noise perturbation has more effect on the IMF extraction. However, the generality of the extraction is still acceptable.

	Periodic STFR	TV STFR	Non-Periodic STFR
Accuracy in envelope Extraction	High	Medium-High	Medium
Accuracy in IF Extraction	High	Medium-High	Medium
Accuracy in IMF Extraction	High	High	High
Speed	High	Medium	Medium
Non-Periodic Data	No	Yes	Yes
Noise Stability	High	Low	High
Boundary Error for Non-Periodic Data	High	Medium	Low
First Guess Initialization	No	Yes	No

TABLE 3.1. Comparison of the STFR Methods

CHAPTER 4

Extraction of Intrawave, Sharp, and Rare Event Signals using Sparse Time-Frequency Method

In general, intrawave signals are oscillatory signals that have intense frequency modulation in at least one θ-coordinate. By intense modulation we mean that the IF has oscillations that is comparable with the oscillation of the IMF itself. If the frequency modulation becomes even more intense, the resulting signal is called a sharp signal. The difference between the intrawave and sharp signal is in the intensity of the frequency modulation. Analyzing these IMFs has so far been a challenging problem in signal processing, specifically for adaptive methods.

Rare events are IMFs with compact support in the time domain. We consider them to be signals with constant very high frequency occurring at a very small time period. A rare event is essentially like a spike. Since they have a compact support, analysis of rare events is relatively hard. Even adaptive methods like EMD/EEMD are not able to process such signals with acceptable accuracy. In this chapter, we propose algorithms to analyze intrawave, sharp and rare event signals.

4.1. Intrawave Signals

Few previous methods have attempted the difficult task of extracting intrawave signals. One approach [54] uses the concept of wave shape function to tackle the problem. However, it is not clear how the shape function is learned. The EMD method can extract one intrawave IMF in the absence of noise. However, neither EMD nor EEMD can extract even one intrawave IMF in the presence of noise. As can be seen from the sparsity of literature in this field, no method that we know of can faithfully extract intrawave IMFs, and this theoretically challenging problem has

been an open problem for many years. Nevertheless, these signals are abundant in second order ordinary differential equations, specifically conservative systems, and if one is eager to identify the original ODE from an observed signal, one must be able to extract intrawave signals properly. Here, we present a method for extracting an intrawave signal with the highest possible accuracy. This means that we propose a method that does not break an intrawave signal into subharmonics.

Before formally defining an intrawave signal in mathematical terms, we consider a clarifying example, using a model intrawave problem:

$$x(t) = \cos\left(\omega t + \frac{\triangle\,\omega}{p}\sin(pt)\right), \tag{4.1.1}$$

where ω is the carrier frequency, $\triangle\,\omega < \omega$ is the strength of frequency modulation, and p is the frequency of the frequency modulation. This example follows initial work by Van der Pol [51]. By taking $\theta(t) = \omega t$, one can map the signal (4.1.1) into the θ-coordinate,

$$x(\theta) = \cos\left(\theta + \frac{\triangle\,\omega}{p}\sin\left(\frac{p}{\omega}\theta\right)\right). \tag{4.1.2}$$

Simplifying this we have

$$x(\theta) = \cos\theta\cos\left(\frac{\triangle\,\omega}{p}\sin\left(\frac{p}{\omega}\theta\right)\right) - \sin\theta\sin\left(\frac{\triangle\,\omega}{p}\sin\left(\frac{p}{\omega}\theta\right)\right).$$

Expanding the terms $\cos\left(\frac{\triangle\omega}{p}\sin\left(\frac{p}{\omega}\theta\right)\right)$ and $\sin\left(\frac{\triangle\omega}{p}\sin\left(\frac{p}{\omega}\theta\right)\right)$, using Fourier transform, we have

$$\begin{aligned} x(\theta) = &\left\{J_0\left(\tfrac{\triangle\omega}{p}\right) + 2\textstyle\sum_{k=1}^{\infty} J_{2k}\left(\tfrac{\triangle\omega}{p}\right)\cos\left(2k\tfrac{p}{\omega}\theta\right)\right\}\cos\theta - \\ &\left\{2\textstyle\sum_{k=1}^{\infty} J_{2k-1}\left(\tfrac{\triangle\omega}{p}\right)\sin\left((2k-1)\tfrac{p}{\omega}\theta\right)\right\}\sin\theta, \end{aligned} \tag{4.1.3}$$

where J_k is the the Bessel function of the first kind of order k. Hence, the original signal (4.1.1) can be expressed in the form $x(\theta) = a(\theta)\cos\theta + b(\theta)\sin\theta$, where

$$a(\theta) = \left\{J_0\left(\frac{\triangle\,\omega}{p}\right) + 2\sum_{k=1}^{\infty} J_{2k}\left(\frac{\triangle\,\omega}{p}\right)\cos\left(2k\frac{p}{\omega}\theta\right)\right\}, \tag{4.1.4}$$

$$b(\theta) = -\left\{2\sum_{k=1}^{\infty} J_{2k-1}\left(\frac{\triangle\,\omega}{p}\right)\sin\left((2k-1)\,\frac{p}{\omega}\theta\right)\right\}. \tag{4.1.5}$$

For a fixed $\frac{\triangle\omega}{p}$, as $k \to \infty$, $J_k\left(\frac{\triangle\omega}{p}\right)$ decreases. However, if $\frac{\triangle\omega}{p}$ is large enough, many of the high frequency terms in this formula would make a considerable tail in the Fourier domain. This thick tail is the main source of difficulty in intrawave IMF extraction. In the STFR method, an explicit band filter is used at each iteration [25]. This band filter is essentially the numerical counter part of (1.2.3). At the same time, in EMD, an implicit filter is used [20, 57, 55]. The narrow band filters used in these methods decompose a wide band signal into subharmonics instead of a single IMF. Intrawave signals, then, would be elusive for analysis in any method that uses a narrow band filter. In what follows, we show the problem in STFR terminology in detail.

At each iteration with θ as an initial guess, the STFR algorithm picks terms that are in $V(\theta)$ in order to extract the envelope (see (1.2.3)). Hence, having $\theta(t) = \omega t$ in (4.1.1), the maximum number of terms that can be collected by the STFR method for the envelope $a(\theta)$, based on dictionary (1.2.3) and equation (4.1.3), is $k_{max} = \left\lfloor \frac{\omega}{4p} \right\rfloor$. For $b(\theta)$, this value is $k_{max} = \left\lfloor \frac{\omega}{4p} + \frac{1}{2} \right\rfloor$. In other words, the approximated envelopes $\tilde{a}(\theta)$, $\tilde{b}(\theta)$ would look like

$$\tilde{a}(\theta) = \left\{J_0\left(\frac{\triangle\,\omega}{p}\right) + 2\sum_{k=1}^{\lfloor \frac{\omega}{4p} \rfloor} J_{2k}\left(\frac{\triangle\,\omega}{p}\right)\cos\left(2k\frac{p}{\omega}\theta\right)\right\}, \tag{4.1.6}$$

$$\tilde{b}(\theta) = -\left\{2\sum_{k=1}^{\lfloor \frac{\omega}{4p} + \frac{1}{2} \rfloor} J_{2k-1}\left(\frac{\triangle\,\omega}{p}\right)\sin\left((2k-1)\,\frac{p}{\omega}\theta\right)\right\}. \tag{4.1.7}$$

Obviously, if $\frac{\omega}{4p} \ll 1$, the only term that will be collected by the STFR algorithm is $J_0\left(\frac{\triangle\omega}{p}\right)$. The other terms would be discarded. In other words, the STFR algorithm will break an intrawave signal into many subharmonics in extraction. The formal definition for an intrawave signal, then, is:

Definition 7. A signal $x(t) = a(t)\cos\theta(t)$, having only one extremum between two consecutive zeros, is an *Intrawave Signal* in coordinate θ, if the envelope $a > 0$ in θ-coordinate, when decomposed by Fourier transform, has terms like $\cos\nu\theta$ and $\sin\nu\theta$ for $\nu \geq \frac{1}{2}$.

Since the main difficulty in dealing with intrawave signals comes from their wide band representation in the frequency domain, which cannot be properly analyzed using methods with explicit or implicit narrow band filters, we propose a method that modifies the normal envelope dictionary (1.2.3) in an STFR framework in order to extract intrawave signals with high accuracy.

Specially, in order to capture more terms of $a(\theta)$ and $b(\theta)$, we need to modify the dictionary (1.2.3), or more precisely the filter, as follows

$$V(\theta) = span\left\{1, \cos\left(\frac{\theta}{\lambda}\right), \sin\left(\frac{\theta}{\lambda}\right) \mid \lambda \geqslant \lambda_0 > 0\right\}. \tag{4.1.8}$$

Here λ_0 should be as small as possible such that enough terms in $a(\theta)$ and $b(\theta)$ can be collected. In this case, the maximum number of terms that can be collected by the STFR method, based on dictionary (4.1.8), is $k_{max} = \left\lfloor \frac{\omega}{4p\lambda_0} \right\rfloor$ for $a(\theta)$. For $b(\theta)$, this value is $k_{max} = \left\lfloor \frac{\omega}{4p\lambda_0} + \frac{1}{2} \right\rfloor$. Obviously, the choice of λ_0 depends on the decay rate of $J_k\left(\frac{\triangle\omega}{p}\right)$ and the ratio $\frac{\omega}{p}$. At the same time, if the recording of the signal is noisy, this parameter should be picked carefully not to collect noise.

This small modification would allow us to treat an intrawave signals without major changes to the original STFR Algorithms 3, and 4. Furthermore, we show that although enlarging the filter requires that the IMF components of the IS must have enough separate time-frequency representations, the method is not problematic when extracting non-separable time-frequency IMFs from a signal provided they are extracted simultaneously. The convergence of the modified algorithm will be proved for the case of periodic signals in the next chapter.

4.2. Algorithmic Analysis

Our filter modification for capturing the thick tail of the intrawave signals has not been implemented in EMD [32] or the Synchrosqueezed wavelet method [17, 50]. However, for both periodic and non-periodic STFR method [25], it is straightforward. The filters or envelope dictionaries must be made wider. This fact manifests itself in the shape of the cutoff function in the periodic Algorithm 3,

$$\chi(\omega) = \begin{cases} 1, & -\frac{1}{\lambda} < \omega < \frac{1}{\lambda}, \\ 0, & otherwise, \end{cases} \tag{4.2.1}$$

where λ is a measure of the span of the filter. In the next chapter, we prove that convergence enhances if $\lambda \to 1^+$. However, our numerical results will show that under intense frequency modulations $\lambda \to 0^+$ can be used legitimately, though with extreme caution when noise is present. For the non-periodic STFR Algorithm 3.1, the modification should be implemented where we define the envelope dictionary

$$\{k_l\}_{l=1}^m \subset [\lambda, \infty). \tag{4.2.2}$$

The filter parameter λ must be set based on the characteristics of the intrawave signal that we aim to extract. However, $\lambda = 1^+$ is always a very good practical starting guess.

4.3. Numerical Examples

The following test cases illustrate the credibility of our STFR method for extracting intrawave IMFs.

Example 8. Here we study the effect of the filter width on one intrawave IMF with constant envelope. Consider the mild intrawave signal

$$x(t) = \cos\left(8\pi t + \frac{1}{\pi}\sin(4\pi t)\right).$$

The IF of this signal is $8\pi+4\sin(4\pi t)$. The strength of frequency modulation is 4, and the frequency of the frequency modulation is 4π. Compared to the carrier frequency, 8π, the modulated frequency is not negligible. Hence, we expect to get incomplete and inaccurate extraction if a narrow-band filter is used. As expected for a normal narrow-band filter with parameter $\lambda = 2$, the IMF extraction error is relatively large (see Figure 4.1). Increasing the length of the filter enhances the extraction drastically: a wider filter with parameter $\lambda = 1^+$ reduces the maximum error by 10^{-2} (see Figure 4.2). While a narrow-band filter only extracts one of the harmonics of the intrawave signal, acting like a one-mode Fourier transform, then, making the filter wider collects more subharmonics as a whole. Consequently, the result looks much better.

Finally, there is an observation that we want to mention regarding Figure 4.1. We recall that when using a narrow band filter, the total number of terms that can be collected by the algorithm is $k^a_{max} = \left\lfloor \frac{\omega}{4p} \right\rfloor$ and $k^b_{max} = \left\lfloor \frac{\omega}{4p} + \frac{1}{2} \right\rfloor$ in the best case, for a good constant initial guess. In our case, $k^a_{max} = 0$, $k^b_{max} = 1$. Hence, the main contribution comes from $J_0\left(\frac{\triangle\omega}{p}\right)$, which is 0.9748, in this example. This observation can be seen perfectly well in Figure 4.1. The plot in the bottom left corner of the figure shows that the extracted envelope is roughly 0.975. Also, the extracted IMF in the bottom right corner of the same figure shows that there is a tiny phase shift, which we can certainly say comes from $k^b_{max} = 1$.

Example 9. Here, we mix the intrawave signal from Example 8 with a high-frequency signal

$$v(t) = (1 + 0.3\cos(8\pi t))\left(\cos\left(40\pi t + \frac{5}{2}\sin(2\pi t)\right)\right).$$

This high-frequency signal has small intrawave modulation as well. However, it is not comparable to the frequency modulation of $\cos\left(8\pi t + \frac{1}{\pi}\sin(4\pi t)\right)$. The final form of the signal is

$$x(t) = \cos\left(8\pi t + \frac{1}{\pi}\sin(4\pi t)\right) + (1 + 0.3\cos(8\pi t))\left(\cos\left(40\pi t + \frac{5}{2}\sin(2\pi t)\right)\right).$$

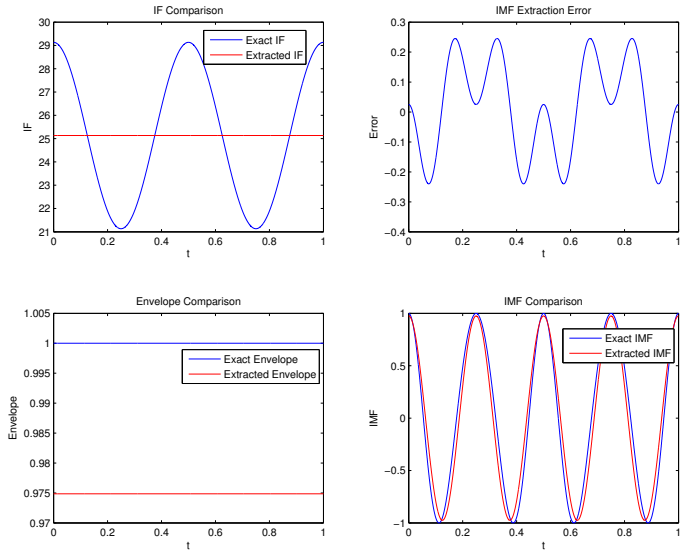

FIGURE 4.1. Mild Intrawave Signal vs Narrow Band Filter

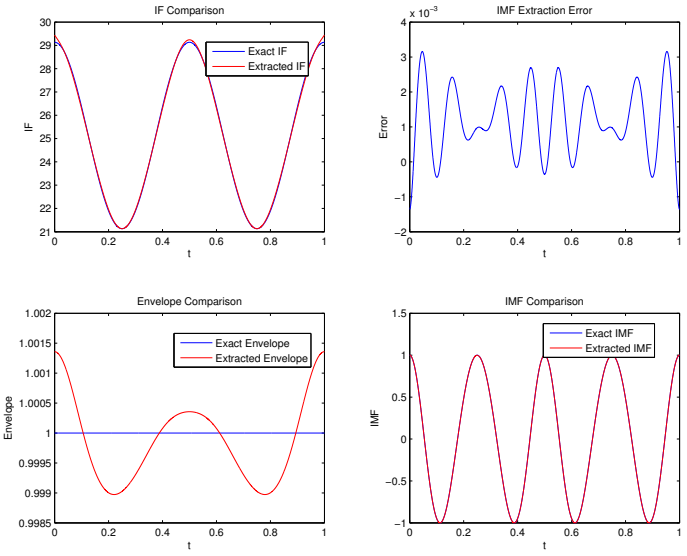

FIGURE 4.2. Mild Intrawave Signal vs Wide Band Filter

We successfully extracted the intrawave part, $\cos\left(8\pi t + \frac{1}{\pi}\sin(4\pi t)\right)$ using a filter with parameter $\lambda = 1^+$ (see Figure 4.3) and successfully extracted the rest of the

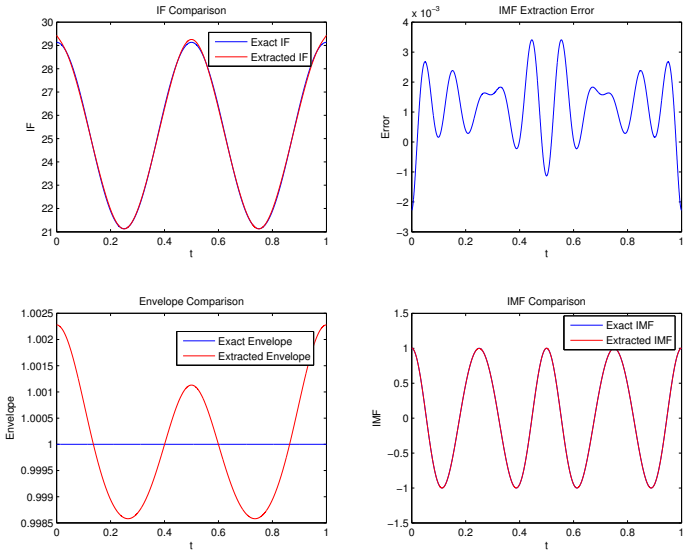

FIGURE 4.3. Intrawave Part of the Mixed Signal

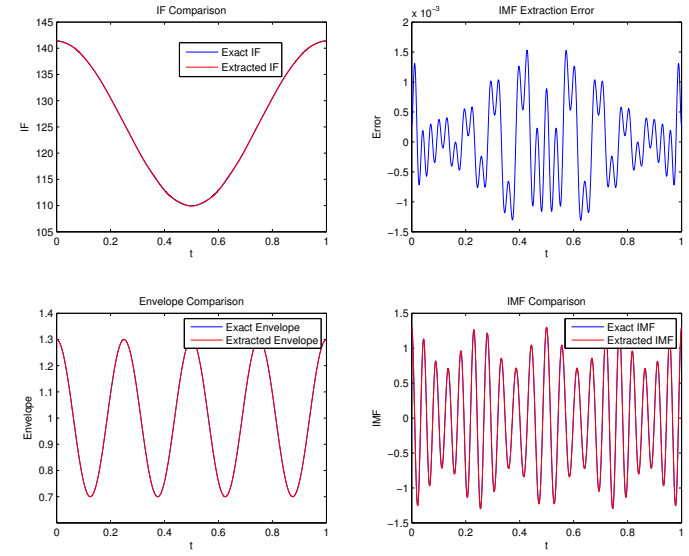

FIGURE 4.4. High Frequency Part of the Mixed Signal

signal using a normal filter with parameter of $\lambda = 2$ (see Figure 4.4). The maximum error in extraction is less than 5×10^{-3} for both IMFs.

Example 10. To further test our approach, we considered a case in which the intrawave modulation is intense. The signal has very sharp peaks and flat troughs. The mathematical form of the signal is

$$x(t) = \cos\left(8\pi t + \frac{21}{8\pi t}\sin(8\pi t)\right).$$

Here, even if the filter parameter is taken to be $\lambda = 1^+$, the results are not yet acceptable (see Figure 4.5). When filter parameter is decreased to $\lambda = 0.3$, the extraction is now possible (see Figure 4.6), although the envelope is not recovered exactly (Figure 4.6 bottom left). To compare our results with other adaptive methods, we analyzed the signal with the Synchrosqueezed Wavelet transform method [17, 50], which unfortunately decomposes the intrawave signal into its harmonics (see Figure 4.7). Matlab code for the Synchrosqueezed Wavelet transform can be found in [9]. To further test the method's accuracy, we considered an intense intrawave signal with a changing envelope using $\lambda = 0.3$. Again, although the envelope extraction is not accurate, the final extraction of the signal is reliable (See Figure 4.8). In contrast, the EMD was unable to decompose the intrawave in one piece (see Figure 4.9) suggesting that our method is more general than other adaptive methods.

Example 11. To investigate the effect of noise on the extraction of an intrawave signal we used an intrawave signal with strong frequency modulation (see Figure 4.10) of form

$$x(t) = \cos\left(8\pi t + \frac{21}{8\pi t}\sin(8\pi t)\right) + 0.1\mathcal{N}(0,1).$$

Here, $\mathcal{N}(0,1)$ is a normal Gaussian random variable. Again using $\lambda = 0.3$, the final extraction is faithful, even though there minor error in the extraction of the envelope and the instantaneous frequency (Figure 4.11). In this case, the noise perturbation was not large compared to the energy of the signal. To test the signal with more noise, we used

$$x(t) = \cos\left(8\pi t + \frac{21}{8\pi t}\sin(8\pi t)\right) + \mathcal{N}(0,1).$$

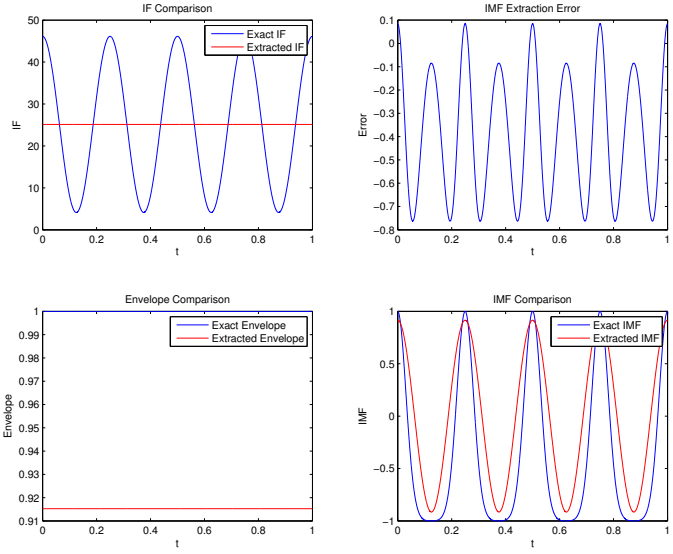

FIGURE 4.5. Intense Intrawave Extraction Failure

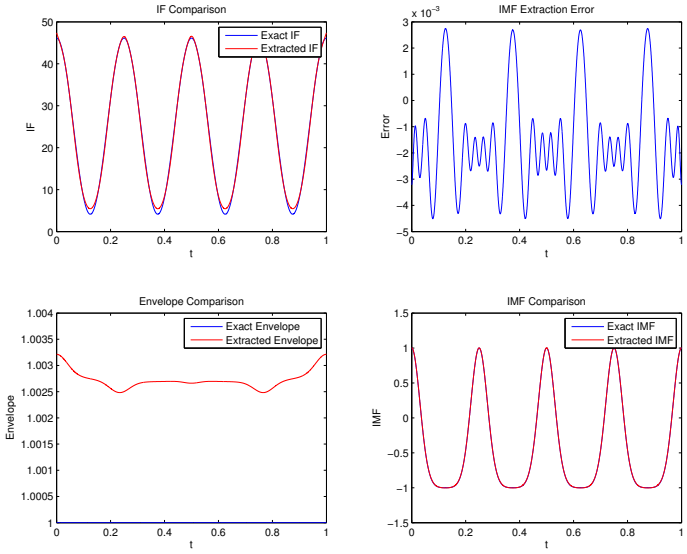

FIGURE 4.6. Intense Intrawave Extraction

This signal is buried in noise (see Figure 4.12). Using $\lambda = 0.8$ the extraction is acceptable in spite of errors in extracted envelope and IF even in the presence of intense noise perturbation (see Figure 4.13). We would like to point out that if we

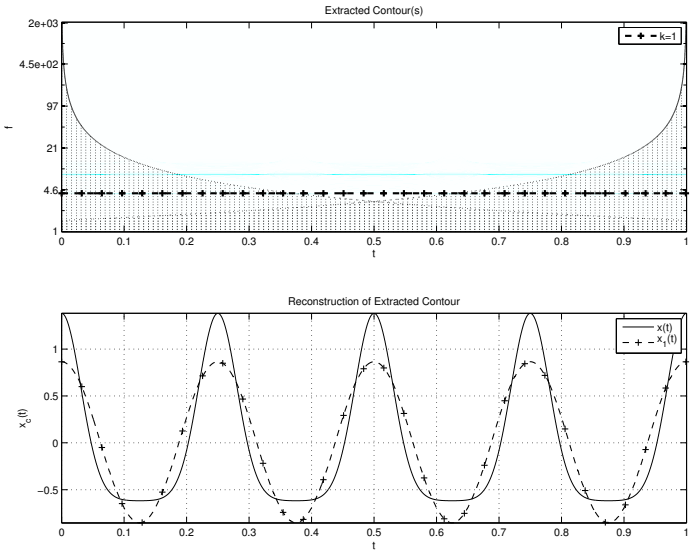

FIGURE 4.7. Synchrosqueezed Wavelet Comparison. Top: The frequency spectrum shows that the Synchrosqueezed method detects two major frequency trends. Bottom: The first IMF extracted using this analysis is like the first dominant harmonics. In this analysis, Morlet wavelet was used.

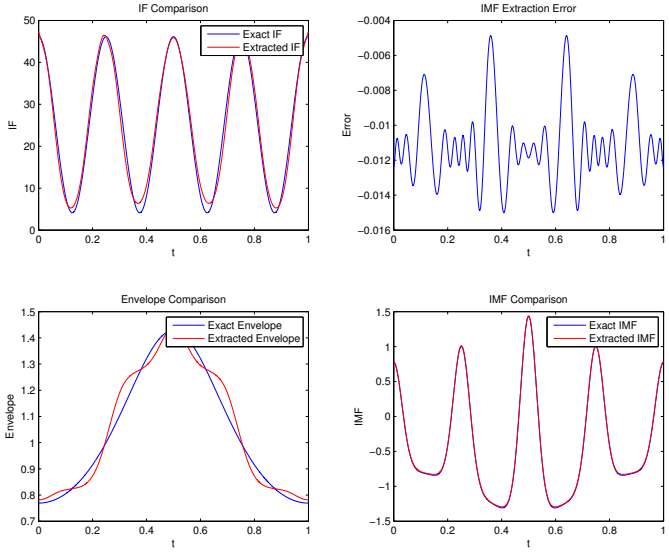

FIGURE 4.8. Intense Intrawave with Non-Constant Envelope

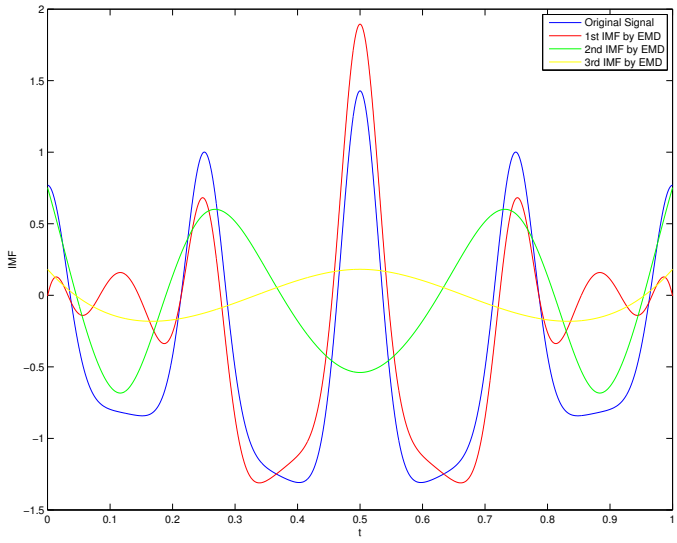

FIGURE 4.9. EMD Extraction Result for Intrawave Signal with Moving Envelope

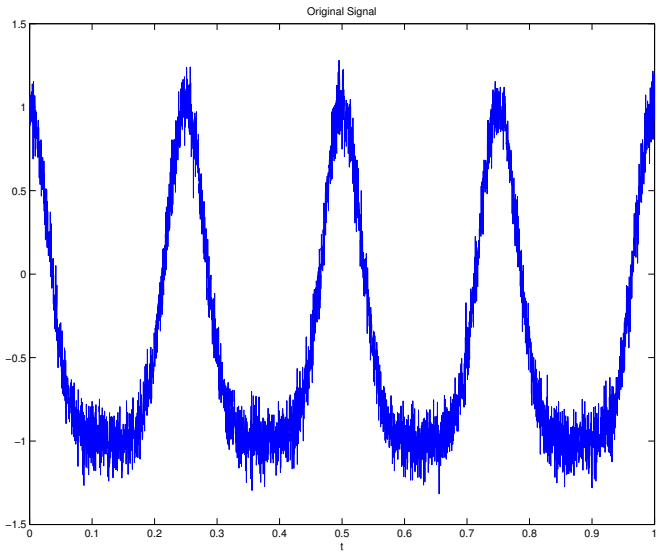

FIGURE 4.10. Original Signal

had taken smaller values of λ, the algorithm would have diverged due to the strong presence of noise.

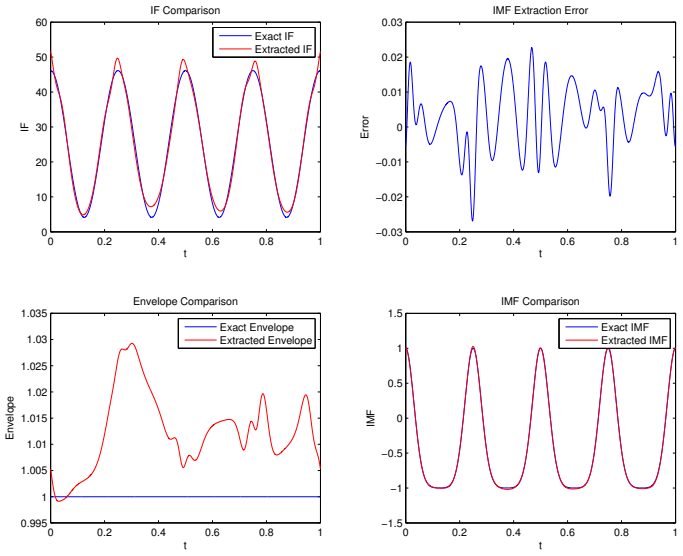

FIGURE 4.11. Extraction Result of the Mildly Noisy Signal

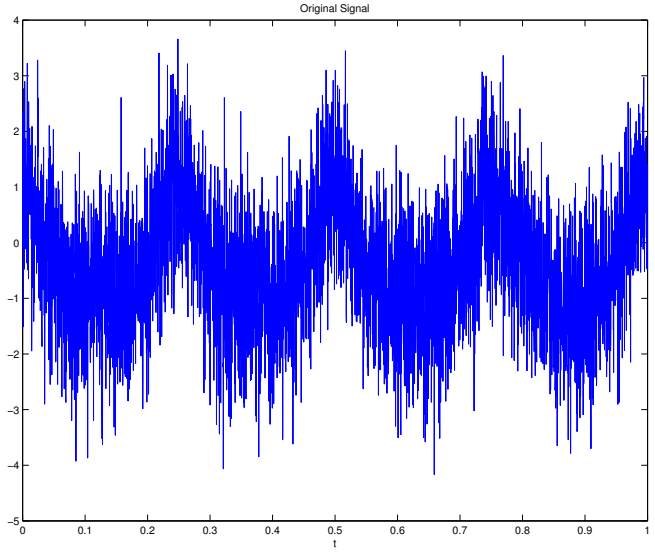

FIGURE 4.12. Original Signal

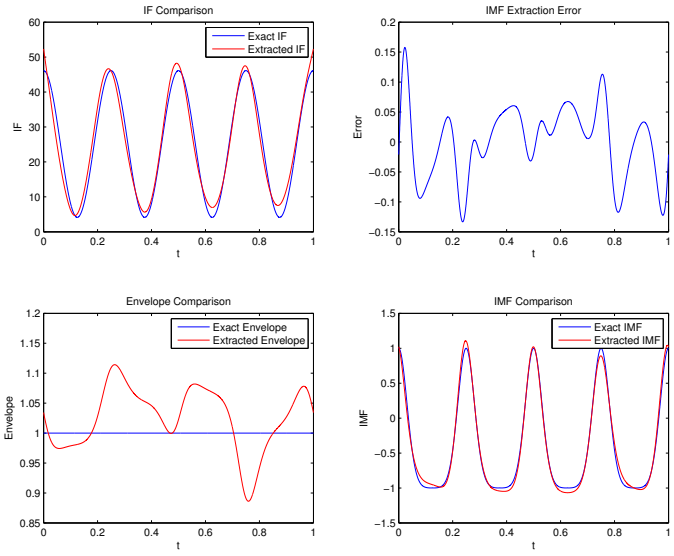

FIGURE 4.13. Extraction Result of the Intensely Noisy Signal

4.4. Mixed Intrawaves, Sharp Signals, Rare Events

In this section, we work on examples where we have a mixture of intrawave, sharp signals and rare events. Mode mixture[1] is a challenging problem for all adaptive data processing methods. In this section, we propose an approximate method to extract IMFs with mode mixture. The essence of the method is simple: when two or more IMFs have mode mixture in an IS, we can extract them simultaneously since we know the number of mixed IMFs.

4.4.1. Algorithm. For a signal with only two IMFs with mode mixture in the specified IS $\mathbf{s}$, Algorithm 4 would still be applicable with only a minor change: extract both IMFs simultaneously. Mathematically, we have

$$Minimize\ \ \delta \left\| \begin{pmatrix} \mathbf{x}_{1,k}^{(n)} \\ \mathbf{x}_{2,k}^{(n)} \end{pmatrix} \right\|_1 + \left\| \mathbf{s} - \left(\mathbf{H}_{1,k}^{(n)}, \mathbf{H}_{2,k}^{(n)} \right) \begin{pmatrix} \mathbf{x}_{1,k}^{(n)} \\ \mathbf{x}_{2,k}^{(n)} \end{pmatrix} \right\|_2^2, \tag{4.4.1}$$

[1]Mode mixture is a case when two or more IMFs in an IS have overlapping or very close IFs at some instances in time.

where the subscripts $1, 2$ belong to the IMFs that have mode mixture. In the presence of intrawave IMFs, the width of the filter must also be changed. Similar for the periodic Algorithm 3, both IMFs should be extracted simultaneously with a wide band filter but the rest of the algorithm would remain intact. None of the sparse time frequency methods can extract them successfully, to the best of our knowledge. Here we present test cases that show how this method works.

4.4.2. Numerical Examples.

Example 12. Here, we investigate the extraction of two intrawave signals with strong frequency modulation that do not have well separated scales. The signal is of the form $f(t) = \cos\theta_1 + \cos\theta_2$ (see Figure 4.14). The IMFs have intense intrawave characteristics and the mode mixture is apparent in the signal. since the signals are not separate in the time-frequency domain, the extraction of such IMFs is challenging. To overcome this difficulty, we use a wide band filter and we extract the IMFs simultaneously (see Figure 4.15) resulting in a fairly successful extraction. The error of the extraction is acceptable, as no other method is capable of such performance. The only *a priori* knowledge that we have used is that there are two IMFs in the original signal. However, this assumption can be dropped in an iterative search method.

It is straightforward to show formally why the extraction in Example 12 works. Let's assume that the signal can be written as

$$x(t) = \cos\left(\omega_1 t + \frac{\triangle\,\omega_1}{p_1}\sin(p_1 t)\right) + \cos\left(\omega_2 t + \frac{\triangle\,\omega_2}{p_2}\sin(p_2 t)\right). \tag{4.4.2}$$

Assume that we intend to extract both IMFs simultaneously. Assume that the initial guesses are $\theta_1(t) = \omega_1 t$ and $\theta_2(t) = \omega_2 t$. Using the expansion in (4.1.3), we can write (4.4.2) as

$$x(\theta_1, \theta_2) = \left\{ J_0\left(\frac{\triangle\,\omega_1}{p_1}\right) + 2\sum_{k=1}^{\infty} J_{2k}\left(\frac{\triangle\,\omega_1}{p_1}\right)\cos\left(2k\frac{p_1}{\omega_1}\theta_1\right)\right\}\cos\theta_1$$

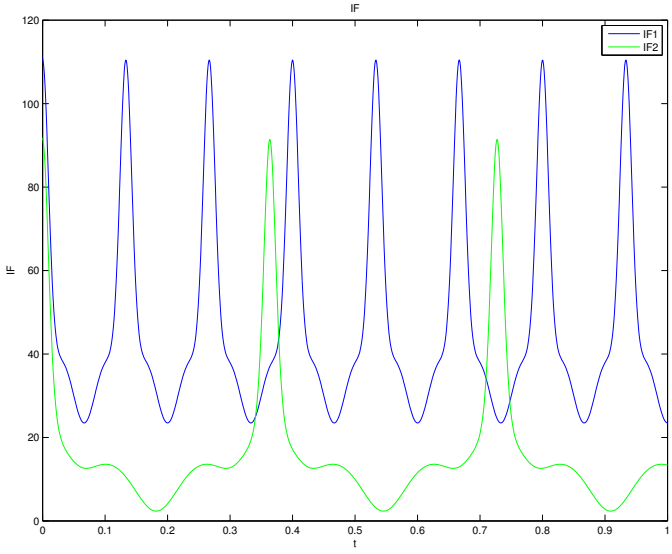

FIGURE 4.14. IF: The horizontal axis is the time variable and the vertical axis shows the IF.

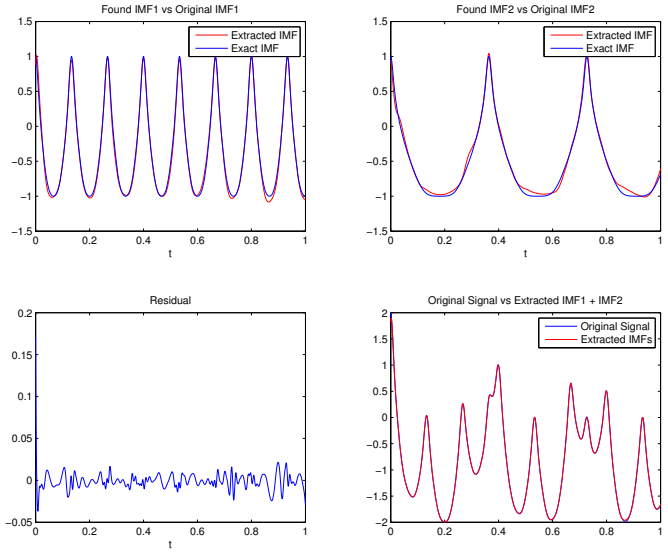

FIGURE 4.15. Results of the Extraction

$$-\left\{2\sum_{k=1}^{\infty} J_{2k-1}\left(\frac{\triangle\,\omega_1}{p_1}\right)\sin\left((2k-1)\,\frac{p_1}{\omega_1}\theta_1\right)\right\}\sin\theta_1$$

$$+\left\{J_0\left(\frac{\triangle\,\omega_2}{p_2}\right)+2\sum_{k=1}^{\infty}J_{2k}\left(\frac{\triangle\,\omega_2}{p_2}\right)\cos\left(2k\frac{p_2}{\omega_2}\theta_2\right)\right\}\cos\theta_2 \tag{4.4.3}$$

$$-\left\{2\sum_{k=1}^{\infty}J_{2k-1}\left(\frac{\triangle\,\omega_2}{p_2}\right)\sin\left(\left(2k-1\right)\frac{p_2}{\omega_2}\theta_2\right)\right\}\sin\theta_2.$$

The sparse representation of (4.4.3) suggests the use of definition (4.1.8) for both θ_1, θ_2 in a singular form (i.e. with a small λ_0). In the presence of noise, one should penalize the sparsity more than ever or the noise would be absorbed in the IMFs, which occurs due to the presence of high frequency mixture of the noise and high frequency terms in (4.4.3).

Example 13. To test the stability of the algorithm in the presence of noise perturbation, we take the original signal to be

$$f(t)=\cos\left(12\pi t+0.4\sin\left(12\pi t\right)\right)+\cos\left(16\pi t+0.3\cos\left(16\pi t\right)\right)+\mathcal{N}(0,1).$$

In spite of some blurring near the peaks and troughs, the result of the extraction is acceptable compared to the original signal (see Figure 4.16), showing that even in the presence of heavy noise, the algorithm works well.

Not only is the algorithm that we use to extract the IMFs stable, then, it is also stable to noise perturbation. The EMD/EEMD methods fail to extract the two IMFs properly. In fact all other adaptive methods fail to extract one IMF with intrawave frequency modulation in the presence of noise, let alone two IMFs with intrawave characteristics mode mixture.

4.4.2.1. *Rare Events.* So far, we have only considered intrawave and sharp signals. The method is also applicable to rare event signals. In this dissertation, a rare event should satisfy these conditions:

(1) The signal is compactly supported.

(2) Since the rare event is only happening in a short period of time, we assume that the envelope and the IF of the event is constant.

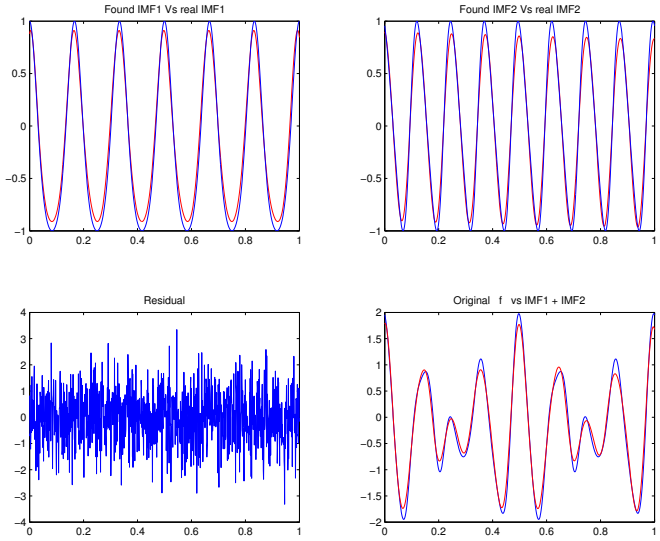

FIGURE 4.16. Results

With these assumptions, extraction of a rare event is trivial as long as the location of the event is known. However, if the rare event is mixed with other IMFs, the extraction becomes nonlinear and hard. Here, we present this issue in an example and explain how to solve this problem.

Example 14. Here we take two constant envelope IMFs; one is a sharp signal, and the other is a rare event (see Figure 4.17 and 4.18). After using a WFT to find the approximate location of the rare event (see Figure 4.19), we extract both IMFs using a simultaneous extraction. In contrast to formulation (4.4.1), here one dictionary is the adaptive STFR dictionary with a wide band filter $\mathbf{H}_{1,k}^{(n)}$, and the second is a fixed dictionary of all rare events in the approximate location of the rare event. The signals were successfully extracted with good accuracy (see Figure 4.20).

We have shown in this chapter that the problems of mode mixture, intrawave extraction, and rare events can be addressed efficiently and accurately using modified STFR methods. No other adaptive signal processing method can address all these problems in one package.

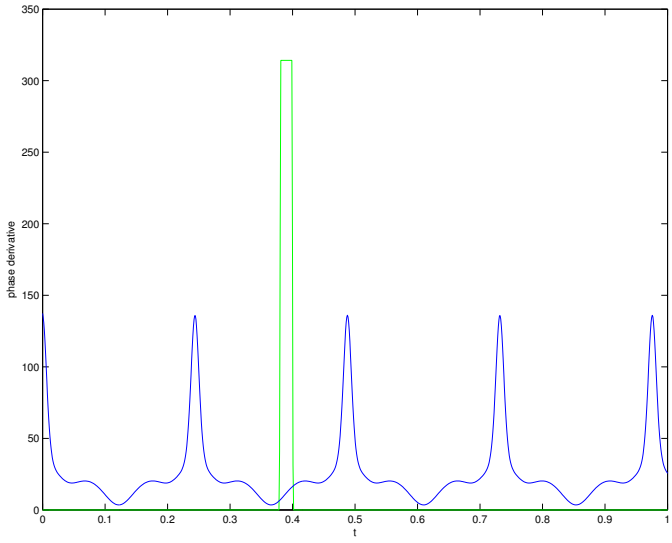

FIGURE 4.17. IF of a rare event and an intrawave signal.

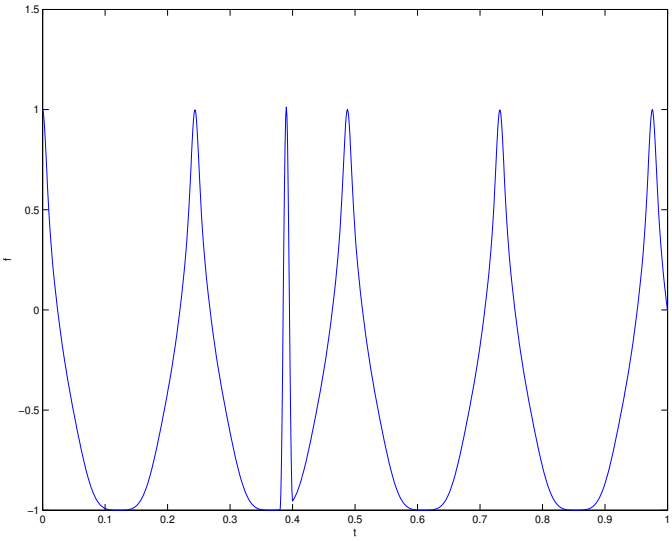

FIGURE 4.18. Original Signal

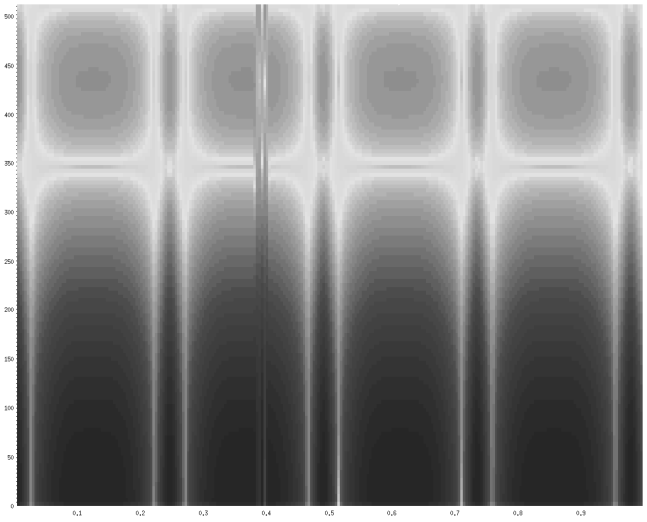

FIGURE 4.19. Spectrum

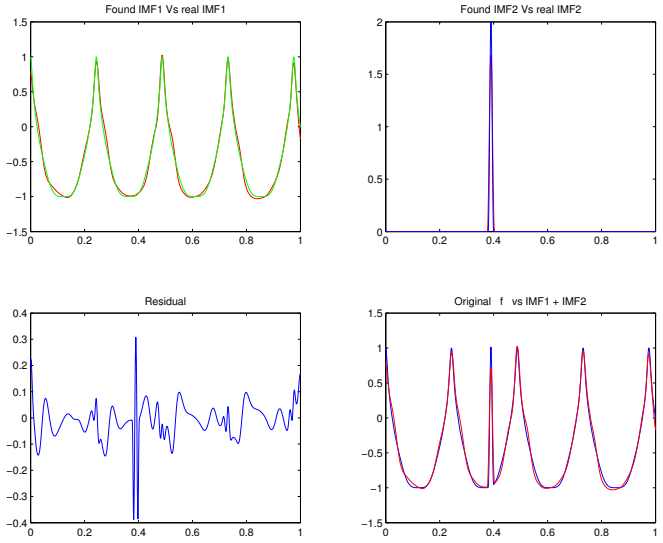

FIGURE 4.20. Results of the extraction. The red curves correspond to the extracted IMFs.

CHAPTER 5

Analysis of Convergence of Sparse Time-Frequency Method

So far in this thesis, we have concentrated on the algorithms and numerical methods. It is now time to prove why some of these algorithms work. In this chapter, we prove the convergence of the periodic STFR method. This proof is useful both for the case of ordinary IMF and intrawave extractions. This proof is based on our recent results in [26].

5.1. Convergence Analysis

In this section, we prove that for any signal, whether intrawave or not, increasing the filter span reduces the extraction error. We will show that the STFR method will converge to an IMF that is close to one of the IMF representations, but with an error associated with the width (span) of the filter. Notation for this section is listed in Table (5.1).

In this section, we assume that an IS can be represented in the following format:

$$f(t) = f_0(t) + f_1(t)\cos\theta(t), \tag{5.1.1}$$

for $f_1(t) > 0$, $\theta' > 0$ and $t \in [0,1]$. We assume that the signal is periodic with mean zero. In fact, even if the signal $f(t)$ is periodic with a non-zero mean, we can redefine it by reducing the mean of the signal from the signal itself. Lemma 2 helps us to bound the n^{th} derivative of the m^{th} approximated phase function $\bar{\theta}^m$ and is later used to bound some of the integrals in the main theorem.

Lemma 2. *If* $\left(\bar{\theta}^m\right)' \in V_{M_0}$, *where* $V_{M_0} = span\left\{e^{i2\pi kt}, |k| \le M_0, k \in \mathbb{Z}\right\}$ *then*

$$\left|\left(\bar{\theta}^m\right)^{(n)}(t)\right| \le (2\pi M_0)^{n-1}\left\|\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right\|_1, \tag{5.1.2}$$

Domain	Symbol	Analysis	Synthesis
θ	$\mathcal{F}_\theta(.)$	$(\mathcal{F}_\theta(g))_k = \frac{\int_{\theta_0}^{\theta_{end}} g(\theta)\exp\left(\frac{-i2\pi k\theta}{\lvert\theta_{end}-\theta_0\rvert}\right)d\theta}{\lvert\theta_{end}-\theta_0\rvert}$	$g(\theta) = \sum_{k=-\infty}^{\infty}(\mathcal{F}_\theta(g))_k \exp\left(\frac{i2\pi k\theta}{\lvert\theta_{end}-\theta_0\rvert}\right)$
$t \in [0,1]$	$\mathcal{F}(.)$	$(\mathcal{F}(g))_k = \int_0^1 g(t)\, e^{-i2\pi kt}dt$	$g(t) = \sum_{k=-\infty}^{\infty}(\mathcal{F}(g))_k\, e^{i2\pi kt}$
$\bar{\theta}$	$\hat{()}_{\bar{\theta}}$	$\hat{f}_{\bar{\theta}}(k) = \int_0^1 f\left(\bar{\theta}\right) e^{-i2\pi k\bar{\theta}}d\bar{\theta}$	$f\left(\bar{\theta}\right) = \sum_{k=-\infty}^{\infty}\hat{f}_{\bar{\theta}}(k)\, e^{i2\pi k\bar{\theta}}$
t	$P_{V_{M_0}}(.)$	$P_{V_{M_0}}(g) = \sum_{k=-M_0}^{M_0}(\mathcal{F}(g))_k\, e^{i2\pi kt}$	$R_{V_{M_0}}(g) = \sum_{\lvert k\rvert>M_0}(\mathcal{F}(g))_k\, e^{i2\pi kt}$
$k \in \mathbb{Z}$	$\lVert.\rVert_{1,M_0}$	$\lVert z\rVert_{1,M_0} = \sum_{\lvert k\rvert\le M_0}\lvert z(k)\rvert$	

TABLE 5.1. Coordinates and Symbols

for $n \in \mathbb{N}$.

PROOF. By definition, we have

$$\begin{aligned}
\left|\left(\bar{\theta}^m\right)^{(n)}(t)\right| &= \left|\sum_{|k|\le M_0}\left(\mathcal{F}\left(\left(\bar{\theta}^m\right)^{(n)}\right)\right)_k e^{i2\pi kt}\right| \\
&\le \sum_{|k|\le M_0}\left|\left(\mathcal{F}\left(\left(\bar{\theta}^m\right)^{(n)}\right)\right)_k\right| \\
&\le \sum_{|k|\le M_0}\left|\left(\mathcal{F}\left(\left(\bar{\theta}^m\right)^{(n)}\right)\right)_k\right| \\
&= \sum_{|k|\le M_0}|i2\pi k|^{n-1}\left|\left(\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right)_k\right| \\
&= (2\pi M_0)^{n-1}\left\|\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right\|_1 .
\end{aligned}$$

□

Lemma 2 leads to Lemma 3, which bounds integrals like $\left|\int_0^1 e^{i\delta\triangle\theta}e^{-i\epsilon\bar{\theta}^m}d\bar{\theta}^m\right|$ that occur frequently in the main theorem. These integrals would be bounded by the norm of the Fourier transform of the phase correction $\triangle\theta'$. In fact, this bound would help us construct a contraction in the main theorem.

Lemma 3. *If* $\left(\bar{\theta}^m\right)' > 0$, $t \in [0,1]$, $\bar{\theta}^m(0) = 0$, $\bar{\theta}^m(1) = 1$, *and* $\left(\bar{\theta}^m\right)', \triangle\theta' \in V_{M_0}$. *Also if* $e^{i\delta\triangle\theta}e^{-i\epsilon\bar{\theta}^m}$ *is periodic, then for* $\epsilon \neq 0$, *we have*

$$\left|\int_0^1 e^{i\delta\triangle\theta}e^{-i\epsilon\bar{\theta}^m}d\bar{\theta}^m\right| \leq \frac{P_m^n M_0^n}{|\epsilon|^n}\sum_{j=1}^{n}|\delta|^j\left(2\pi M_0\right)^{-j}\left(\|\mathcal{F}\left(\triangle\theta'\right)\|_1\right)^j. \tag{5.1.3}$$

Proof. Using integration by parts, we have

$$\int_0^1 e^{i\delta\triangle\theta}e^{-i\epsilon\bar{\theta}^m}d\bar{\theta}^m = \frac{1}{(i\epsilon)^n}\int_0^1\left(\frac{d^n}{d\left(\bar{\theta}^m\right)^n}e^{i\delta\triangle\theta}\right)e^{-i\epsilon\bar{\theta}^m}d\bar{\theta}^m.$$

Now, using Lemma 2, we have

$$\begin{aligned}\left|\int_0^1 e^{i\delta\triangle\theta}e^{-i\epsilon\bar{\theta}^m}d\bar{\theta}^m\right| \leq& \left|\frac{1}{|\epsilon|^n}\int_0^1\left(\frac{d^n}{d\left(\bar{\theta}^m\right)^n}e^{i\delta\triangle\theta}\right)d\bar{\theta}^m\right| \\ \leq& \frac{P\left(\frac{\left\|\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right\|_1}{\min\left(\bar{\theta}^m\right)'}, n\right)M_0^n}{\left(\min\left(\bar{\theta}^m\right)'\right)^n|\epsilon|^n}\sum_{j=1}^{n}|\delta|^j\left(2\pi M_0\right)^{-j}\left(\|\mathcal{F}\left(\triangle\theta'\right)\|_1\right)^j \\ =& \frac{P_m^n M_0^n}{|\epsilon|^n}\sum_{j=1}^{n}|\delta|^j\left(2\pi M_0\right)^{-j}\left(\|\mathcal{F}\left(\triangle\theta'\right)\|_1\right)^j.\end{aligned}$$

Here $P(x,n)$ is a polynomial of degree $n-1$ and $P_m^n = P\left(\frac{\left\|\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right\|_1}{\min\left(\bar{\theta}^m\right)'}, n\right)$. This completes the proof. □

Remark 1. Here we present a simple calculation on how to compute the polynomial $P(x,n)$ for small n. For example for $n=2$, we have

$$\begin{aligned}\left|\frac{d^2}{d\left(\bar{\theta}^m\right)^2}e^{i\delta\triangle\theta}\right| =& \left|i\left(\frac{(\delta\triangle\theta)''}{\left(\left(\bar{\theta}^m\right)'\right)^2} - \frac{(\delta\triangle\theta)'\left(\bar{\theta}^m\right)''}{\left(\left(\bar{\theta}^m\right)'\right)^3} + i\frac{\left((\delta\triangle\theta)'\right)^2}{\left(\left(\bar{\theta}^m\right)'\right)^2}\right)e^{i\delta\triangle\theta}\right| \\ \leq& \left|\frac{(\delta\triangle\theta)''}{\left(\left(\bar{\theta}^m\right)'\right)^2}\right| + \left|\frac{(\delta\triangle\theta)'\left(\bar{\theta}^m\right)''}{\left(\left(\bar{\theta}^m\right)'\right)^3}\right| + \left|\frac{\left((\delta\triangle\theta)'\right)^2}{\left(\left(\bar{\theta}^m\right)'\right)^2}\right|\end{aligned}$$

$$\leq \frac{\max\left|(\delta\triangle\theta)''\right|}{\min\left(\left(\bar{\theta}^m\right)'\right)^2} + \frac{\max\left|(\delta\triangle\theta)'\right|\max\left|\left(\bar{\theta}^m\right)''\right|}{\min\left(\left(\bar{\theta}^m\right)'\right)^3} + \frac{\max\left|\left((\delta\triangle\theta)'\right)^2\right|}{\min\left(\left(\bar{\theta}^m\right)'\right)^2}$$

$$\leq \frac{\left[\left(1+\frac{\left\|\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right\|_1}{\min\left(\bar{\theta}^m\right)'}\right)2\pi M_0\left\|\mathcal{F}\left((\delta\triangle\theta)'\right)\right\|_1 + \left\|\mathcal{F}\left((\delta\triangle\theta)'\right)\right\|_1^2\right]}{\min\left(\left(\bar{\theta}^m\right)'\right)^2},$$

where we have used $\left(\bar{\theta}^m\right)', \triangle\theta' \in V_{M_0}$. In other words, we have $P(x,2) = K(x+1)$ for some positive constant K.

Here we present the convergence theorem. The essence of the algorithm is as follows. We try to construct a contraction iterative scheme on $\left\|\mathcal{F}(\theta-\theta^m)'\right\|_1$, where θ^m is the approximate value of θ at the m^{th} step. This contraction is built upon the error bounds of the extracted envelopes at each iteration of the algorithm. The notations of this proof follow the notations of Algorithm 3.

Theorem 5. *[Convergence Theorem] Assume that the instantaneous frequency in equation (5.1.1) is M_0-sparse; i.e. $\theta' \in V_{M_0}$. Furthermore, assume that*

$$\left|\hat{f}_{0,\bar{\theta}}(k)\right| \leq \frac{C_0}{|k|^p},\ \left|\hat{f}_{1,\bar{\theta}}(k)\right| \leq \frac{C_0}{|k|^p},$$

for $C_0 > 0$ and $p \geq 4$. If the initial guess satisfies $\frac{\left\|\mathcal{F}\left(\triangle\theta^0\right)'\right\|_1}{2\pi M_0} \leq \frac{1}{4}$, then there exists an $\eta_0 > 0$ such that for $L > \eta_0$ we have

$$\left\|\mathcal{F}(\theta-\theta^{m+1})'\right\|_1 \ \leq\ \Gamma_1\lambda^{2p-2}L^{-p+2} + \tfrac{1}{2}\left\|\mathcal{F}(\theta-\theta^m)'\right\|_1,$$

for $\lambda > 1$ and $\Gamma_1 > 0$.

Proof. We know that if $\triangle\theta^m = \theta - \theta^m$, then $a^m = f_1\cos\Delta\theta^m$, and $b^m = -f_1\sin\Delta\theta^m$. Let $\tilde{a}^m$, $\tilde{b}^m$ be approximate envelope functions. Set the error in envelopes as $\triangle a^m = a - \tilde{a}^m$, and $\triangle b^m = b - \tilde{b}^m$. Using these, we get $f = f_0 +$

$a^m \cos\theta^m + b^m \sin\theta^m$. Let $L_m = \frac{\theta^m(1)-\theta^m(0)}{2\pi}$ and $\bar{\theta}^m = \frac{\theta^m}{2\pi L_m}$. Then, we have $f = f_0 + a^m \cos 2\pi L_m \bar{\theta}^m + b^m \sin 2\pi L_m \bar{\theta}^m$. If we take the Fourier transform in $\bar{\theta}^m$ coordinate (See Table (5.1)), we get

$$\hat{f}_{\bar{\theta}^m}(k) = \hat{f}_{0,\bar{\theta}^m}(k) + \frac{\left(\hat{a}^m_{\bar{\theta}^m}(k-L_m) + \hat{a}^m_{\bar{\theta}^m}(k+L_m)\right)}{2} + \frac{\left(\hat{b}^m_{\bar{\theta}^m}(k-L_m) - \hat{b}^m_{\bar{\theta}^m}(k+L_m)\right)}{2i}.$$

Consequently, one can find

$$\begin{aligned}
\hat{a}^m_{\bar{\theta}^m}(k) &= \hat{f}_{\bar{\theta}^m}(k+L_m) + \hat{f}_{\bar{\theta}^m}(k-L_m) \\
&\quad -\hat{f}_{0,\bar{\theta}^m}(k+L_m) - \hat{f}_{0,\bar{\theta}^m}(k-L_m) \\
&\quad +\frac{1}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m) - \hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right) \\
&\quad +\frac{1}{2i}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m) - \hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right),
\end{aligned}$$

and

$$\begin{aligned}
\hat{b}^m_{\bar{\theta}^m}(k) &= i\hat{f}_{\bar{\theta}^m}(k+L_m) - i\hat{f}_{\bar{\theta}^m}(k-L_m) \\
&\quad -i\hat{f}_{0,\bar{\theta}^m}(k+L_m) + i\hat{f}_{0,\bar{\theta}^m}(k-L_m) \\
&\quad +\frac{i}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m) + \hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right) \\
&\quad +\frac{1}{2}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m) + \hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right).
\end{aligned}$$

In the periodic STFR algorithm, $\hat{\tilde{a}}^m_{\bar{\theta}^m}$, $\hat{\tilde{b}}^m_{\bar{\theta}^m}$ are approximated as

$$\hat{\tilde{a}}^m_{\bar{\theta}^m}(k) = \begin{cases} \hat{f}_{\bar{\theta}^m}(k+L_m) + \hat{f}_{\bar{\theta}^m}(k-L_m), & -\frac{L_m}{\lambda} \le k \le \frac{L_m}{\lambda}, \\ 0, & otherwise, \end{cases}$$

$$\hat{\tilde{b}}^m_{\bar{\theta}^m}(k) = \begin{cases} i\left(\hat{f}_{\bar{\theta}^m}(k+L_m) - \hat{f}_{\bar{\theta}^m}(k-L_m)\right), & -\frac{L_m}{\lambda} \le k \le \frac{L_m}{\lambda}, \\ 0, & otherwise. \end{cases}$$

Here, $1 < \lambda$ defines the width of the filter. Hence, for $\triangle\hat{a}^m_{\bar{\theta}^m}$ and $\triangle\hat{b}^m_{\bar{\theta}^m}$ we have

$$\triangle\hat{a}^m_{\bar{\theta}^m}(k) = \begin{cases} \{-\hat{f}_{0,\bar{\theta}^m}(k+L_m) - \hat{f}_{0,\bar{\theta}^m}(k-L_m) \\ +\frac{1}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m) - \hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right) & |k| \le \frac{L_m}{\lambda}, \\ +\frac{1}{2i}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m) - \hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right)\}, \\ \hat{a}^m_{\bar{\theta}^m}(k), & |k| > \frac{L_m}{\lambda}, \end{cases}$$

$$\triangle\hat{b}^m_{\bar{\theta}^m}(k) = \begin{cases} \{-i\hat{f}_{0,\bar{\theta}^m}(k+L_m) + i\hat{f}_{0,\bar{\theta}^m}(k-L_m) \\ +\frac{i}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m) + \hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right) & |k| \le \frac{L_m}{\lambda}, \\ +\frac{1}{2}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m) + \hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right)\}, \\ \hat{b}^m_{\bar{\theta}^m}(k), & |k| > \frac{L_m}{\lambda}. \end{cases}$$

The ideal case for updating the phase function is $\frac{d\theta_{new}}{dt} = \frac{d\theta^m}{dt} - \frac{d}{dt}\arctan\frac{\tilde{b}^m}{\tilde{a}^m}$. However, the algorithm works in a way that we must choose $\frac{d\theta^{m+1}}{dt}$ in V_{M_0}. Hence, $\frac{d\theta^{m+1}}{dt} = P_{V_{M_0}}\left(\frac{d\theta_{new}}{dt}\right)$. So, at each step, we force $\frac{d\theta^m}{dt}$ to be in V_{M_0}. In other words, $\frac{d\theta^m}{dt} \in V_{M_0}$ for all $m \ge 0$, $m \in \mathbb{Z}$. This short analysis tells us that $\frac{d\theta^{m+1}}{dt} = \frac{d\theta^m}{dt} - P_{V_{M_0}}\left(\frac{d}{dt}\arctan\frac{\tilde{b}^m}{\tilde{a}^m}\right)$. Since $\theta \in V_{M_0}$ is sufficiently differentiable, and $\frac{d\theta^m}{dt} \in V_{M_0}$, then $(\theta - \theta^m = \triangle\theta^m) \in V_{M_0}$. Having these in mind, we can find

$$\begin{aligned} \frac{d}{dt}\triangle\theta^{m+1} =& \frac{d}{dt}\left(\theta - \theta^{m+1}\right) = \frac{d\theta}{dt} - \frac{d\theta^m}{dt} + P_{V_{M_0}}\left(\frac{d}{dt}\arctan\frac{\tilde{b}^m}{\tilde{a}^m}\right) \\ =& P_{V_{M_0}}\frac{d}{dt}\left(\arctan\frac{\tilde{b}^m}{\tilde{a}^m} - \arctan\frac{b^m}{a^m}\right) + R_{V_{M_0}}\left(\frac{d}{dt}\triangle\theta^m\right). \end{aligned}$$

We know $\arctan\frac{\tilde{b}^m}{\tilde{a}^m} - \arctan\frac{b^m}{a^m}$ is in $\mathcal{C}^1$. For any $g \in \mathcal{C}^1$, we have $P_{V_{M_0}}\frac{d}{dt}(g) = \frac{d}{dt}P_{V_{M_0}}(g)$. Hence, the Fourier transform of the IF error is

$$\begin{aligned} \left(\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right)_k =& \left(\mathcal{F}\left(P_{V_{M_0}}\frac{d}{dt}\left(\arctan\frac{\tilde{b}^m}{\tilde{a}^m} - \arctan\frac{b^m}{a^m}\right)\right)\right)_k \\ &+ \left(\mathcal{F}\left(R_{V_{M_0}}\left(\triangle\theta^m\right)'\right)\right)_k \\ &= (i2\pi k)\left(\mathcal{F}\left(P_{V_{M_0}}\left(\arctan\frac{\tilde{b}^m}{\tilde{a}^m} - \arctan\frac{b^m}{a^m}\right)\right)\right)_k \\ &+ \left(\mathcal{F}\left(R_{V_{M_0}}\left(\triangle\theta^m\right)'\right)\right)_k. \end{aligned}$$

As $\left(\mathcal{F}\left(P_{V_{M_0}}\left(\arctan\frac{\tilde{b}^m}{\tilde{a}^m}-\arctan\frac{b^m}{a^m}\right)\right)\right)_k = 0$ for $|k|>M_0$, then

$$\begin{aligned}\left\|\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right\|_1 &\leq (2\pi M_0)\left\|\mathcal{F}\left(P_{V_{M_0}}\left(\arctan\frac{\tilde{b}^m}{\tilde{a}^m}-\arctan\frac{b^m}{a^m}\right)\right)\right\|_1 \\ &\quad+\left\|\mathcal{F}\left(R_{V_{M_0}}\left(\triangle\theta^m\right)'\right)\right\|_1.\end{aligned}$$

Now, we know that for any function $g(t)$, we have

$$\begin{aligned}\left\|\mathcal{F}\left(P_{V_{M_0}}(g)\right)\right\|_1 &= \sum_{k=-M_0}^{M_0}\left|\left(\mathcal{F}(g)\right)_k\right| = \sum_{k=-M_0}^{M_0}\left|\int_0^1 g(t)\,e^{-i2\pi kt}dt\right| \\ &\leq \sum_{k=-M_0}^{M_0}\int_0^1|g(t)|\,dt \leq \sum_{k=-M_0}^{M_0}\|g\|_\infty = (2M_0+1)\|g\|_\infty.\end{aligned}$$

Hence, the l_1 norm of the Fourier of the IF error is

$$\begin{aligned}\left\|\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right\|_1 &\leq (2\pi M_0)(2M_0+1)\left\|\arctan\frac{\tilde{b}^m}{\tilde{a}^m}-\arctan\frac{b^m}{a^m}\right\|_\infty \\ &\quad+\left\|\mathcal{F}\left(R_{V_{M_0}}\left(\triangle\theta^m\right)'\right)\right\|_1.\end{aligned}$$

Since θ' is sparse in V_{M_0}, then the last term vanishes. In fact, we previously showed that $\frac{d\theta^m}{dt}\in V_{M_0}$, hence

$$\left(\mathcal{F}\left(\triangle\theta^m\right)'\right)_k = \left(\mathcal{F}\left(\theta-\theta^m\right)'\right)_k = \begin{cases}\left(\mathcal{F}\left(\theta'\right)\right)_k-\left(\mathcal{F}\left(\left(\theta^m\right)'\right)\right)_k, & |k|\leq M_0,\\ \left(\mathcal{F}\left(\theta'\right)\right)_k=0, & |k|>M_0,\end{cases}$$

so,

$$R_{V_{M_0}}\left(\triangle\theta^m\right)' = \sum_{|k|>M_0}\left(\mathcal{F}\left(\triangle\theta^m\right)'\right)_k e^{i2\pi kt} = \sum_{|k|>M_0}\left(\mathcal{F}\left(\theta'\right)\right)_k e^{i2\pi kt} = 0,$$

and then $\left(\mathcal{F}\left(R_{V_{M_0}}\left(\triangle\theta^m\right)'\right)\right)_k = 0$. Finally, we have the following bound on the IF error:

$$\left\|\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right\|_1 \leq (2\pi M_0)(2M_0+1)\left\|\arctan\frac{\tilde{b}^m}{\tilde{a}^m}-\arctan\frac{b^m}{a^m}\right\|_\infty.$$

The last term on the right hand side of this bound can be simplified further. If we use the fact that $x^2 + xy \leq \frac{x^2}{2} - y^2$ for real x and y, we have

$$\begin{aligned}\left|\arctan\frac{\tilde{b}^m}{\tilde{a}^m} - \arctan\frac{b^m}{a^m}\right| &= \left|\arctan\left(\frac{\tilde{b}^m a^m - \tilde{a}^m b^m}{a^m\tilde{a}^m + b^m\tilde{b}^m}\right)\right| \leq \left|\frac{\tilde{b}^m a^m - \tilde{a}^m b^m}{a^m\tilde{a}^m + b^m\tilde{b}^m}\right| \\ &= \left|\frac{\left(a^m + \triangle a^m\right)\triangle b^m - \left(b^m + \triangle b^m\right)\triangle a^m}{\left(a^m\right)^2 + \left(b^m\right)^2 + \left(\triangle a^m\right)a^m + \left(\triangle b^m\right)b^m}\right| \\ &\leq \frac{\left(|a^m| + |\triangle a^m|\right)|\triangle b^m| + \left(|b^m| + |\triangle b^m|\right)|\triangle a^m|}{\frac{\left(a^m\right)^2+\left(b^m\right)^2}{2} - \left(\left(\triangle a^m\right)^2 + \left(\triangle b^m\right)^2\right)} \\ &\leq D\left(|\triangle a^m| + |\triangle b^m|\right),\end{aligned}$$

for $D = \max\left(\frac{f_1+|\triangle a^m|}{\frac{f_1^2}{2}-\left((\triangle a^m)^2+(\triangle b^m)^2\right)}, \frac{f_1+|\triangle b^m|}{\frac{f_1^2}{2}-\left((\triangle a^m)^2+(\triangle b^m)^2\right)}\right)$, in which we have taken into account that $f_1 > 0$. Now, consider the fact that

$$\begin{aligned}|\triangle a^m(t)| &= |\triangle a^m(\theta)| = \left|\triangle a^m\left(\bar{\theta}^m\right)\right| = \left|\sum_{k=-\infty}^{\infty}\triangle\hat{a}^m_{\bar{\theta}^m}(k)\, e^{i2\pi k\bar{\theta}^m}\right| \\ &\leq \sum_{k=-\infty}^{\infty}|\triangle\hat{a}^m_{\bar{\theta}^m}(k)| = \left\|\triangle\hat{a}^m_{\bar{\theta}^m}\right\|_1,\end{aligned}$$

and as, in general $\hat{g}_{\bar{\theta}}(-k) = \overline{\hat{g}_{\bar{\theta}}(k)}$ for any g, then $|\hat{g}_{\bar{\theta}}(-k)| = |\hat{g}_{\bar{\theta}}(k)|$. Consequently, the bound on the envelopes errors can be expressed as

$$\begin{aligned}|\triangle a^m| &\leq \left\|\triangle\hat{a}^m_{\bar{\theta}^m}\right\|_1 \\ &\leq 2\sum_{\left(1-\frac{1}{\lambda}\right)L_m\leq k\leq\left(1+\frac{1}{\lambda}\right)L_m}\left|\hat{f}_{0,\bar{\theta}^m}(k)\right| \\ &+ \sum_{\left(2-\frac{1}{\lambda}\right)L_m\leq k\leq\left(2+\frac{1}{\lambda}\right)L_m}\left(|\hat{a}^m_{\bar{\theta}^m}(k)| + \left|\hat{b}^m_{\bar{\theta}^m}(k)\right|\right) \\ &+ \sum_{|k|>\frac{L_m}{\lambda}}|\hat{a}^m_{\bar{\theta}^m}(k)|,\end{aligned}$$

and

$$\begin{aligned}
|\triangle b^m| &\leq \left\|\triangle \hat{b}^m_{\bar{\theta}^m}\right\|_1 \\
&\leq 2 \sum_{\left(1-\frac{1}{\lambda}\right)L_m \leq k \leq \left(1+\frac{1}{\lambda}\right)L_m} \left|\hat{f}_{0,\bar{\theta}^m}(k)\right| \\
&+ \sum_{\left(2-\frac{1}{\lambda}\right)L_m \leq k \leq \left(2+\frac{1}{\lambda}\right)L_m} \left(\left|\hat{a}^m_{\bar{\theta}^m}(k)\right| + \left|\hat{b}^m_{\bar{\theta}^m}(k)\right|\right) \\
&+ \sum_{|k|>\frac{L_m}{\lambda}} \left|\hat{b}^m_{\bar{\theta}^m}(k)\right|.
\end{aligned}$$

Before we use approximations on these terms, recall that according to one of our assumptions, we have: "The observation is periodic with mean zero." Hence, $\hat{f}_{0,\bar{\theta}^m}(0) = 0$. Furthermore, as $e^{i2\pi k\bar{\theta}}e^{-i2\pi\omega\bar{\theta}^m}$ is periodic, so is $e^{i2\pi\bar{\theta}^m\left(k\frac{L_m}{L}-\omega\right)}e^{ik\frac{\triangle\theta^m}{L}}$, and then we have the following estimate using these facts and Lemma 3:

$$\begin{aligned}
\left|\hat{f}_{0,\bar{\theta}^m}(\omega)\right| \leq\ & 2C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} \\
& +2C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\left(\sum_{j=1}^{p-2}\left(\frac{\gamma}{L}\right)^j\frac{\pi^2}{3}+2\sum_{j=p}^{n}\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^j\right) \\
& +2C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\left(2\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right)\right).
\end{aligned} \tag{5.1.4}$$

Here, $\gamma = \frac{\left\|\mathcal{F}(\triangle\theta^m)'\right\|_1}{2\pi M_0}$. The proof of this inequality can be found in in Appendix A. To find the estimate on $\hat{a}^m_{\bar{\theta}^m}(k)$, we follow a similar approach. The approximation procedure is detailed in Appendix B.

$$\begin{aligned}
|\hat{a}^m_{\bar{\theta}^m}(\omega)| \leq & 2C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} + \left|\hat{f}_{1,\bar{\theta}}(0)\right|P_m^n\left(\frac{M_0}{2\pi|\omega|}\right)^n\sum_{j=1}^{n}\gamma^j \\
& + C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\frac{2p}{p-1}\sum_{j=1}^{n}2^j\gamma^j \\
& + C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\left(\sum_{j=1}^{p-2}2^j\left(\frac{\gamma}{L}\right)^j\frac{\pi^2}{3}+\sum_{j=p}^{n}2^{j+1}\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^j\right) \\
& + C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\left(2^p\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right)\right).
\end{aligned} \tag{5.1.5}$$

There is a similar bound for $\hat{b}^m_{\bar{\theta}^m}(\omega)$ as well:

$$\begin{aligned}\left|\hat{b}^m_{\bar{\theta}^m}(\omega)\right| \leq & 2C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} + \left|\hat{f}_{1,\bar{\theta}}(0)\right| P_m^n\left(\frac{M_0}{2\pi|\omega|}\right)^n \sum_{j=1}^{n}\gamma^j \\ & + C_0 P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \frac{2p}{p-1}\sum_{j=1}^{n}2^j\gamma^j \\ & + C_0 P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \left(\sum_{j=1}^{p-2}2^j\left(\frac{\gamma}{L}\right)^j\frac{\pi^2}{3} + \sum_{j=p}^{n}2^{j+1}\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^j\right) \\ & + C_0 P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \left(2^p\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right)\right).\end{aligned}$$

Now we approximate the terms in the inequality for $|\triangle a^m|$. One can find the detail of this approximation in Appendix C. We have a similar inequality for $|\triangle b^m|$. These can be expressed in the following way

$$|\triangle a^m| \leq C_1\lambda^{2p-2}L^{-p+2} + C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma\ , \tag{5.1.6}$$

$$|\triangle b^m| \leq C_1\lambda^{2p-2}L^{-p+2} + C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma\ , \tag{5.1.7}$$

where C_1 and C_2 depend on C_0, M_0, ε, α, P_m^n, n, p. Before moving forward, we need to show that parameters like α and P_m^n are uniformly bounded. We start with α. We have

$$\begin{aligned}|1-\alpha| = & \left|1-\frac{L_m}{L}\right| = \left|1-\frac{\frac{\theta^m(1)-\theta^m(0)}{2\pi}}{\frac{\bar{\theta}^m(1)-\bar{\theta}^m(0)}{2\pi}}\right| = \left|\frac{\Delta\theta^m(1)-\Delta\theta^m(0)}{2\pi L}\right| \\ \leq & \left\|\frac{(\Delta\theta^m)'}{2\pi L}\right\|_\infty \leq \left\|\frac{\mathcal{F}(\Delta\theta^m)'}{2\pi L}\right\|_1 = \frac{\gamma M_0}{L} \leq \frac{M_0}{4L} \leq \frac{1}{8}.\end{aligned} \tag{5.1.8}$$

Here, we have used the fact that $\gamma \leq \frac{1}{4}$. At the last part of the proof, we will show that this condition remains intact for all iterations. The last inequality in (5.1.8) can be true for the ***condition*** $L \geq 2M_0$. Hence, $\frac{7}{8} \leq \alpha \leq \frac{9}{8}$. In order to prove the boundedness of P_m^n at every step, we need to find bounds on $\left\|\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right\|_1$ and $\min\left(\bar{\theta}^m\right)'$. If we take the ***condition*** $\min\left(\bar{\theta}'\right) \geq \frac{M_0}{2L}$, we get $\bar{\theta}' \geq \frac{M_0}{2L} > \frac{M_0}{4L} \geq \frac{\gamma M_0}{L} =$

$\left\|\frac{\mathcal{F}(\Delta\theta^m)'}{2\pi L}\right\|_1 \geq \left|\frac{(\Delta\theta^m)'}{2\pi L}\right|$. Using this, we have

$$\begin{aligned}\left|\left(\bar{\theta}^m\right)'\right| =& \left|\frac{(\theta^m)'}{2\pi L_m}\right| = \left|\frac{(\theta-\Delta\theta^m)'}{2\pi\alpha L}\right| \geq \frac{8}{9}\left|\frac{(\theta-\Delta\theta^m)'}{2\pi L}\right| \\ =& \frac{8}{9}\left|\bar{\theta}' - \frac{(\Delta\theta^m)'}{2\pi L}\right| \geq \frac{8}{9}\left(\bar{\theta}' - \frac{(\Delta\theta^m)'}{2\pi L}\right) \geq \frac{8}{9}\left(\bar{\theta}' - \frac{M_0}{4L}\right) \\ \geq& \frac{8}{9}\left(\bar{\theta}' - \frac{M_0}{4L}\right) \geq \frac{8}{9}\left(\bar{\theta}' - \frac{\bar{\theta}'}{2}\right) \geq \frac{4}{9}\min\left(\bar{\theta}'\right).\end{aligned}$$

We also have

$$\begin{aligned}\left\|\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right\|_1 =& \frac{1}{\alpha}\left\|\mathcal{F}\left(\frac{(\theta-\Delta\theta^m)'}{2\pi L}\right)\right\|_1 = \frac{1}{\alpha}\left\|\mathcal{F}\left(\bar{\theta}'\right) - \frac{\mathcal{F}\left((\Delta\theta^m)'\right)}{2\pi L}\right\|_1 \\ \leq& \frac{8}{7}\left(\left\|\mathcal{F}\left(\bar{\theta}'\right)\right\|_1 + \frac{\left\|\mathcal{F}\left((\Delta\theta^m)'\right)\right\|_1}{2\pi L}\right) = \frac{8}{7}\left(\left\|\mathcal{F}\left(\bar{\theta}'\right)\right\|_1 + \frac{M_0\gamma}{L}\right) \\ \leq& \frac{8}{7}\left(\left\|\mathcal{F}\left(\bar{\theta}'\right)\right\|_1 + \frac{M_0}{4L}\right).\end{aligned}$$

These two estimates pave the way to rigorously prove that P_m^n is bounded at every single step. Now it is time for a bound on D. In fact, if we bound D uniformly over all steps, we can set a contraction. Assuming that $1 < \lambda < \lambda_0$, and since $f_1 > 0$, taking the ***condition***

$$\frac{C_1\lambda_0^{2p-2}}{L^{-p+1}} + \frac{C_2\lambda_0^{2n}}{4L^{-\max(-n,-2p+1,-n-p+2,-2p)}} \leq \frac{\sqrt{2}}{4}\min f_1,$$

would result in[1]

$$\begin{aligned}|\triangle a^m| \leq& C_1\lambda^{2p-2}L^{-p+2} + C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma \\ \leq& \frac{C_1\lambda_0^{2p-2}}{L^{p-2}} + \frac{C_2\lambda_0^{2n}}{4L^{-\max(-n,-2p+1,-n-p+2,-2p)}} \leq \frac{\sqrt{2}}{4}\min f_1.\end{aligned}$$

[1]Remember that if f_1 has a zero crossing, this condition would never be satisfied. Hence, the rest of the proof will not be valid.

Similarly, we have the same bound for $|\triangle b^m|$, and then

$$\begin{aligned} D =& \max\left(\frac{f_1+|\triangle a^m|}{\frac{f_1^2}{2}-\left(\left(\triangle a^m\right)^2+\left(\triangle b^m\right)^2\right)}, \frac{f_1+|\triangle b^m|}{\frac{f_1^2}{2}-\left(\left(\triangle a^m\right)^2+\left(\triangle b^m\right)^2\right)}\right) \\ \leq& \frac{\max f_1+\frac{\sqrt{2}}{4}\min f_1}{\min\left(\frac{f_1^2}{2}-\left(\left(\triangle a^m\right)^2+\left(\triangle b^m\right)^2\right)\right)} \leq \frac{\max f_1+\frac{\sqrt{2}}{4}\min f_1}{\min\frac{f_1^2}{2}-\frac{1}{4}\min f_1^2} \\ =& \frac{4\max f_1+\sqrt{2}\min f_1}{\min f_1^2} = E_0. \end{aligned}$$

The latter shows that D is also bounded. So, we get

$$\begin{aligned} \left\|\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right\|_1 \leq& \left(2\pi M_0\right)\left(2M_0+1\right)\left\|\arctan\frac{\tilde{b}^m}{\tilde{a}^m}-\arctan\frac{b^m}{a^m}\right\|_\infty \\ \leq& \left(2\pi M_0\right)\left(2M_0+1\right)D\left(|\triangle a^m|+|\triangle b^m|\right) \\ \leq& E_0\left(4\pi M_0\right)\left(2M_0+1\right)C_1\lambda^{2p-2}L^{-p+2} \\ &+ E_0\left(4\pi M_0\right)\left(2M_0+1\right)C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma. \end{aligned}$$

The last inequality is nothing but

$$\begin{aligned} \left\|\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right\|_1 \leq& \Gamma_1\lambda^{2p-2}L^{-p+2} \\ &+\Gamma_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\left\|\mathcal{F}\left(\triangle\theta^{m}\right)'\right\|_1. \end{aligned}$$

If we have the ***condition*** that $\Gamma_2\lambda_0^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)} \leq \frac{1}{2}$, we have the contraction that we were looking for. Before finishing, we need to state the following: $\exists\eta_0 > 0$ such that $L > \eta_0$, then all the conditions $L \geq 2M_0$, $\min\left(\bar{\theta}'\right) \geq \frac{M_0}{2L}$, $\frac{C_1\lambda_0^{2p-2}}{L^{p-2}}+\frac{C_2\lambda_0^{2n}}{4L^{-\max(-n,-2p+1,-n-p+2,-2p)}} \leq \frac{\sqrt{2}}{4}\min f_1$, and $\Gamma_2\lambda_0^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)} \leq \frac{1}{2}$ would be satisfied. As a result, we have

$$\left\|\mathcal{F}\left(\theta-\theta^{m+1}\right)'\right\|_1 \leq \Gamma_1\lambda^{2p-2}L^{-p+2}+\frac{1}{2}\left\|\mathcal{F}\left(\theta-\theta^{m}\right)'\right\|_1.$$

Having this bound, the condition $\gamma \leq \frac{1}{4}$ would remain intact for all iterations. In other words, when there is a contraction on $\left\|\mathcal{F}\left(\theta-\theta^{m+1}\right)'\right\|_1$, this term would remain

bounded. Hence, if $\frac{\left\|\mathcal{F}\left(\triangle\theta^0\right)'\right\|_1}{2\pi M_0} \leq \frac{1}{4}$ for the first iteration, it will remain bounded by $\frac{1}{4}$ for all iterations. This completes the proof. □

A corollary of this is that if f_0 has only high-frequency components, there would be no interference between f_0 and $f_1 \cos\theta$. Hence, the convergence proof would be true for $\lambda > 0$, as well. Furthermore, since it is possible that an Intrinsic Signal has multiple representations and this theorem states merely that the algorithm converges to an IMF in one of these representations, the algorithm's result is not necessary unique. The theorem does not mention which representation the algorithm converges to. The theorem that we observed, in this part, says that if in a representation we have $\left|\hat{f}_{0,\bar{\theta}}(k)\right| \leq \frac{C_0}{|k|^p}$, $\left|\hat{f}_{1,\bar{\theta}}(k)\right| \leq \frac{C_0}{|k|^p}$, then increasing the width of the filter will reduce the error in extraction. More specifically, if in a representation we have a wide band signal (like an intrawave signal), we are more likely to capture it by widening the filter width. This approach helps us to find a sparser representation compared to the case in which one uses a normal envelope dictionary 1.2.3. From the algorithmic point of view, the initial guess θ_0 and the parameter λ define the representation to which we converge. In practice, if the IMFs that constitute the signal are separated enough in the time-frequency domain, the reduction of the parameter λ is always beneficial. This idea is illustrated in the following example.

Example 15. Again consider our generic intrawave IMF $x(t) = \cos\left(\omega t + \frac{\triangle\omega}{p}\sin(pt)\right)$. We can have many representations for this signal. The first representation would be

$$x(t) = \cos\left(\omega t + \frac{\triangle\,\omega}{p}\sin(pt)\right), \tag{5.1.9}$$

for $\bar{\theta} = \omega t + \frac{\triangle\omega}{p}\sin(pt)$. The second representation would be

$$\begin{aligned} x(\theta) = &\left\{J_0\left(\frac{\triangle\,\omega}{p}\right) + 2\sum_{k=1}^{\infty} J_{2k}\left(\frac{\triangle\,\omega}{p}\right)\cos(2kpt)\right\}\cos(\omega t) \\ &- \left\{2\sum_{k=1}^{\infty} J_{2k-1}\left(\frac{\triangle\,\omega}{p}\right)\sin((2k-1)\,pt)\right\}\sin(\omega t), \end{aligned}$$

for $\bar{\theta} = \omega t$. So, with a constant initial guess for θ_0 close to $\bar{\theta} = \omega t$, the second representation is the one that is seen by the algorithm at the first iteration. Hence, if a narrow-band filter (large value of λ) is used, only the first harmonic is extracted by the algorithm. In order to capture more terms in the envelope of the second representation, we need a smaller value of λ.

5.2. Recovery of Signals Polluted by Noise

The next theorem shows that if the observation is polluted by noise, under mild presence of noise, the algorithm still converges. Take

$$f(t) = f_0(t) + f_1(t)\cos\theta(t) + \Im, \tag{5.2.1}$$

where $\Im$ is a periodic perturbation to the original signal (5.1.1).

Theorem 6. *[Convergence Theorem in the Presence of Noise] Assume that the instantaneous frequency in equation (5.2.1) is M_0-sparse; i.e. $\theta' \in V_{M_0}$. Furthermore, assume that*

$$\left|\hat{f}_{0,\bar{\theta}}(k)\right| \leq \frac{C_0}{|k|^p},\ \left|\hat{f}_{1,\bar{\theta}}(k)\right| \leq \frac{C_0}{|k|^p},$$

for $C_0 > 0$ and $p \geq 4$. If the initial guess satisfies $\frac{\left\|\mathcal{F}\left(\triangle\theta^0\right)'\right\|_1}{2\pi M_0} \leq \frac{1}{4}$, then there exists an $\eta_0 > 0$ such that for $L > \eta_0$, and $\|\Im\|_\infty \leq \epsilon_0$ (ϵ_0 sufficiently small) we have

$$\left\|\mathcal{F}\left(\theta - \theta^{m+1}\right)'\right\|_1 \ \leq\ \Upsilon_0(L,\lambda)\|\Im(t)\|_\infty + \Gamma_1\lambda^{2p-2}L^{-p+2} + \tfrac{1}{2}\left\|\mathcal{F}\left(\theta - \theta^m\right)'\right\|_1,$$

for $\lambda > 1$, $\Gamma_1 > 0$ and $\Upsilon_0(L,\lambda)$.

Proof. The proof of this theorem is essentially the same as the proof of the Convergence Theorem in the absence of noise. There is only one minor change. Using the

same trend of proof as before, we have

$$\triangle\hat{a}^m_{\bar{\theta}^m}(k)=\begin{cases}\{-\hat{\Im}_{\bar{\theta}^m}(k+L_m)-\hat{\Im}_{\bar{\theta}^m}(k-L_m)\\ -\hat{f}_{0,\bar{\theta}^m}(k+L_m)-\hat{f}_{0,\bar{\theta}^m}(k-L_m)\\ +\frac{1}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m)-\hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right)\\ +\frac{1}{2i}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m)-\hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right)\}, & |k|\le\frac{L_m}{\lambda},\\ \hat{a}^m_{\bar{\theta}^m}(k), & |k|>\frac{L_m}{\lambda},\end{cases}$$

$$\triangle\hat{b}^m_{\bar{\theta}^m}(k)=\begin{cases}\{-i\hat{\Im}_{\bar{\theta}^m}(k+L_m)+i\hat{\Im}_{\bar{\theta}^m}(k-L_m)\\ -i\hat{f}_{0,\bar{\theta}^m}(k+L_m)+i\hat{f}_{0,\bar{\theta}^m}(k-L_m)\\ +\frac{i}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m)+\hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right)\\ +\frac{1}{2}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m)+\hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right)\}, & |k|\le\frac{L_m}{\lambda},\\ \hat{b}^m_{\bar{\theta}^m}(k), & |k|>\frac{L_m}{\lambda}.\end{cases}$$

Here, in simple words, we have $\triangle a^m(t)=\Im^m_{a,trunc}(t)+\triangle a^m_{old}(t)$, where $\Im^m_{a,trunc}\left(\bar{\theta}^m\right)$ is the truncated part of $\Im$ in a:

$$\begin{aligned}\Im^m_{a,trunc}\left(\bar{\theta}^m\right) &= \sum_{k=-\infty}^{\infty}\hat{\Im}^m_{a,trunc,\bar{\theta}^m}(k)\,e^{i2\pi k\bar{\theta}^m}\\ &= \sum_{|k|\le\frac{L_m}{\lambda}}\left(-\hat{\Im}_{\bar{\theta}^m}(k+L_m)-\hat{\Im}_{\bar{\theta}^m}(k-L_m)\right)e^{i2\pi k\bar{\theta}^m}.\end{aligned}$$

Also, $\triangle a^m_{old}(t)$, in Fourier domain, is defined as

$$\triangle\hat{a}^m_{old,\bar{\theta}^m}(k)=\begin{cases}\{-\hat{f}_{0,\bar{\theta}^m}(k+L_m)-\hat{f}_{0,\bar{\theta}^m}(k-L_m)\\ +\frac{1}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m)-\hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right)\\ +\frac{1}{2i}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m)-\hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right)\}, & |k|\le\frac{L_m}{\lambda},\\ \hat{a}^m_{\bar{\theta}^m}(k), & |k|>\frac{L_m}{\lambda}.\end{cases}$$

This formula is telling us that we only need to find a bound on $\Im^m_{trunc}(t)$, and then use the results of the Theorem 5 to finalize the proof. One can simplify $\Im^m_{a,trunc}\left(\bar{\theta}^m\right)$

, using

$$\chi_\lambda(k) = \begin{cases} 1, & |k| \leq \frac{L_m}{\lambda}, \\ 0, & |k| > \frac{L_m}{\lambda}, \end{cases}$$

as

$$\begin{aligned} \Im^m_{a,trunc}\left(\bar{\theta}^m\right) = & \sum_{k=-\infty}^{\infty} \left(-\hat{\Im}_{\bar{\theta}^m}(k+L_m) - \hat{\Im}_{\bar{\theta}^m}(k-L_m)\right) \chi_\lambda(k)\, e^{i2\pi k\bar{\theta}^m} \\ = & e^{-i2\pi L_m\bar{\theta}^m} \sum_{k=-\infty}^{\infty} \left(-\hat{\Im}_{\bar{\theta}^m}(k)\right) \chi_\lambda(k-L_m)\, e^{i2\pi k\bar{\theta}^m} \\ & + e^{i2\pi L_m\bar{\theta}^m} \sum_{k=-\infty}^{\infty} \left(-\hat{\Im}_{\bar{\theta}^m}(k)\right) \chi_\lambda(k+L_m)\, e^{i2\pi k\bar{\theta}^m}. \end{aligned}$$

Here, we need to pay attention to $\sum_{k=-\infty}^{\infty} \hat{\Im}_{\bar{\theta}^m}(k)\, \chi_\lambda(k+L_m)\, e^{i2\pi k\bar{\theta}^m}$ for further simplifications. Rename $\hat{\Im}_{\bar{\theta}^m}(k)\, \chi_\lambda(k+L_m)$ as $\hat{P}_{1,\bar{\theta}^m}(k)$. Hence, we have

$$P_1\left(\bar{\theta}^m\right) = \int_0^1 \Im(\tau)\, \check{\chi}_{1,\lambda}\left(\bar{\theta}^m - \tau\right) d\tau,$$

where

$$\check{\chi}_{1,\lambda}\left(\bar{\theta}^m\right) = \sum_{k=-\infty}^{\infty} \chi_\lambda(k+L_m)\, e^{i2\pi k\bar{\theta}^m} = \sum_{L_m\left(-\frac{1}{\lambda}-1\right)\leq k\leq L_m\left(\frac{1}{\lambda}-1\right)} e^{i2\pi k\bar{\theta}^m}.$$

Also, rename $-\hat{\Im}_{\bar{\theta}^m}(k)\, \chi_\lambda(k-L_m)$ as $\hat{P}_{2,\bar{\theta}^m}(k)$. Hence, we get

$$P_2\left(\bar{\theta}^m\right) = \int_0^1 \Im(\tau)\, \check{\chi}_{2,\lambda}\left(\bar{\theta}^m - \tau\right) d\tau,$$

where

$$\check{\chi}_{2,\lambda}\left(\bar{\theta}^m\right) = \sum_{k=-\infty}^{\infty} \chi_\lambda(k-L_m)\, e^{i2\pi k\bar{\theta}^m} = \sum_{L_m\left(-\frac{1}{\lambda}+1\right)\leq k\leq L_m\left(\frac{1}{\lambda}+1\right)} e^{i2\pi k\bar{\theta}^m}.$$

As a consequence, one can find

$$\begin{aligned}
\left|\Im^m_{a,trunc}\left(\bar{\theta}^m\right)\right| &\leq \left|e^{-i2\pi L_m\bar{\theta}^m}\sum_{k=-\infty}^{\infty}\left(-\hat{\Im}_{\bar{\theta}^m}(k)\right)\chi_\lambda(k-L_m)\,e^{i2\pi k\bar{\theta}^m}\right| \\
&\quad+\left|e^{-i2\pi L_m\bar{\theta}^m}\sum_{k=-\infty}^{\infty}\left(-\hat{\Im}_{\bar{\theta}^m}(k)\right)\chi_\lambda(k-L_m)\,e^{i2\pi k\bar{\theta}^m}\right| \\
&=\left|\sum_{k=-\infty}^{\infty}\hat{P}_{2,\bar{\theta}^m}(k)\,e^{i2\pi k\bar{\theta}^m}\right|+\left|\sum_{k=-\infty}^{\infty}\hat{P}_{1,\bar{\theta}^m}(k)\,e^{i2\pi k\bar{\theta}^m}\right| \\
&=\left|\int_0^1\Im(\tau)\,\check{\chi}_{2,\lambda}\left(\bar{\theta}^m-\tau\right)d\tau\right|+\left|\int_0^1\Im(\tau)\,\check{\chi}_{1,\lambda}\left(\bar{\theta}^m-\tau\right)d\tau\right| \\
&\leq\|\Im(t)\|_\infty\left(\int_0^1\left|\check{\chi}_{2,\lambda}\left(\bar{\theta}^m-\tau\right)\right|d\tau+\int_0^1\left|\check{\chi}_{1,\lambda}\left(\bar{\theta}^m-\tau\right)\right|d\tau\right).
\end{aligned}$$

Since, $\check{\chi}_{2,\lambda}\left(\bar{\theta}^m-\tau\right)$ and $\check{\chi}_{1,\lambda}\left(\bar{\theta}^m-\tau\right)$ are periodic, we have

$$\begin{aligned}
\int_0^1\left|\check{\chi}_{i,\lambda}\left(\bar{\theta}^m-\tau\right)\right|d\tau &= -\int_{\bar{\theta}^m}^{\bar{\theta}^m-1}\left|\check{\chi}_{i,\lambda}(\tau)\right|d\tau \\
&= \int_{\bar{\theta}^m-1}^{\bar{\theta}^m}\left|\check{\chi}_{i,\lambda}(\tau)\right|d\tau=\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\check{\chi}_{i,\lambda}(\tau)\right|d\tau,
\end{aligned}$$

for $i=1,2$. So, we get

$$\begin{aligned}
\left|\Im^m_{a,trunc}\left(\bar{\theta}^m\right)\right| &\leq\|\Im(t)\|_\infty\left(\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\check{\chi}_{1,\lambda}(\tau)\right|d\tau+\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\check{\chi}_{2,\lambda}(\tau)\right|d\tau\right) \\
&\leq\|\Im(t)\|_\infty\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\sum_{L_m\left(-\frac{1}{\lambda}-1\right)\leq k\leq L_m\left(\frac{1}{\lambda}-1\right)}e^{i2\pi k\tau}\right|d\tau \\
&\quad+\|\Im(t)\|_\infty\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\sum_{L_m\left(-\frac{1}{\lambda}+1\right)\leq k\leq L_m\left(\frac{1}{\lambda}+1\right)}e^{i2\pi k\tau}\right|d\tau.
\end{aligned}$$

Now, using the identity $e^{i2\pi kt} = \frac{\pi t}{\sin \pi t} \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} e^{i2\pi\omega t} d\omega$, we get

$$\begin{aligned}
\left|\Im^m_{a,trunc}\left(\bar{\theta}^m\right)\right| \leq & \left\|\Im\left(t\right)\right\|_\infty \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \sum_{L_m\left(-\frac{1}{\lambda}-1\right)\leq k\leq L_m\left(\frac{1}{\lambda}-1\right)} \frac{\pi t}{\sin \pi t} \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} e^{i2\pi\omega t} d\omega \right| dt \\
& + \left\|\Im\left(t\right)\right\|_\infty \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \sum_{L_m\left(-\frac{1}{\lambda}+1\right)\leq k\leq L_m\left(\frac{1}{\lambda}+1\right)} \frac{\pi t}{\sin \pi t} \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} e^{i2\pi\omega t} d\omega \right| dt \\
\leq & \frac{\pi}{2} \left\|\Im\left(t\right)\right\|_\infty \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \frac{e^{i2\pi\left(L_m\left(\frac{1}{\lambda}-1\right)+\frac{1}{2}\right)t} - e^{i2\pi\left(L_m\left(-\frac{1}{\lambda}-1\right)-\frac{1}{2}\right)t}}{2\pi i t} \right| dt \\
& + \frac{\pi}{2} \left\|\Im\left(t\right)\right\|_\infty \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \frac{e^{i2\pi\left(L_m\left(\frac{1}{\lambda}+1\right)+\frac{1}{2}\right)t} - e^{i2\pi\left(L_m\left(-\frac{1}{\lambda}+1\right)-\frac{1}{2}\right)t}}{2\pi i t} \right| dt.
\end{aligned}$$

In the last inequality, we have used the fact that $\max_{-\frac{1}{2}\leq t\leq\frac{1}{2}} \left(\frac{\pi t}{\sin \pi t}\right) \leq \frac{\pi}{2}$. From Theorem 5, we know that $\frac{L_m}{L} \leq \frac{9}{8}$. Hence, the above bound would be

$$\left|\Im^m_{a,trunc}\left(\bar{\theta}^m\right)\right| \leq \Upsilon\left(L,\lambda\right) \left\|\Im\left(t\right)\right\|_\infty,$$

for some constant $\Upsilon\left(L,\lambda\right)$. The same thing is true for $\left|\Im^m_{b,trunc}\left(\bar{\theta}^m\right)\right|$:

$$\left|\Im^m_{b,trunc}\left(\bar{\theta}^m\right)\right| \leq \Upsilon\left(L,\lambda\right) \left\|\Im\left(t\right)\right\|_\infty.$$

Hence, we get

$$|\triangle a^m| \leq \Upsilon\left(L,\lambda\right) \left\|\Im\left(t\right)\right\|_\infty + C_1\lambda^{2p-2}L^{-p+2} + C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma,$$

$$|\triangle b^m| \leq \Upsilon\left(L,\lambda\right) \left\|\Im\left(t\right)\right\|_\infty + C_1\lambda^{2p-2}L^{-p+2} + C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma.$$

Now, if $\frac{\left\|\mathcal{F}(\triangle\theta^m)'\right\|_1}{2\pi M_0} = \gamma \leq \frac{1}{4}$, one can find $\eta_0 > 0$ such that $L > \eta_0$ in a way that all the following conditions are satisfied for sufficiently small ϵ_0 (in $\left\|\Im\left(t\right)\right\|_\infty \leq \epsilon_0$),

$$L \geq 2M_0,$$

$$\min\left(\bar{\theta}'\right) \geq \frac{M_0}{2L},$$

$$\Upsilon\left(L,\lambda\right)\left\|\Im\left(t\right)\right\|_{\infty}+\frac{C_{1}\lambda_{0}^{2p-2}}{L^{-p+1}}+\frac{C_{2}\lambda_{0}^{2n}}{4L^{-\max(-n,-2p+1,-n-p+2,-2p)}}\leq\frac{\sqrt{2}}{4}\min f_{1},$$

$$\Gamma_{2}\lambda_{0}^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\leq\frac{1}{2},$$

then we have, for some real positive constant $\Upsilon_{0}\left(L,\lambda\right)$,

$$\left\|\mathcal{F}\left(\theta-\theta^{m+1}\right)'\right\|_{1}\ \leq\ \Upsilon_{0}\left(L,\lambda\right)\left\|\Im\left(t\right)\right\|_{\infty}+\Gamma_{1}\lambda^{2p-2}L^{-p+2}+\tfrac{1}{2}\left\|\mathcal{F}\left(\theta-\theta^{m}\right)'\right\|_{1}.$$

□

5.3. Envelope and Mean Properties

Two of the assumptions used in Theorems 5 and 6 were $\left|\hat{f}_{0,\bar{\theta}}\left(k\right)\right|\leq\frac{C_{0}}{|k|^{p}}$, $\left|\hat{f}_{1,\bar{\theta}}\left(k\right)\right|\leq\frac{C_{0}}{|k|^{p}}$. The condition $\left|\hat{f}_{\bar{\theta}}\left(k\right)\right|\leq\frac{C}{|k|^{p}}$ for $C>0$ and $p\geq4$ sets a connection between f and $\bar{\theta}$. On the other hand, one might be interested in the connection between f and t. We first start with a L_{t}^{2}/L_{t}^{1} Bound. Here, we use the assumption that $p\geq4$ in order to be consistent with the convergence theorems of this section. . However, this assumption can be relaxed to $p\geq1$.

Theorem 7. *If* $\left|\hat{f}_{\bar{\theta}}\left(k\right)\right|\leq\frac{C}{|k|^{p}}$ *for* $C>0$ *and* $p\geq4$*, then*

$$\frac{\|f\|_{L_{t}^{2}}}{\|f\|_{L_{t}^{1}}}\ \leq\ \sqrt{3+\tfrac{2}{2p-1}}\frac{\left(\max\left(\bar{\theta}'\right)\right)}{\sqrt{\min\left(\bar{\theta}'\right)}}. \tag{5.3.1}$$

Proof. As $\left|\hat{f}_{\bar{\theta}}\left(k\right)\right|\leq\frac{C}{|k|^{p}}$ for $C>0$ and $p\geq4$, the function f is at least in $\mathcal{C}_{\bar{\theta}}^{2}$. Consequently, it is also in $\mathcal{C}_{t}^{2}$. Hence, f belongs to L_{t}^{q} and $L_{\bar{\theta}}^{q}$ for $q\geq1$. So we have

$$\begin{aligned}\left|\hat{f}_{\bar{\theta}}\left(k\right)\right|&=\left|\int_{0}^{1}f\left(\bar{\theta}\right)e^{-i2\pi k\bar{\theta}}d\bar{\theta}\right|\\&\leq\max\left(\bar{\theta}'\right)\left(\int_{0}^{1}\left|f\left(t\right)\right|dt\right).\end{aligned}$$

In other words, we have

$$\left\|\hat{f}_{\bar{\theta}}\right\|_{\infty}\ \leq\ \max\left(\bar{\theta}'\right)\|f\|_{L_{t}^{1}}.$$

Since $\left\|\hat{f}_{\bar{\theta}}\right\|_{\infty} = \max\left\{\left|\hat{f}_{\bar{\theta}}(0)\right|, C\right\}$, for the 2-norm of $\hat{f}_{\bar{\theta}}$, we get

$$\begin{aligned}\left\|\hat{f}_{\bar{\theta}}\right\|_2^2 =& \sum_k \left|\hat{f}_{\bar{\theta}}(k)\right|^2 = \left|\hat{f}_{\bar{\theta}}(0)\right|^2 + \sum_{|k|\geq 1}\left|\hat{f}_{\bar{\theta}}(k)\right|^2 \\ &\leq \left|\hat{f}_{\bar{\theta}}(0)\right|^2 + 2C^2\sum_{k\geq 1}\frac{1}{k^{2p}} \\ &\leq \left\|\hat{f}_{\bar{\theta}}\right\|_{\infty}^2\left(3+\frac{2}{2p-1}\right).\end{aligned}$$

Here we approximated the sum by an integral. By Parseval's identity, we know $\left\|\hat{f}_{\bar{\theta}}\right\|_2 = \|f\|_{L^2_{\bar{\theta}}}$, hence we find out that

$$\|f\|_{L^2_{\bar{\theta}}}^2 \;\leq\; \left\|\hat{f}_{\bar{\theta}}\right\|_{\infty}^2\left(3+\tfrac{2}{2p-1}\right).$$

Now, we need to find the lower-bound of $\|f\|_{L^2_{\bar{\theta}}}$ in terms of $\|f\|_{L^2_t}$. We have

$$\begin{aligned}\|f\|_{L^2_{\bar{\theta}}}^2 &= \int_0^1 \left|f\left(\bar{\theta}\right)\right|^2 d\bar{\theta} \\ &= \int_0^1 |f(t)|^2 \frac{d\bar{\theta}}{dt}dt \\ &\geq \min\left(\bar{\theta}'\right)\|f\|_{L^2_t}^2.\end{aligned}$$

Now, combining all together, we have

$$\min\left(\bar{\theta}'\right)\|f\|_{L^2_t}^2 \leq \|f\|_{L^2_{\bar{\theta}}}^2 \;\leq\; \left\|\hat{f}_{\bar{\theta}}\right\|_{\infty}^2\left(3+\tfrac{2}{2p-1}\right) \leq \left(\max\left(\bar{\theta}'\right)\right)^2\|f\|_{L^1_t}^2\left(3+\tfrac{2}{2p-1}\right),$$

which is

$$\frac{\|f\|_{L^2_t}}{\|f\|_{L^1_t}} \;\leq\; \left(3+\tfrac{2}{2p-1}\right)^{\frac{1}{2}}\frac{\left(\max\left(\bar{\theta}'\right)\right)}{\sqrt{\min\left(\bar{\theta}'\right)}}\;.$$

□

This can be seen as a condition that we need for the envelope or trend (mean) in the convergence theorem. It is now time for a more general bound.

Theorem 8. *Assume* $\left|\hat{f}_{\bar{\theta}}(k)\right| \leq \frac{C}{|k|^p}$ *for* $C > 0$ *and* $p \geq 4$. *Also take* $1 \leq r \leq \infty$ *and* $1 \leq q < \infty$ *then*

$$\frac{\|f\|_{L_t^2}}{\|f\|_{L_t^r}} \leq \left(3 + \frac{2}{qp-1}\right)^{\frac{1}{2q}} \left(3 + \frac{2}{\frac{qp}{q-1}-1}\right)^{\frac{q-1}{2q}} \frac{\left\|\bar{\theta}'\right\|_{L_t^{\frac{r}{r-1}}}}{\sqrt{\min\left(\bar{\theta}'\right)}} . \tag{5.3.2}$$

Proof. Again, first of all, we need a bound on $\left|\hat{f}_{\bar{\theta}}(k)\right|$ for $1 \leq r \leq \infty$:

$$\begin{aligned}\left|\hat{f}_{\bar{\theta}}(k)\right| &= \left|\int_0^1 f\left(\bar{\theta}\right) e^{-i2\pi k\bar{\theta}} d\bar{\theta}\right| = \left|\int_0^1 f(t)\, e^{-i2\pi k\bar{\theta}(t)} \bar{\theta}' dt\right| \\ &\leq \|f\|_{L_t^r} \left\|\bar{\theta}'\right\|_{L_t^{\frac{r}{r-1}}} .\end{aligned}$$

Hence, we get

$$\left\|\hat{f}_{\bar{\theta}}\right\|_\infty \leq \|f\|_{\mathcal{L}_t^r} \left\|\bar{\theta}'\right\|_{\mathcal{L}_t^{\frac{r}{r-1}}} .$$

Since $\left\|\hat{f}_{\bar{\theta}}\right\|_\infty = \max\left\{\left|\hat{f}_{\bar{\theta}}(0)\right|, C\right\}$, for the q-norm, $1 \leq q < \infty$, of $\hat{f}_{\bar{\theta}}$ we have

$$\begin{aligned}\left\|\hat{f}_{\bar{\theta}}\right\|_q^q &= \sum_k \left|\hat{f}_{\bar{\theta}}(k)\right|^q = \left|\hat{f}_{\bar{\theta}}(0)\right|^q + \sum_{|k|\geq 1} \left|\hat{f}_{\bar{\theta}}(k)\right|^q \\ &\leq \left|\hat{f}_{\bar{\theta}}(0)\right|^q + 2C^q \sum_{k\geq 1} \frac{1}{k^{qp}} \\ &\leq \left\|\hat{f}_{\bar{\theta}}\right\|_\infty^q \left(3 + \frac{2}{qp-1}\right).\end{aligned}$$

In the last inequality, the sum is bounded by an integral. By Parseval's identity, we know $\left\|\hat{f}_{\bar{\theta}}\right\|_2 = \|f\|_{L_{\bar{\theta}}^2}$. Now, using the Holder inequality, we have

$$\begin{aligned}\left\|\hat{f}_{\bar{\theta}}\right\|_2^2 = \|f\|_{L_{\bar{\theta}}^2}^2 &= \sum_k \left|\hat{f}_{\bar{\theta}}(k)\right|^2 = \sum_k \hat{f}_{\bar{\theta}}(k)\, \overline{\hat{f}_{\bar{\theta}}(k)} \\ &\leq \left\|\hat{f}_{\bar{\theta}}\right\|_q \left(\sum_k \left|\hat{f}_{\bar{\theta}}(k)\right|^{\frac{q}{q-1}}\right)^{\frac{q-1}{q}} \\ &\leq \left\|\hat{f}_{\bar{\theta}}\right\|_q \left\|\hat{f}_{\bar{\theta}}\right\|_\infty \left(3 + \frac{2}{\frac{qp}{q-1}-1}\right)^{\frac{q-1}{q}} .\end{aligned}$$

Now, using $\left\|\hat{f}_{\bar{\theta}}\right\|_q^q \leq \left\|\hat{f}_{\bar{\theta}}\right\|_\infty^q \left(3+\frac{2}{qp-1}\right)$ and $\left\|\hat{f}_{\bar{\theta}}\right\|_\infty \leq \|f(t)\|_{L_t^r} \left\|\bar{\theta}'\right\|_{L_t^{\frac{r}{r-1}}}$, we get

$$\|f\|_{L_{\bar{\theta}}^2}^2 \leq \|f\|_{L_t^r}^2 \left\|\bar{\theta}'\right\|_{L_t^{\frac{r}{r-1}}}^2 \left(3+\frac{2}{qp-1}\right)^{\frac{1}{q}} \left(3+\frac{2}{\frac{qp}{q-1}-1}\right)^{\frac{q-1}{q}} .$$

Finally, since $\min\left(\bar{\theta}'\right)\|f\|_{L_t^2}^2 \leq \|f\|_{L_{\bar{\theta}}^2}^2$, we find out that

$$\frac{\|f\|_{L_t^2}}{\|f\|_{L_t^r}} \leq \left(3+\frac{2}{qp-1}\right)^{\frac{1}{2q}} \left(3+\frac{2}{\frac{qp}{q-1}-1}\right)^{\frac{q-1}{2q}} \frac{\left\|\bar{\theta}'\right\|_{L_t^{\frac{r}{r-1}}}}{\sqrt{\min\left(\bar{\theta}'\right)}}.$$

□

5.4. Uniqueness Issues

The theorems, proven in this chapter explain only that we have a convergent algorithm, not that the algorithm's extraction is unique. Uniqueness is a difficult theoretical problem. In this section, we explain insights about how we can handle this topic in implementation and numerical analysis.

Here, we reduce uniqueness into sparsity. In fact, any signal can have many representations. In our STFR methodology, we prefer a representation in which there is the least number of IMFs extracted from the signal. In other words, a sparse representation is the unique representation in STFR numerical terminology. If two different extractions would result in the same number of IMFs, both of them are acceptable. However, we pick the one in which the components have the smoothest envelopes compared to the IFs.

Most importantly the result of a numerical implementation of STFR must be sparse. In other words, we pick the set of IMFs as the constituent blocks of a signal if we cannot find another set that contains fewer IMFs. A preferred extraction is the one that has fewer IMF components. In fact, as we observed, in the case of intrawave signals, widening the filter width would help to find the smallest number of IMFs in an extraction.

One should keep in mind that if there is any mode mixture present (the IFs are not separate in time-frequency domain), it is highly possible to have a non-sparse extraction. In order to tackle this problem, one should find an estimate of the number of mode mixtures and then try to extract the mode mixed IMFs simultaneously, which would impose sparsity. Quantitatively, having two extractions with the same number of IMFs, the smoothness measure (SM) can be one of the following.

$$SM_{TV} = \frac{TV\left(\tilde{a}\left(t\right)\right)}{TV\left(\tilde{\dot{\theta}}\left(t\right)\right)},$$

$$SM_{max-local} = \max\left(\frac{\tilde{\dot{a}}\left(t\right)}{\tilde{\dot{\theta}}\left(t\right)}\right).$$

In these definitions, $\tilde{a}\left(t\right)$ is the extracted envelope, $\tilde{\dot{\theta}}\left(t\right)$ is the extracted IF, and TV is the total variation. In case of many IMFs, one can sum up the SM of all extracted IMFs to judge the quality of extraction.

5.5. Appendices

In this section we provide the omitted parts of the proof presented in this chapter.

5.5.1. Appendix A (Approximating $\left|\hat{f}_{0,\bar{\theta}^m}\left(\omega\right)\right|$). In order to find a bound on $\left|\hat{f}_{0,\bar{\theta}^m}\left(\omega\right)\right|$, in (5.1.4), during the following calculations, we will bound some of the summations by integrals. In detail, we have

$$\begin{aligned}
\left|\hat{f}_{0,\bar{\theta}^m}\left(\omega\right)\right| &= \left|\int_0^1 f_0\left(\bar{\theta}^m\right) e^{-i2\pi\omega\bar{\theta}^m} d\bar{\theta}^m\right| = \left|\int_0^1 f_0\left(\bar{\theta}^m\left(\bar{\theta}\right)\right) e^{-i2\pi\omega\bar{\theta}^m} d\bar{\theta}^m\right| \\
&= \left|\sum_{k\neq 0} \hat{f}_{0,\bar{\theta}}\left(k\right) \int_0^1 e^{i2\pi\bar{\theta}^m(\alpha k-\omega)} e^{ik\frac{\triangle\theta^m}{L}} d\bar{\theta}^m\right| \\
&\leq \sum_{|\alpha k|>\frac{|\omega|}{\lambda}} \left|\hat{f}_{0,\bar{\theta}}\left(k\right)\right| + \sum_{0<|\alpha k|\leq\frac{|\omega|}{\lambda}} \left|\hat{f}_{0,\bar{\theta}}\left(k\right)\right| \left|\int_0^1 e^{i2\pi\bar{\theta}^m(\alpha k-\omega)} e^{ik\frac{\triangle\theta^m}{L}} d\bar{\theta}^m\right|.
\end{aligned}$$

Since we have $\left|\hat{f}_{0,\bar{\theta}}(k)\right| \leq |k|^{-p}$ for $k \neq 0$, and using Lemma 2 and 3 in Section 3, we get

$$\begin{aligned}
\left|\hat{f}_{0,\bar{\theta}^m}(\omega)\right| \leq & C_0 \sum_{|k|>\frac{|\omega|}{\alpha\lambda}} |k|^{-p} + C_0 \sum_{0<|k|\leq\frac{|\omega|}{\alpha\lambda}} |k|^{-p} \left| \int_0^1 e^{i2\pi\bar{\theta}^m(\alpha k-\omega)} e^{ik\frac{\triangle\theta^m}{L}} d\bar{\theta}^m \right| \\
\leq & 2C_0 (\alpha\lambda)^{p-1} |\omega|^{-p+1} \\
& + 2C_0 P_m^n \left(\frac{M_0\lambda}{2\pi}\right)^n |\varepsilon\omega|^{-n} \sum_{0<|k|\leq\frac{|\omega|}{\alpha\lambda}} \sum_{j=1}^{n} |k|^{j-p} \left(\frac{\left\|\mathcal{F}\left(\triangle\theta^m\right)'\right\|_1}{2\pi M_0 L}\right)^j \\
\leq & 2C_0 \left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} \\
& + 2C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \left(\sum_{0<|k|\leq\frac{|\omega|}{\alpha\lambda}} \left(\sum_{j=1}^{p-2} |k|^{j-p} \left(\frac{\gamma}{L}\right)^j \right) \right) \\
& + 2C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \left(\sum_{0<|k|\leq\frac{|\omega|}{\alpha\lambda}} \left(\sum_{j=p}^{n} |k|^{j-p} \left(\frac{\gamma}{L}\right)^j + |k|^{-1} \left(\frac{\gamma}{L}\right)^{p-1} \right) \right) \\
\leq & 2C_0 \left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} \\
& + 2C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \left(\sum_{j=1}^{p-2} \left(\frac{\gamma}{L}\right)^j \frac{\pi^2}{3} \right) \\
& + 2C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \left(2\sum_{j=p}^{n} \left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1} \left(\frac{\gamma}{L}\right)^j + 2\left(\frac{\gamma}{L}\right)^{p-1} \left(1+\frac{|\omega|}{\alpha\lambda}\right) \right).
\end{aligned}$$

In the inequalities above, we have also used the fact that if $0 < k \leq \frac{\omega}{\alpha\lambda}$ then $\frac{1}{\omega-\alpha k} < \frac{\lambda}{\lambda-1}\frac{1}{\omega} < \frac{\lambda}{\varepsilon}\frac{1}{\omega}$ provided $0 < \varepsilon \ll 1$ and $\lambda \geq 1+\varepsilon$. Also we called $\frac{\left\|\mathcal{F}(\triangle\theta^m)'\right\|_1}{2\pi M_0} = \gamma$. We recall that in these calculations $n > p$. This finalized the derivation of the bound on (5.1.4).

5.5.2. Appendix B (Approximating $\left|\hat{a}^m_{\bar{\theta}^m}(\omega)\right|$). This approximation (see (5.1.5)) is essentially the same as the one we saw in Appendix A, except that we also use

Jensen's inequality. In the main theorem of this section, we extensively use this lemma. So, we present the following lemma to introduce the inequality.

Lemma 4. *[Jensen's Inequality] If φ is a convex function, and $a_i \in \mathbb{R}^+$ for $i \in \mathbb{N}$, $i > 1$, then*

$$\varphi\left(\frac{\sum a_i x_i}{\sum a_i}\right) \leq \frac{\sum a_i \varphi(x_i)}{\sum a_i}. \tag{5.5.1}$$

A proof of this lemma can be found in [8]. A consequence of this lemma is that, for $j \in \mathbb{N} + \{0\}$, we have

$$2^{-j}\left|1 + \frac{k}{L}\right|^j = \left|\frac{1}{2}1 + \frac{1}{2}\frac{k}{L}\right|^j \leq \frac{1}{2}\left(|1|^j + \left|\frac{k}{L}\right|^j\right). \tag{5.5.2}$$

This bound is used several times in the next approximations. As mentioned above, we also bound summations by integrals. So, for $\left|\hat{a}^m_{\bar{\theta}^m}(\omega)\right|$ we have

$$\begin{aligned}
|\hat{a}^m_{\bar{\theta}^m}(\omega)| =& \left|\int_0^1 a^m\left(\bar{\theta}^m\right) e^{-i2\pi\omega\bar{\theta}^m} d\bar{\theta}^m\right| = \left|\int_0^1 a^m\left(\bar{\theta}^m\left(\bar{\theta}\right)\right) e^{-i2\pi\omega\bar{\theta}^m} d\bar{\theta}^m\right| \\
=& \frac{1}{2}\left|\int_0^1 \left(e^{i\Delta\theta^m} + e^{-i\Delta\theta^m}\right) \sum_{k=-\infty}^{\infty} \hat{f}_{1,\bar{\theta}}(k)\, e^{i2\pi\bar{\theta}^m(\alpha k-\omega)} e^{ik\frac{\triangle\theta^m}{L}} d\bar{\theta}^m\right| \\
\leq& \frac{1}{2}\left|\int_0^1 \sum_{k=-\infty}^{\infty} \hat{f}_{1,\bar{\theta}}(k)\, e^{i2\pi\bar{\theta}^m(\alpha k-\omega)} e^{i\triangle\theta^m\left(\frac{k}{L}+1\right)} d\bar{\theta}^m\right| \\
&+ \frac{1}{2}\left|\int_0^1 \sum_{k=-\infty}^{\infty} \hat{f}_{1,\bar{\theta}}(k)\, e^{i2\pi\bar{\theta}^m(\alpha k-\omega)} e^{i\triangle\theta^m\left(\frac{k}{L}-1\right)} d\bar{\theta}^m\right|.
\end{aligned}$$

Now, we break up the the summation $\sum_{k=-\infty}^{\infty}$ into summations over $k=0$, $|\alpha k|>\frac{|\omega|}{\lambda}$, and $0<|\alpha k|\leq\frac{|\omega|}{\lambda}$ and also use the first and second lemma in Section 3 to get

$$\begin{aligned}
|\hat{a}^m_{\bar{\theta}^m}(\omega)| \leq & 2C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1}+\left|\hat{f}_{1,\bar{\theta}}(0)\right|P_m^n\left(\frac{M_0}{2\pi|\omega|}\right)^n\sum_{j=1}^n\gamma^j \\
& +C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\sum_{0<|k|\leq\frac{|\omega|}{\alpha\lambda}}|k|^{-p}\sum_{j=1}^n\left(\left|1+\frac{k}{L}\right|^j+\left|1-\frac{k}{L}\right|^j\right)\gamma^j \\
\leq & 2C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1}+\left|\hat{f}_{1,\bar{\theta}}(0)\right|P_m^n\left(\frac{M_0}{2\pi|\omega|}\right)^n\sum_{j=1}^n\gamma^j \\
& +C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\sum_{0<|k|\leq\frac{|\omega|}{\alpha\lambda}}|k|^{-p}\sum_{j=1}^n2^j\left(1+\left|\frac{k}{L}\right|^j\right)\gamma^j \\
\leq & 2C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1}+\left|\hat{f}_{1,\bar{\theta}}(0)\right|P_m^n\left(\frac{M_0}{2\pi|\omega|}\right)^n\sum_{j=1}^n\gamma^j \\
& +C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\frac{2p}{p-1}\sum_{j=1}^n2^j\gamma^j \\
& +C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\left(\sum_{j=1}^{p-2}2^j\left(\frac{\gamma}{L}\right)^j\frac{\pi^2}{3}\right) \\
& +C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\left(\sum_{j=p}^n2^{j+1}\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^j+2^p\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right)\right).
\end{aligned}$$

The latter finalized the derivation of the bound on $\left|\hat{a}^m_{\bar{\theta}^m}(\omega)\right|$; see (5.1.5).

5.5.3. Appendix C (Bounds on $|\triangle a^m|$, $|\triangle b^m|$). In order to find the bound on $|\triangle a^m|$ (see (5.1.6)), we first check the term involving $\left|\hat{f}_{0,\bar{\theta}^m}(\omega)\right|$. Taking

$$2\sum_{\left(1-\frac{1}{\lambda}\right)L_m\leq\omega\leq\left(1+\frac{1}{\lambda}\right)L_m}\left|\hat{f}_{0,\bar{\theta}^m}(\omega)\right|=\beth,$$

we have

$$\begin{aligned}
\beth \leq & 4 \sum_{\left(1-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(1+\frac{1}{\lambda}\right)L_m} C_0 \left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} \\
& + 4 \sum_{\left(1-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(1+\frac{1}{\lambda}\right)L_m} C_0 P_m^n \left(\frac{M_0\lambda}{2\pi\left|\varepsilon\omega\right|}\right)^n \sum_{j=1}^{p-2} \left(\frac{\gamma}{L}\right)^j \frac{\pi^2}{3} \\
& + 8 \sum_{\left(1-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(1+\frac{1}{\lambda}\right)L_m} C_0 P_m^n \left(\frac{M_0\lambda}{2\pi\left|\varepsilon\omega\right|}\right)^n \sum_{j=p}^{n} \left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1} \left(\frac{\gamma}{L}\right)^j \\
& + 8 \sum_{\left(1-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(1+\frac{1}{\lambda}\right)L_m} C_0 P_m^n \left(\frac{M_0\lambda}{2\pi\left|\varepsilon\omega\right|}\right)^n \left(\frac{\gamma}{L}\right)^{p-1} \left(1+\frac{|\omega|}{\alpha\lambda}\right).
\end{aligned}$$

Here, we can use a trick to bound the summations with integrals. When $\left(1-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(1+\frac{1}{\lambda}\right)L_m$, we know that ω will at least start from 1. The latter is due to the fact that $\lambda > 1$, and hence, $\left(1-\frac{1}{\lambda}\right)L_m > 0$. So, we can always find a fixed $\varepsilon > 0$ such that $0 < \left|1-\frac{1}{1-\varepsilon}\right| L_m < \min\left(\left(1-\frac{1}{\lambda}\right)L_m, 1\right)$. We can take this ε to be the same as the one we used before. So, the whole inequality, using the fact that $L_m = \alpha L$, becomes

$$\begin{aligned}
\beth \leq & \frac{4C_0\alpha}{p-2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-p+2}\lambda^{p-1}L^{-p+2} \\
& + \frac{4\pi^2 C_0 P_m^n \left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^n \alpha^{-n+1}}{3\left(n-1\right)} L^{-n+1}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+1}\sum_{j=1}^{p-2}\left(\frac{\gamma}{L}\right)^j \\
& + 8C_0\left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^n P_m^n \alpha^{-n+2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+2} L^{-p+2}\frac{1}{p+2}\sum_{j=p}^{n}\left(\frac{\gamma}{L}\right)^j \\
& + 8C_0 P_m^n\left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^n\left(\frac{\gamma}{L}\right)^{p-1}\left(\frac{\left(\alpha L\left(\frac{\varepsilon}{1-\varepsilon}\right)\right)^{-n+1}}{n-1} + \frac{\left(\alpha L\left(\frac{\varepsilon}{1-\varepsilon}\right)\right)^{-n+2}}{\alpha\lambda\left(n-2\right)}\right).
\end{aligned}$$

Here, we use the assumption that $\gamma \leq \frac{1}{4}$. This assumption would remain intact throughout the steps for large enough L. In fact, the condition $\gamma \leq \frac{1}{4}$ would remain intact for all iterations. In other words, when there is a contraction on $\left\|\mathcal{F}\left(\theta-\theta^{m+1}\right)'\right\|_1$, this term would remain bounded. Hence, if $\frac{\left\|\mathcal{F}\left(\triangle\theta^0\right)'\right\|_1}{2\pi M_0} \leq \frac{1}{4}$ for the first iteration, it will remain bounded by $\frac{1}{4}$ for all iterations. Hence, using this,

we have $\sum_{j=1}^{p-2}\left(\frac{\gamma}{L}\right)^{j} \leq \left(\frac{\gamma}{L}\right)\sum_{j=1}^{\infty}\left(\frac{\gamma}{L}\right)^{j-1} \leq \frac{4\gamma}{3L}$. For the second sum, we use the same trick, namely $\sum_{j=p}^{n}\left(\frac{\gamma}{L}\right)^{j} \leq \left(\frac{\gamma}{L}\right)^{p}\sum_{j=1}^{\infty}\left(\frac{\gamma}{L}\right)^{j-1} \leq (4L)^{-p}\,\frac{4\gamma}{3L}$. So, we get

$$
\begin{aligned}
\beth \leq & \frac{4C_0\alpha}{p-2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-p+2}\lambda^{p-1}L^{-p+2} \\
& + \frac{4\pi^2 C_0 P_m^n\left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^n\alpha^{-n+1}}{3\,(n-1)}L^{-n+1}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+1}\frac{4\gamma}{3L} \\
& + 8C_0\left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^n P_m^n\alpha^{-n+2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+2}L^{-p+2}\frac{1}{p+2}\,(4L)^{-p}\,\frac{4\gamma}{3L} \\
& + 8C_0 P_m^n\left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^n\left(\frac{\gamma}{L}\right)^{p-1}\left(\frac{\left(\alpha L\left(\frac{\varepsilon}{1-\varepsilon}\right)\right)^{-n+1}}{n-1}+\frac{\left(\alpha L\left(\frac{\varepsilon}{1-\varepsilon}\right)\right)^{-n+2}}{\alpha\lambda\,(n-2)}\right).
\end{aligned}
$$

Simplifying further, we get

$$
\begin{aligned}
\beth \leq & \frac{4C_0\alpha}{p-2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-p+2}\lambda^{p-1}L^{-p+2} \\
& + \frac{16}{9}\pi^2 C_0 P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n+1}}{n-1}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+1}L^{-n}\lambda^n\gamma \\
& + \frac{8}{3}C_0 P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n+2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+2}\frac{4^{-p+1}}{p+2}L^{-2p+1}\lambda^n\gamma \\
& + 8C_0 P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n+1}\frac{\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+1}}{n-1}L^{-n-p+1}\lambda^n\gamma \\
& + 8C_0 P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n+1}\frac{\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+2}}{(n-2)}L^{-n-p+2}\lambda^{n-1}\gamma.
\end{aligned}
$$

Now, we check the term involving $\left|\hat{a}_{\bar{\theta}^m}^m(k)\right| + \left|\hat{b}_{\bar{\theta}^m}^m(k)\right|$. Taking

$$
\sum_{\left(2-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(2+\frac{1}{\lambda}\right)L_m}\left(\left|\hat{a}_{\bar{\theta}^m}^m(\omega)\right| + \left|\hat{b}_{\bar{\theta}^m}^m(\omega)\right|\right) = \daleth
$$

we have

$$
\begin{aligned}
\urcorner \leq & 4 \sum_{\left(2-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(2+\frac{1}{\lambda}\right)L_m} C_0 \left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} \\
& + 2 \sum_{\left(2-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(2+\frac{1}{\lambda}\right)L_m} \left|\hat{f}_{1,\bar{\theta}}(0)\right| P_m^n \left(\frac{M_0}{2\pi|\omega|}\right)^n \sum_{j=1}^{n} \gamma^j \\
& + 2 \sum_{\left(2-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(2+\frac{1}{\lambda}\right)L_m} C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \frac{2p}{p-1} \sum_{j=1}^{n} 2^j \gamma^j \\
& + 2 \sum_{\left(2-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(2+\frac{1}{\lambda}\right)L_m} C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \sum_{j=1}^{p-2} 2^j \left(\frac{\gamma}{L}\right)^j \frac{\pi^2}{3} \\
& + 2 \sum_{\left(2-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(2+\frac{1}{\lambda}\right)L_m} C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \sum_{j=p}^{n} 2^{j+1} \left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1} \left(\frac{\gamma}{L}\right)^j \\
& + 2 \sum_{\left(2-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(2+\frac{1}{\lambda}\right)L_m} C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n 2^p \left(\frac{\gamma}{L}\right)^{p-1} \left(1+\frac{|\omega|}{\alpha\lambda}\right).
\end{aligned}
$$

In this case, due to the presence of the term $\left(2-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(2+\frac{1}{\lambda}\right)L_m$, the substitution of the sum with an integral is much easier. Since $\lambda > 1$, we have $1 < 2-\frac{1}{\lambda} < 2$. As a result, $\alpha L = L_m < \left(2-\frac{1}{\lambda}\right)L_m \leq \omega$. Hence, for any $\zeta > 1$ we have

$$
\begin{aligned}
\sum_{\left(2-\frac{1}{\lambda}\right)L_m \leq \omega \leq \left(2+\frac{1}{\lambda}\right)L_m} \frac{1}{\omega^\zeta} &\leq \sum_{\alpha L < \omega} \frac{1}{\omega^\zeta} \\
&= \frac{1}{(\alpha L)^\zeta} \sum_{1 < \frac{\omega}{\alpha L}} \frac{1}{\left(\frac{\omega}{\alpha L}\right)^\zeta} = \frac{1}{(\alpha L)^\zeta} \sum_{1<k} \frac{1}{k^\zeta} \\
&\leq \frac{1}{(\alpha L)^\zeta} \int_1^\infty \frac{1}{k^\zeta} = \frac{(\alpha L)^{-\zeta}}{\zeta - 1}.
\end{aligned}
$$

Consequently, we get

$$\begin{aligned}
\urcorner \leq & \frac{4C_0}{p-2}\lambda^{p-1}L^{-p+1} \\
& + \frac{8}{3}\left|\hat{f}_{1,\bar{\theta}}(0)\right| P_m^n \left(\frac{M_0}{2\pi}\right)^n \frac{\alpha^{-n}}{n-1}L^{-n}\gamma \\
& + \frac{16p}{p-1}C_0 P_m^n \left(\frac{M_0}{2\pi\varepsilon}\right)^n \frac{\alpha^{-n}}{n-1}L^{-n}\lambda^n\gamma \\
& + \frac{8\pi^2}{3}C_0 P_m^n \left(\frac{M_0}{2\pi\varepsilon}\right)^n \frac{\alpha^{-n}}{n-1}L^{-n-1}\lambda^n\gamma \\
& + 2^{-p+4}C_0 P_m^n \left(\frac{M_0}{2\pi\varepsilon}\right)^n \alpha^{-n}L^{-2p}\lambda^n\gamma \\
& + 2^{-p+3}C_0 P_m^n \left(\frac{M_0}{2\pi\varepsilon}\right)^n \frac{\alpha^{-n}L^{-n-p+1}\lambda^n\gamma}{n-1} \\
& + 2^{-p+3}C_0 P_m^n \left(\frac{M_0}{2\pi\varepsilon}\right)^n \frac{\alpha^{-n}L^{-n-p+2}\lambda^n\gamma}{(n-2)\lambda}.
\end{aligned}$$

Finally, we find the bound on the term that involves $\left|\hat{a}^m_{\bar{\theta}^m}(k)\right|$. This term appears in the bound of $|\triangle a^m|$. We have

$$\begin{aligned}
\sum_{|\omega|>\frac{L_m}{\lambda}} |\hat{a}^m_{\bar{\theta}^m}(\omega)| \leq & 2\sum_{|\omega|>\frac{L_m}{\lambda}} C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} \\
& + \sum_{|\omega|>\frac{L_m}{\lambda}} \left|\hat{f}_{1,\bar{\theta}}(0)\right| P_m^n \left(\frac{M_0}{2\pi|\omega|}\right)^n \sum_{j=1}^{n}\gamma^j \\
& + \sum_{|\omega|>\frac{L_m}{\lambda}} C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \frac{2p}{p-1}\sum_{j=1}^{n}2^j\gamma^j \\
& + \sum_{|\omega|>\frac{L_m}{\lambda}} C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \sum_{j=1}^{p-2}2^j\left(\frac{\gamma}{L}\right)^j \frac{\pi^2}{3} \\
& + \sum_{|\omega|>\frac{L_m}{\lambda}} C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \sum_{j=p}^{n}2^{j+1}\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^j \\
& + \sum_{|\omega|>\frac{L_m}{\lambda}} C_0 P_m^n \left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n 2^p\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right).
\end{aligned}$$

Again we have for any $\zeta > 1$

$$\begin{aligned}\sum_{|\omega|>\frac{L_m}{\lambda}} \frac{1}{|\omega|^{\zeta}} &= 2\left(\frac{\lambda}{\alpha L}\right)^{\zeta} \sum_{\frac{\lambda\omega}{\alpha L}>1} \frac{1}{\left(\frac{\lambda\omega}{\alpha L}\right)^{\zeta}} \\ &= 2\left(\frac{\lambda}{\alpha L}\right)^{\zeta} \sum_{1<k} \frac{1}{k^{\zeta}} \leq 2\left(\frac{\lambda}{\alpha L}\right)^{\zeta} \int_1^{\infty} \frac{1}{k^{\zeta}} = 2\left(\frac{\lambda}{\alpha L}\right)^{\zeta} \frac{1}{\zeta - 1}.\end{aligned}$$

Hence, we can again bound the summations with integrals. So, we get

$$\begin{aligned}\sum_{|\omega|>\frac{L_m}{\lambda}} |\hat{a}^m_{\bar{\theta}^m}(\omega)| \leq & \frac{4C_0}{p-2}\lambda^{2p-2}L^{-p+1} \\ & + \frac{8}{3}\left|\hat{f}_{1,\bar{\theta}}(0)\right| P_m^n \left(\frac{M_0}{2\pi}\right)^n \frac{\alpha^{-n}}{n-1} L^{-n}\lambda^n\gamma \\ & + \frac{16p}{p-1} C_0 P_m^n \left(\frac{M_0}{2\pi\varepsilon}\right)^n \frac{\alpha^{-n}}{n-1} L^{-n}\lambda^{2n}\gamma \\ & + \frac{8\pi^2}{3} C_0 P_m^n \left(\frac{M_0}{2\pi\varepsilon}\right)^n \frac{\alpha^{-n}}{n-1} L^{-n-1}\lambda^{2n}\gamma \\ & + 2^{-p+5} C_0 P_m^n \left(\frac{M_0}{2\pi\varepsilon}\right)^n \alpha^{-n} L^{-2p-1}\lambda^{2n-2}\gamma \\ & + 2^{-p+5} C_0 P_m^n \left(\frac{M_0}{2\pi\varepsilon}\right)^n \frac{\alpha^{-n}}{n-1} L^{-n-p+1}\lambda^{2n}\gamma \\ & + 2^{-p+5} C_0 P_m^n \left(\frac{M_0}{2\pi\varepsilon}\right)^n \frac{\alpha^{-n}}{(n-2)\lambda} L^{-n-p+2}\lambda^{2n-2}\gamma.\end{aligned}$$

Finally, combining all, we have

$$
\begin{aligned}
|\triangle a^m| \leq & \frac{4C_0\alpha}{p-2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-p+2}\lambda^{p-1}L^{-p+2}+\frac{4C_0}{p-2}\lambda^{p-1}L^{-p+1} \\
& +\frac{4C_0}{p-2}\lambda^{2p-2}L^{-p+1} \\
& +\frac{16}{9}\pi^2 C_0 P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n+1}}{n-1}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+1}L^{-n}\lambda^n\gamma \\
& +\frac{8}{3}\left|\hat{f}_{1,\bar{\theta}}(0)\right|P_m^n\left(\frac{M_0}{2\pi}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n}\gamma \\
& +\frac{8}{3}\left|\hat{f}_{1,\bar{\theta}}(0)\right|P_m^n\left(\frac{M_0}{2\pi}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n}\lambda^n\gamma \\
& +\frac{8}{3}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n+2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+2}\frac{4^{-p+1}}{p+2}L^{-2p+1}\lambda^n\gamma \\
& +\frac{16p}{p-1}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n}\lambda^n\gamma \\
& +8C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n+1}\frac{\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+1}}{n-1}L^{-n-p+1}\lambda^n\gamma \\
& +\frac{8\pi^2}{3}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n-1}\lambda^n\gamma \\
& +2^{-p+4}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n}L^{-2p}\lambda^n\gamma \\
& +8C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n+1}\frac{\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+2}}{(n-2)}L^{-n-p+2}\lambda^{n-1}\gamma \\
& +2^{-p+3}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n-p+1}\lambda^n\gamma \\
& +2^{-p+3}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{(n-2)\lambda}L^{-n-p+2}\lambda^n\gamma \\
& +\frac{16p}{p-1}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n}\lambda^{2n}\gamma \\
& +\frac{8\pi^2}{3}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n-1}\lambda^{2n}\gamma \\
& +2^{-p+5}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n}L^{-2p-1}\lambda^{2n-2}\gamma \\
& +2^{-p+5}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n-p+1}\lambda^{2n}\gamma \\
& +2^{-p+5}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{(n-2)\lambda}L^{-n-p+2}\lambda^{2n-2}\gamma.
\end{aligned}
$$

The latter is nothing but

$$|\triangle a^m| \;\; \leq \;\; C_1 \lambda^{2p-2} L^{-p+2} + C_2 \lambda^{2n} L^{\max(-n,-2p+1,-n-p+2,-2p)} \gamma \;\; .$$

This finishes the derivation of the bound for (5.1.6). This bound can be used, as it is, for the term that includes $\sum_{|k|>\frac{L_m}{\lambda}} \left| \hat{b}^m_{\theta^m}(k) \right|$, in (5.1.7).

CHAPTER 6

Applications of Sparse Time-Frequency Method in Dynamical Systems

For signal analysis in many scientific applications such as biological investigations, the complexity of the underlying physical problem is perplexing, and the appropriate governing equation that describes its dynamics is unknown. While several dominating components could contribute to complex phenomena, it is believed that every dominating component can be characterized by a dynamical system. Even if the characteristics of the underlying system are unknown, data collection makes it possible to roughly sketch the characteristics. Researchers would like to be able to quantify those characteristics and then determine whether the underlying dynamical system is linear or nonlinear, including quantifying the degree of any nonlinearity. This chapter proposes one possible approach to these problems via STFR methods [25, 24, 26].

The definition of linearity, that the output of a system is linearly dependent on the input, is not practical since the governing system is often not known precisely, which makes it difficult even to define what is input and what is output. Furthermore, the solution typically consists of several dominating components. Possibly, they each account for a different physical mechanism. Since some components could be linear and others nonlinear, working on the entire data is not recommended. Instead, the data should be decomposed into several dominating components and each one analyzed separately. Extracting these intrinsic physical components from the data without compromising their hidden physical structure and integrity is highly nontrivial.

In this chapter, we present a method to quantify the nonlinearity of the IMFs given by the STFR method. The main idea is to establish a connection between the IMFs and classical second order differential equations. One of the main results of this

chapter is that we show that each IMF can be associated with a solution of a second order ordinary differential equation of the form $\ddot{x} + p(x,t)\dot{x} + q(x,t) = f(t)$. We do this both theoretically and numerically. We further assume that the coefficients $p(x,t)$, $q(x,t)$ and $f(t)$ are slowly varying with respect to t. Thus, we can freeze these coefficients locally in time and absorb the forcing function into q. This leads to the reduced autonomous second order ODE, i.e. $\ddot{x} + p(x)\dot{x} + q(x) = 0$. Further, we can reformulate the second order ODE in a conservative form: $\ddot{x} + \dot{P}(x) + q(x) = 0$, where $\frac{dP(x)}{dx} = p(x)$. We then have the following weak formulation of the equation by integrating by parts:

$$\left\langle x, \ddot{\phi} \right\rangle - \left\langle P(x), \dot{\phi} \right\rangle + \langle q(x), \phi \rangle = 0,$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product, and ϕ is a smooth test function of compact support. Assuming that $p(x)$ and $q(x)$ have a sparse representation in terms of the polynomial basis, we can represent $P(x)$ and $q(x)$ as follows: $P(x) = \sum_{k=0}^{M} p_k x^{k+1}$, $q(x) = \sum_{k=0}^{M} q_k x^k$ for some integer $M > 0$. Then we obtain the following weak formulation:

$$\left\langle x, \ddot{\phi} \right\rangle - \sum_{k=0}^{M} p_k \left\langle x^{k+1}, \dot{\phi} \right\rangle + \sum_{k=0}^{M} q_k \left\langle x^k, \phi \right\rangle = 0.$$

Based on the above weak formulation, we can design a l^1-based optimization method to solve for p_k and q_k, for an extracted and known IMF,

$$\begin{aligned} (p_k, q_k) \;=\; & \arg\min_{\alpha_k, \beta_k} \gamma \sum_{k=0}^{M} (|\alpha_k| + |\beta_k|) \\ & + \sum_{i=1}^{N} \left| \left\langle x, \ddot{\phi}_i \right\rangle - \sum_{k=0}^{M} \alpha_k \left\langle x^{k+1}, \dot{\phi}_i \right\rangle + \sum_{k=0}^{M} \beta_k \left\langle x^k, \phi_i \right\rangle \right|^2, \end{aligned}$$

where ϕ_i's are smooth test functions of compact support and N is the number of the functions. We will provide some guidance how to choose these test functions optimally.

This method provides a new way to interpret the hidden intrinsic information contained in the extracted IMF of an IS. Depending on the local form of nonlinearity

in $p(x,t)$ and $q(x,t)$, we can define the degree of nonlinearity for each associated IMF. We can also recover accurately the coefficients for the nonlinear terms in p and q. What we sketch in this chapter generalizes a similar concept recently introduced by Huang et. al. [29].

6.1. IMFs and Second Order ODEs

6.1.1. Looking for a physical explanation. The main question that one could legitimately ask is: why should one pick dictionaries with elements looking like $r\cos\theta$? In this chapter we use mathematical theory and numerical examples to confirm the assertion that many physical signals have IMF manifestations. As a result, this specific choice of dictionary is better motivated.

The canonical forms of linear and nonlinear equations used in this dissertation are detailed here before we move onto continuing the assertion. Usually a second order homogeneous Ordinary Differential Equation (ODE) is expressed as

$$u'' + a(x)\,u' + b(x)\,u = 0. \tag{6.1.1}$$

However, in this dissertation, equations like

$$\left(p(x)\,v'\right)' + q(x)\,v = 0, \tag{6.1.2}$$

or

$$v'' + q(x)\,v = 0 \tag{6.1.3}$$

are of more interest in theory[1]. To convert (6.1.1) into (6.1.2), one needs to take $v = u$, $p(x) = e^{\int^x a(\xi)d\xi}$, and $q(x) = b(x)\,e^{\int^x a(\xi)d\xi}$. In order to convert (6.1.1) into (6.1.3), one needs to take $v = e^{\frac{1}{2}\int^x a(\xi)d\xi}u$, and $q(x) = b(x) - \frac{1}{4}a^2(x) - \frac{1}{2}a'(x)$. The latter transformation is interesting since $e^{\pm\frac{1}{2}\int^x a(\xi)d\xi}$ is always positive and will not change the zeros of the solution. Hence, if u has an IMF representation, so does v,

[1] $\frac{d()}{dx}$ and $()'$ are used interchangeably in the chapter. The same thing applies to $\frac{d()}{dt}$ and $\dot{()}$.

and vice versa. In this section, the general form of the nonlinear second order ODE is taken to be as $\ddot{x} + q(x, \dot{x}) = 0$.

As will be shown, the solutions of many of the second order ODEs are IMFs. Using Fourier Transformation to analyze a given IMF would generally produce a large number of Fourier coefficients, suggesting the need for a method that produces a sparser representation.

Theory accompanied by examples show how linear and nonlinear homogeneous second order ODEs have solutions that are essentially IMFs. STFR, combined with the ODE-IMF solution mentality, can help to extract information about the governing ODE of a signal. This is called *System Identification*.

6.1.2. Linear Homogeneous Second Order ODEs. Here, we show how we can have IMF solutions for Linear Second Order ODEs by Prüfer transformation.

6.1.2.1. *Prüfer transformation for Linear Second Order ODEs.* Take the following differential equation in the canonical form

$$\begin{aligned} \tfrac{d}{dx}\left(P(x)\tfrac{du}{dx}\right) + Q(x)u = 0, \\ u(a) = u_0, \\ u'(a) = u_0', \end{aligned} \tag{6.1.4}$$

where $P > 0$, $P \in C^1[a, b]$ and $Q \in C[a, b]$. Remember that it is possible to have $b = \infty$. The solution of this differential equation can be represented in a new coordinate system using the Prüfer transformation [46]

$$\begin{aligned} u &= r\sin\theta, \\ P\tfrac{du}{dx} &= r\cos\theta. \end{aligned} \tag{6.1.5}$$

This transformation[2] explicitly shows that the envelope r is strictly positive:

$$r = \sqrt{u^2 + (Pu')^2} > 0.$$

[2] θ and r are called the Prüfer variables; namely the Prüfer angle and Prüfer radius, respectively.

The last inequality holds true since for any non-zero initial conditions, the solution should not get to zero for both u and $u^{'}$. Using (6.1.5), one can convert (6.1.4) into

$$\begin{cases} \theta^{'} = Q\sin^2\theta + \frac{1}{P}\cos^2\theta, \\ r^{'} = \left(\frac{1}{P} - Q\right) r\cos\theta\sin\theta, \\ u_0 = r_0\sin\theta_0, \\ P\left(a\right)u_0^{'} = r_0\cos\theta_0, \\ \theta_0 \in \left[0, 2\pi\right). \end{cases} \tag{6.1.6}$$

If $Q > 0$, the phase derivative is always positive and θ is strictly increasing. In other words, considering $r > 0$, the solution is nothing but an IMF. This transformation shows that a large class of second order linear ODEs have solutions that are in the form of IMFs. Furthermore, this transformation shows that the oscillatory solutions of Legendre, Hermite, Laguerre and Chebychev equations are all of the IMF type in certain domains. This first shows that IMFs are more prevalent in Physics and Applied Mathematics than previously thought.

In addition, the solution can be an IMF in the complete classical sense. This can be observed as $Pu^{'} = r\cos\theta$ shows that the derivative of the solution u goes to zero only once between two consecutive zeros of the solution itself. In the following, a few examples support the ideas expressed so far.

Example 16. (Linear IMFs) The solution of

$$\frac{d}{dx}\left(\frac{du}{dx}\right) + u = 0 \tag{6.1.7}$$

is an IMF of a constant envelope and constant IF: $u = c_1\cos\left(x\right) + c_2\sin\left(x\right)$.

Example 17. (Chebychev IMFs) The Chebyshev's Differential Equation, for $|x| < 1, \alpha > 0$ is

$$\frac{d}{dx}\left(\sqrt{1-x^2}\frac{du}{dx}\right) + \frac{\alpha^2}{\sqrt{1-x^2}}u = 0. \tag{6.1.8}$$

The solution of this equation is

$$u = c_1 \cos\left(-\alpha \cos^{-1} x\right) + c_2 \sin\left(-\alpha \cos^{-1} x\right) \tag{6.1.9}$$

for some constant envelopes c_1, c_2. Taking $\theta = -\alpha \cos^{-1} x$,

$$\frac{d\theta}{dx} = \frac{\alpha}{\sqrt{1-x^2}} > 0.$$

In other words, the solution of the Chebychev differential equation is a constant envelope Frequency Modulated (FM) IMF. It is seen that the solution of the Chebychev differential equation is sparse if (6.1.9) is considered. In contrast, the presentation of the same solution in the Fourier Transform domain is not sparse at all.

Example 18. (Bessel IMFs) Take the zero order Bessel ODE on $(0, \infty)$

$$x^2 \frac{d^2u}{dx^2} + x\frac{du}{dx} + x^2 u = 0.$$

It is possible to convert this into

$$\frac{d}{dx}\left(x\frac{du}{dx}\right) + xu = 0.$$

Using the Prüfer transformation, one gets

$$\theta' = x \sin^2\theta + \frac{\cos^2\theta}{x} > 0,\ \forall x \in (0, \infty),$$

which clearly shows the IMF behavior of the solution in $(0, \infty)$.

The Prüfer transformation is an excellent mathematical transformation that proves that the solution of certain linear second order ODEs are IMFs. This classic transformation is certainly a great motivation to use STFR methods to extract IMFs produced by ODEs.

6.1.2.2. *Fundamental IMF Solutions of Linear Second Order ODEs.* Previously, the Prüfer transformation was used to express an IMF solution for a general initial value problem. Here, we investigate the *fundamental IMF solutions* of Second Order

ODEs. The solution of $y^{''} + q(x)\, y = 0$ in the form of $r\cos\theta$ is detailed in [40]. Here, we present a formalization of fundamental solutions of such equations. Take the following Second Order ODE in the form of (6.1.3)

$$y^{''} + q(x)\, y = 0, \tag{6.1.10}$$

where $q(x) \in C$. Having this, the solution exists for any initial value problem. Furthermore, all solutions of (6.1.10) could show IMF behavior:

Theorem 9. *$y_1 = r\cos\theta$ and $y_2 = r\sin\theta$ are two linearly independent solutions of (6.1.10) if and only if*

$$\begin{cases} r^{''} - r\theta'^2 + rq = 0, \\ r^2\theta' = 1. \end{cases} \tag{6.1.11}$$

Proof. First assume that $y_1 = r\cos\theta$ and $y_2 = r\sin\theta$ are linearly independent solutions. One can easily put them in (6.1.10) separately to find:

$$\begin{aligned} -\left(r^{''} - r\theta'^2 + rq\right)\sin\theta + \left(-r\theta^{''} - 2r^{'}\theta^{'}\right)\cos\theta &= 0, \\ \left(r^{''} - r\theta'^2 + rq\right)\cos\theta + \left(-r\theta^{''} - 2r^{'}\theta^{'}\right)\sin\theta &= 0, \end{aligned}$$

which has the unique solution

$$\begin{cases} r^{''} - r\theta'^2 + rq = 0, \\ -r\theta^{''} - 2r^{'}\theta^{'} = 0. \end{cases}$$

r cannot be zero since this would make y_1 and y_2 linearly dependent. Hence, $r\theta^{''} + 2r^{'}\theta^{'} = 0 \Rightarrow \left(r^2\theta^{'}\right)^{'} = 0$. $\theta^{'}$ cannot be zero either, since this would make y_1 and y_2 linearly dependent. Without loss of generality, one can take $r^2\theta^{'} = 1$. In the other direction, put $y_1 = r\cos\theta$ and $y_2 = r\sin\theta$ in $y^{''} + q(x)\, y$ and check that based on (6.1.11), the original equation (6.1.10) is satisfied. Finally, one can check independence by constructing the Wronskian

$$W(y_1, y_2) = y_1 y_2^{'} - y_1^{'} y_2 = r^2\theta^{'} = 1 \neq 0.$$

This argument is saying that all solutions of (6.1.10) can be written as

$$y = c_1 r\cos\theta + c_2 r\sin\theta \tag{6.1.12}$$

for some real constants c_1, c_2, and r, θ satisfying (6.1.11). □

Remember that the *fundamental solution* (6.1.12), and the *fundamental conditions* (6.1.11) are not necessarily conditions for the existence of IMFs as solutions. They are proposing solutions represented in $r\cos\theta$ and $r\sin\theta$ format. If certain conditions are satisfied on $q(x)$, then the solutions are IMFs.

We can further analyze the fundamental conditions (6.1.11). Assume that the fundamental solutions are IMFs. Based on $r^2\theta' = 1$, one can observe that the envelope and the instantaneous frequency are not independent from each other. They are a dependent part of the IMF produced by the linear second order ODE. Equation (6.1.11) shows that the IF is strictly positive. Another consequence is that as the IF becomes small, the envelope r increases to compensate, and as the IF increases in time, the envelope is damped.

It is important to note that when the solution of an initial value problem is needed, one should use (6.1.11) with caution. Assume that the solution of (6.1.10) is to be found under the initial conditions

$$y(0) = A,$$
$$y'(0) = B.$$

In (6.1.11), for simplicity and without loss of generality, one can set $\theta(0) = 0$ and $r(0) = 1$. However, there is no way to find $r'(0)$ using the initial conditions given. In fact, $r'(0) \triangleq r_0'$ remains as a free parameter. Using the initial conditions given, the solution (6.1.12) would become

$$y = Ar\cos\theta + \left(B - Ar_0'\right) r\sin\theta.$$

The dependency of θ and r in (6.1.11) would bring about the nonlinear second order ODE

$$r'' - r^{-3} + rq(x) = 0,$$
$$r(0) = 1,$$

having $r'(0)$ as a free parameter. This free parameter is the source of different representations (manifestations) of the same IMF. This can also be seen in

$$\theta(x) = \int^x \frac{d\xi}{r^2(\xi)}. \tag{6.1.13}$$

However, having $r'(0)$ as a free parameter will not deter the uniqueness of the IMF solution itself.

There is significant difference between the Prüfer transformation (6.1.5), and the fundamental solution conditions (6.1.11). The fundamental solution conditions (6.1.11) provide an observation that the solution of *all* linear second order ODEs are necessarily the sum of two linearly independent IMF-like functions. Equation (6.1.11) does not say that the solution is necessarily an IMF. However, the Prüfer transformation (6.1.5) is a good method to perceive whether or not the solution is essentially an IMF.

6.1.2.3. *WKB Theory and IMF Solutions.* If $q(x) \gg 1$ in (6.1.10), WKB theory can be used to find leading order approximate solution(s) of (6.1.10) (see [3]). To see this, we consider the following linear second order ODE:

$$\ddot{x} + b(t)\dot{x} + c(t)x = 0. \tag{6.1.14}$$

This can be rewritten as:

$$\ddot{v} + Q(t)v = 0, \tag{6.1.15}$$

where

$$v = e^{\frac{1}{2}\int_0^t b(\xi)d\xi}x, \quad Q(t) = c(t) - \frac{1}{4}b^2(t) - \frac{1}{2}\dot{b}(t).$$

Assume that $Q(t) \gg 1$. Using the WKB method [3], we can get the asymptotic approximation of $v(t)$,

$$v(t) \sim c_1 \cos\left(\int_0^t \sqrt{Q(\xi)}d\xi\right) + c_2 \sin\left(\int_0^t \sqrt{Q(\xi)}d\xi\right).$$

In terms of the original variables, the solution of (6.1.14) has the form:

$$x(t) \sim e^{-\frac{1}{2}\int_0^t b(\xi)d\xi}\left(c_1 \cos\left(\int_0^t \sqrt{Q(\xi)}d\xi\right) + c_2 \sin\left(\int_0^t \sqrt{Q(\xi)}d\xi\right)\right),$$

which is an IMF. On the other hand, for those IMFs $a(t)\cos\theta(t)$ that do not have intrawave frequency modulation (meaning that both $a(t)$ and $\dot{\theta}(t)$ are smoother than $\cos\theta(t)$), it is easy to see that the coefficients b and c given in (6.1.14) are smooth functions with respect to t. This seems to suggest that there is a close connection between oscillatory solutions of a linear second order ODE with smooth coefficients and IMFs without intrawave frequency modulation. All in all, the WKB method is solely an approximate asymptotic method and not a method proving IMF solutions.

6.1.2.4. *Sturm-Liouville-Type Problems.* Most of the time, the Sturm-Liouville (SL) problem is seen in the theory of Partial Differential equations (PDEs). Hence, analyzing the SL problem from the IMF-STFR point of view could be a good starting point for future research in PDE systems. Remember that a (SL) problem is

$$\begin{gathered} \frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + (\lambda\rho(x) - q(x))u = 0,\ x \in (a,b) \\ \alpha_1 u(a) + \alpha_2 u'(a) = 0,\ (\alpha_1)^2 + (\alpha_2)^2 > 0 \\ \beta_1 u(b) + \beta_2 u'(b) = 0,\ (\beta_1)^2 + (\beta_2)^2 > 0 \end{gathered} \tag{6.1.16}$$

having $p(x) \in C^1$, $p(x) > 0$ on $[a,b]$, $\rho(x) > 0$ on $[a,b]$, and $q, \rho \in C$ on $[a,b]$. If

$$\lambda\rho(x) - q(x) > 0, \tag{6.1.17}$$

then the solution of the SL equation is necessarily of the IMF type. Condition (6.1.17) is satisfied in most of the SL problems if the eigenvalue λ is sufficiently large. Furthermore, this condition is satisfied in certain SL problems even for small eigenvalues.

For example, the one dimensional Laplace equation, in both cylindrical or Cartesian coordinates, has the form of (6.1.16). If the Dirichlet boundary conditions are imposed, then the condition (6.1.17) is satisfied as $q = 0$ and $\lambda_n > 0$. In fact, eigenvalues of symmetric elliptic operators are positive for a Dirichlet boundary condition [19].

6.1.2.5. *Oscillatory Solutions of ODEs in Literature.* As mentioned in [43], it is not hard to show that if $x(t) = a(t)\cos\theta(t)$, then one possible governing differential equation is

$$\ddot{x} + \left(-\frac{\ddot{\theta}}{\dot{\theta}} - 2\frac{\dot{a}}{a}\right)\dot{x} + \left(\dot{\theta}^2 + \frac{\dot{a}\ddot{\theta}}{a\dot{\theta}} + 2\left(\frac{\dot{a}}{a}\right)^2 - \frac{\ddot{a}}{a}\right)x = 0. \tag{6.1.18}$$

Knowing the coefficients in front of $\dot{x}$ and x, one can solve for the unknown envelope a and the phase function θ. However, equation (6.1.18) is not always easy to solve for a and θ, so a better method is needed for finding possible IMF solutions. Furthermore, one needs methods to prove that equations of the form $u'' + a(x)u' + b(x)u = 0$ are sources of IMFs. Some of the theorems and transformations in Oscillation Theory that prove the oscillatory nature of linear equation solutions are reviewed here. In particular, we present theorems from Oscillation Theory literature that show the possible oscillatory nature of linear second order ODEs. Their results can be checked by means of the system introduced in (6.1.5).

The first theorem expresses the interval increase in the phase function if the solution is oscillatory. The second explains how it is possible to understand whether an equation has oscillatory solutions and how one can find the minimum number of zeros. The third helps to determine whether the solution is oscillatory on a semi-infinite domain, which is useful when one is trying to use the fundamental solution conditions (6.1.11).

Theorem. *There is at most one value of* $x_n \in [a, b]$ *such that* $\theta(x_n) = n\pi$, $n \in \mathbb{N}$. $\theta(x)$ *will remain strictly above that in* (x_n, b), *and strictly below that in* (a, x_n) *for* (6.1.4) *using* (6.1.5).

For the proof of this theorem, see [1] theorem 8.4.3.

Theorem. *Let p_1, p_2, g_1 and g_2 be piecewise continuous on $[a,b]$, satisfying $0 < p_2(x) \leq p_1(x)$, $g_1(x) \leq g_2(x)$ $\forall x \in [a,b]$. Let u_1 and u_2 be solutions to the equation $\frac{d}{dx}\left(p_i(x)\frac{du_i}{dx}\right) + g_i(x)u_i = 0$ for $i = 1,2$. We know θ_1 and θ_2 are defined by $\tan\theta_i = \frac{u_i}{p_i u_i'}$. Let $\theta_2(a) \geq \theta_1(a)$. Then $\theta_2(x) \geq \theta_1(x)$ $\forall x \in [a,b]$. Moreover, if $g_1(x) < g_2(x)$ $\forall x \in [a,b]$, then $\theta_2(x) > \theta_1(x)$ $\forall x \in [a,b]$.*

For the proof of this theorem, see [13] Theorem 8.1.2. In this theorem, if one chooses $p_1 = \max\limits_{x\in[a,b]}\{p_2(x)\}$, $g_1 = \min\limits_{x\in[a,b]}\{g_2(x)\}$ and if $\frac{g_1}{p_1} > 0$, then the solution of

$$\frac{d^2u}{dx^2} + \frac{g_1}{p_1}u = 0 \tag{6.1.19}$$

is necessarily oscillatory. As a result, the solution of $\frac{d}{dx}\left(p_2(x)\frac{du}{dx}\right) + g_2(x)u = 0$ is indeed oscillatory and has at least the same number of zeros as (6.1.19). The following theorems express the conditions under which there will be oscillation in a semi-infinite domain [2, 49]. For an in depth discussion on Oscillation Theory, see [2].

Theorem. *1- In (6.1.4), if $\int_a^\infty \frac{dx}{P(x)} < \infty$ and $\int_a^\infty |Q(x)|\,dx < \infty$, then the solution is non-oscillatory on $[a,\infty)$.*

2- In (6.1.4), if $\int_a^\infty \frac{dx}{P(x)} = \infty$ and $\int_a^\infty Q(x)\,dx = \infty$, then the solution is oscillatory on $[a,\infty)$.

Theorem. *(Kneser's Theorem) Consider the differential equation*

$$-u'' + q(x)u = 0 \tag{6.1.20}$$

on $(0,\infty)$, then

1- $\liminf\limits_{x\to\infty}(x^2q(x)) > -\frac{1}{4}$ implies that the solution of 6.1.20 is not oscillatory;

2- $\liminf\limits_{x\to\infty}(x^2q(x)) < -\frac{1}{4}$ implies that the solution of 6.1.20 is oscillatory.

More literature on the oscillatory nature of these equations for periodic cases (i.e. Hill's equation (6.1.15)) can be found in [40].

6.1.3. Nonlinear Second order Autonomous Systems. Here we present IMF solutions of certain nonlinear second order ODEs using second order conservative ODEs as a starting point. These ODEs, which are prevalent in classical physics have closed phase space trajectories, so the solution of a conservative system is necessarily periodic.

Theorem 10. *The periodic solution of a conservative nonlinear second order ODE is necessarily an IMF.*

PROOF. Consider a conservative system $\ddot{x} = F(x)$ where $F(x) = -\frac{dU(x)}{dx}$, $U(x) \in C(x)$. The total energy of the system is $E = \frac{1}{2}\dot{x}^2 + U(x)$. Assume $E > U(x)$ for all values of x in $D = [x_0, x_1]$, except the end points where $E = U(x_0) = U(x_1)$. $\dot{x} = 0$ only at x_0 and x_1. Take $a = \frac{x_0+x_1}{2}$, $b = \frac{-x_0+x_1}{2}$. Consequently, $\frac{x(t)-a}{b}$ is in $[-1, 1]$. Simply take

$$\begin{aligned} for \quad \dot{x} < 0 : \quad & \theta(t) = \arccos\left(\tfrac{x(t)-a}{b}\right) \\ for \quad \dot{x} > 0 : \quad & \theta(t) = \arccos\left(\tfrac{-x(t)+a}{b}\right) + \pi. \end{aligned}$$

As a result

$$\begin{aligned} for \quad \dot{x} < 0 : \quad & \dot{\theta}(t) = \frac{-\dot{x}(t)}{b\sqrt{1-\left(\frac{x(t)-a}{b}\right)^2}}, \\ for \quad \dot{x} > 0 : \quad & \dot{\theta}(t) = \frac{\dot{x}(t)}{b\sqrt{1-\left(\frac{x(t)-a}{b}\right)^2}}, \end{aligned}$$

and $\dot{\theta} > 0$, if $\frac{x(t)-a}{b} \neq \pm 1$. Now, if $\frac{x(t)-a}{b} \to -1$, then $\dot{\theta}(t) \to \sqrt{\frac{\ddot{x}}{b}}$. Remember that $\ddot{x} > 0$ as $\frac{x(t)-a}{b} \to -1$. The same justification can be used if $\frac{x(t)-a}{b} \to 1$. Hence, take

$$\begin{aligned} for \quad \dot{x} \leq 0 : \quad & \theta(t) = \arccos\left(\tfrac{x(t)-a}{b}\right), \\ for \quad \dot{x} \geq 0 : \quad & \theta(t) = \arccos\left(\tfrac{-x(t)+a}{b}\right) + \pi, \end{aligned} \tag{6.1.21}$$

and finally

$$\begin{aligned} & & \dot{x} < 0 : \quad \dot{\theta}(t) &= \frac{-\dot{x}(t)}{b\sqrt{1-\left(\frac{x(t)-a}{b}\right)^2}}, \\ \frac{x(t)-a}{b} = -1, \quad & \dot{x} = 0 : & \dot{\theta}(t) &= \sqrt{\frac{\ddot{x}}{b}}, \\ & & \dot{x} > 0 : \quad \dot{\theta}(t) &= \frac{\dot{x}(t)}{b\sqrt{1-\left(\frac{x(t)-a}{b}\right)^2}}, \\ \frac{x(t)-a}{b} = 1, \quad & \dot{x} = 0 : & \dot{\theta}(t) &= \sqrt{\frac{\ddot{x}}{-b}}. \end{aligned}$$

In other words, the solution of $\ddot{x} = F(x)$ can be represented as $x(t) = a + b\cos\theta(t)$, where a, b are constants and $\theta(t) \in C^1(t)$, $\dot{\theta}(t) > 0$. The period T of the oscillation, using $\dot{\theta}$, can be defined as a real positive number such that $\int_0^T \dot{\theta} dt = 2\pi$. In other words, $T = \int_0^T dt = \int_0^{2\pi} \frac{d\theta}{\dot{\theta}}$. □

Example 19. (Undamped Duffing Equation) Take the equation $\ddot{x} + x + x^3 = 0$. The energy E of the system is $E = \frac{\dot{x}^2}{2} + \frac{x^2}{2} + \frac{x^4}{4}$, the potential energy is $U(x) = \frac{x^2}{2} + \frac{x^4}{4}$. Assume that the domain is $D = [-A, A]$. Due to symmetry, the solution should look like $x(t) = A\cos\theta(t)$. Putting this in the energy equation gives

$$A^2\cos^2\theta + A^2\dot{\theta}^2\sin^2\theta + \frac{1}{2}A^4\cos^4\theta = A^2 + \frac{A^4}{2},$$

and simplifying it results in the IF of the IMF-solution

$$\dot{\theta}^2 = \frac{A^2}{2}\left(1 + \cos^2\theta\right) + 1.$$

The right hand side of this equation is strictly positive; one can take the positive root to have

$$\dot{\theta} = \sqrt{\frac{A^2}{2}\left(1 + \cos^2\theta\right) + 1}.$$

This shows that the solution is a FM signal. Further properties of the solution can be extracted by simplifying the equation using trigonometric identities. The square of the IF would then be

$$\dot{\theta}^2 = 1 + \frac{3}{4}A^2 + \frac{A^2}{4}\cos 2\theta,$$

showing that the peaks and troughs of the signal coincide with the maximum of the IF. In other words, the signal is sharp at peaks and troughs. It is an intrawave. Finally, the period of the system can also be found by

$$T = \int_0^{2\pi} \frac{d\theta}{\sqrt{\frac{A^2}{2}\left(1+\cos^2\theta\right)+1}}.$$

Using the same methodology, we can observe the presence of IMFs in other types of nonlinear systems. For example, for systems with limit cycles, the original equation $\ddot{x} = f(x, \dot{x})$ can be rewritten

$$\dot{x} = y,$$
$$\dot{y} = f(x, y).$$

The intersection of the phase space trajectory with the x-axis happens if and only if $\dot{x} = 0$. If the crossing points are assumed to be x_0, x_1, with $x_0 < x_1$, the value of y will remain with one single sign for $x_0 < x < x_1$. Using the method described in the proof of Theorem 10, one can find proper maps like (6.1.21). In other words, in a system $\ddot{x} = f(x, \dot{x})$, for a well behaved f that guarantees uniqueness and $f(x, y) = 0$ if and only if $x = y = 0$, if there exists a limit cycle (stable or unstable), then the limit cycle solution can be expressed as $x(t) = a + b\cos\theta(t)$ where a, b are constants and $\dot{\theta}(t) > 0$. This claim can be ratified in a similar way to the conservative case. One can characterize the IF in a closed integral form. If the system $\ddot{x} + h(x, \dot{x}) + g(x) = 0$ has a limit cycle, then the solution is $x(t) = a + b\cos\theta(t)$. Take the energy of the system to be $E(t) = \frac{1}{2}\dot{x}^2 + \int g(x)\,dx$. Hence, $\frac{dE}{dt} = -yh(x, y)$, where $\dot{x} = y$. For a period of the limit cycle one has

$$E(T) - E(0) = \int_0^T -h\left(a + b\cos\theta(t), -bw(t)\sin\theta(t)\right) b\omega(t)\sin\theta(t)\,dt = 0.$$

Using the fact that $\omega(t) = \frac{d\theta(t)}{dt}$, the latter would become

$$\int_0^{2\pi} h\left(a + b\cos\theta, -bw(\theta)\sin\theta\right)\sin\theta d\theta = 0.$$

The same methodology can be used to generate IMFs as solutions to other nonlinear ODE systems. Take a system of the form $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, where $\mathbf{x} \in \mathbb{R}^2$, $\mathbf{f} \in \mathbb{R}^2$ is Lipschitz continuous. If the trajectory of the solution in the phase space is a closed convex curve for a specific initial condition $\mathbf{x}_0$, then there is a solution of the form $x_i = a_i + b_i \cos\theta_i(t)$ with $\theta_i(t)$ strictly increasing for $i = 1, 2$. a_i and b_i are constants for $i = 1, 2$. Consequently, if the phase path of the solution is expressed as $F(\mathbf{x}) = Cte$, where Cte is constant real number, then the IFs for each coordinate solution, x_1 and x_2, are related to each other by

$$\frac{d}{dt}F(\mathbf{x}) = 0 \Rightarrow \sum b_i\dot{\theta}_i \sin\theta_i F_{x_i} = 0.$$

Furthermore, for a Hamiltonian system with the convex Hamiltonian $H(x_1, x_2)$, any solution is of the form $x_i = a_i + b_i \cos\theta_i(t)$. Again, the abundance of IMF solutions make the use of adaptive data analysis methods obligatory when ODE-IMF solutions are encountered.

For constant envelope IMFs produced by a nonlinear second order ODE, the IF is not necessary equal to the polar angle time derivative of the phase space trajectory of the ODE. For example, the polar coordinate phase description of a limit cycle $\psi(t)$ can be related to the IMF angle $\theta(t)$ as follows. Assume that the physical coordinate is centralized, hence the physical displacement and velocity of the system $\ddot{x} + f(x, \dot{x}) = 0$ are

$$x = b\cos\theta(t),$$

$$\dot{x} = y = -b\dot{\theta}\sin\theta(t).$$

These variables can also be represented in the polar coordinate as

$$x = a(t)\cos\psi(t),$$

$$y = a(t)\sin\psi(t).$$

Simplifying these equations gives

$$-\dot{\theta}(t)\tan\theta(t)=\tan\psi(t),$$

$$a^2(t)=b^2\left(\cos^2\theta(t)+\dot{\theta}^2\sin^2\theta(t)\right).$$

For a pure circular motion, $\dot{\theta}=1$, one has $-\tan\theta(t)=\tan\psi(t)$ and $a^2(t)=b^2$. However, in the general case, they specify neither the same angle nor the same envelope.

6.2. Nonlinear Degree Analysis

Using the theory explained thus far, here we propose a new method for analyzing the degree of nonlinearity of IMFs decomposed from a multiscale signal and explain an effective optimization method for constructing a second order ODE for each IMF. Moreover, based on the degree of the nonlinearity of the coefficients associated with the second order ODE, we define the degree of nonlinearity for each IMF.

Consider a second order ODE of the following type:

$$\ddot{x}+p(x,t)\dot{x}+q(x,t)=f(t), \tag{6.2.1}$$

where $p(x,t)$, $q(x,t)$ and $f(t)$ are slowly varying with respect to t. For Duffing equation, for example, one observes that $p(x)=0, q(x)=x+x^3$.

Based on the assumption of slow variation, one can freeze $p(x,t)$, $q(x,t)$, and $f(t)$ locally in time over a local time interval (a few periods). Thus (6.2.1) can be replaced by the corresponding autonomous ODE over this local time interval. The function f can be absorbed into q (meaning that we can set $f=0$):

$$\ddot{x}+p(x)\dot{x}+q(x)=0. \tag{6.2.2}$$

6.2.1. A Strong Formulation. To determine the autonomous ODE (6.2.2) locally, we use polynomials to approximate $p(x)$ and $q(x)$.

$$p(x) = \sum_{k=0}^{M} p_k x^k, \quad q(x) = \sum_{k=0}^{M} q_k x^k, \tag{6.2.3}$$

Here, M is the order of polynomials which is given *a priori*, and p_k, q_k are unknown coefficients. One way to get the coefficients p_k, q_k is to substitute (6.2.3) to (6.2.2):

$$\ddot{x} + \sum_{k=0}^{M} p_k \left(x^k\right) \dot{x} + \sum_{k=0}^{M} q_k x^k = 0. \tag{6.2.4}$$

Then p_k, q_k can be obtained by using a least squares method,

$$(p_k, q_k) = \arg\min_{\alpha_k, \beta_k} \left\| \ddot{x} + \sum_{k=0}^{M} \alpha_k \left(x^k\right) \dot{x} + \sum_{k=0}^{M} \beta_k x^k \right\|_2. \tag{6.2.5}$$

To study the degree of nonlinearity, the highest order terms in p and q are considered. Further, it is assumed that the coefficients p_k and q_k are sparse. Since x and $\dot{x}$ are strongly correlated, the least squares solution is unstable to noise perturbation. To stabilize the optimization algorithm, we add an l^1 term to regularize the least squares solution and look for the sparsest representation,

$$(p_k, q_k) = \arg\min_{\alpha_k, \beta_k} \gamma \sum_{k=0}^{M} \left(|\alpha_k| + |\beta_k|\right) + \left\| \ddot{x} + \sum_{k=0}^{M} \alpha_k \left(x^k\right) \dot{x} + \sum_{k=0}^{M} \beta_k x^k \right\|_2^2, \tag{6.2.6}$$

where γ is a parameter to control the sparsity of the coefficients. In order to capture the leading order term, γ is chosen to be $O(1)$. In the following examples, γ is chosen to be 2.

In this method, we must compute $\ddot{x}$ and $\dot{x}$. Unfortunately, this amplifies the error introduced in the approximation of the IMF, x, indicating that another approach is needed.

6.2.2. A Weak Formulation. Here, we use an l^1-based optimization based on a weak formulation for the second order ODE. Let $P(x)$ be the primitive function of

$p(x)$, i.e. $\dot{P}(x) = p(x)\dot{x}$. Then the ODE can be rewritten in a conservation form as

$$\ddot{x} + \dot{P}(x) + q(x) = 0. \tag{6.2.7}$$

Assume the span of time of the signal under scrutiny is $[0, T]$. For any test function $\phi \in C_0^2[0, T]$ that satisfies $\dot{\phi}(0) = \dot{\phi}(T) = 0$, we have the following weak formulation of the equation by performing integration by parts:

$$\left\langle x, \ddot{\phi} \right\rangle - \left\langle P(x), \dot{\phi} \right\rangle + \langle q(x), \phi \rangle = 0. \tag{6.2.8}$$

If $p(x), q(x)$ can be approximated by polynomials as in the same way as they are in (6.2.3), then $P(x)$ and $q(x)$ can be expanded in terms of polynomial basis:

$$P(x) = \sum_{k=0}^{M} p_k x^{k+1}, \quad q(x) = \sum_{k=0}^{M} q_k x^k. \tag{6.2.9}$$

Then we get

$$\left\langle x, \ddot{\phi} \right\rangle - \sum_{k=0}^{M} p_k \left\langle x^{k+1}, \dot{\phi} \right\rangle + \sum_{k=0}^{M} q_k \left\langle x^k, \phi \right\rangle = 0. \tag{6.2.10}$$

Using this formulation, we can design the following optimization problem to solve for p_k and q_k,

$$\begin{aligned} (p_k, q_k) \quad = \quad & \arg\min_{\alpha_k, \beta_k} \gamma \sum_{k=0}^{M} (|\alpha_k| + |\beta_k|) \\ & + \sum_{i=1}^{N} \left| \left\langle x, \ddot{\phi}_i \right\rangle - \sum_{k=0}^{M} \alpha_k \left\langle x^{k+1}, \dot{\phi}_i \right\rangle + \sum_{k=0}^{M} \beta_k \left\langle x^k, \phi_i \right\rangle \right|^2. \end{aligned} \tag{6.2.11}$$

Here, N is the number of the test functions being used. In our computations, we pick $N = 2M$ making sure to have enough measurements to determine the coefficients. The test functions that we use are given below:

$$\phi_i(t) = \begin{cases} \frac{1}{2}\left(1 + \cos\left(\frac{\pi(t-t_i)}{\lambda}\right)\right), & -\lambda < t - t_i < \lambda, \\ 0, & \text{otherwise}, \end{cases}$$

where t_i's $(i = 1, \ldots, N)$ are the centers of the test functions and the parameter λ determines their support. Choosing a large λ would enhance stability. On the other hand, with a large support for ϕ we cannot get the high frequency information of the signal. Since the high frequencies are essential for capturing the nonlinearity of the signal, λ is determined based on the balance between the stability and resolution: we make λ as large as possible without compromising the resolution. In computations, λ is chosen to be $\frac{1}{5}$ of the local period (or wavelength) of the signal. After determining λ, we choose $t_i, i = 1, \ldots, N$ to be uniformly distributed in $[\lambda, T - \lambda]$.

The choice of λ depends on the regularity of the signal that we want to study. If the signal is nearly singular, we should choose a small λ to make sure that the signal's information can be well captured by the test functions. If the test functions $\phi_i(t)$ are chosen to be the classical piecewise linear finite element basis, then the weak formulation is equivalent to the strong formulation if we approximate $\ddot{x}$ and $\dot{x}$ by a second order central difference approximation.

Based on the coefficients recovered from the signal, we can define two indices associated with each IMF to characterize the nonlinearity of this IMF.

Definition 8. (Degrees of Nonlinearity) The degrees of nonlinearity of an IMF are defined to be the following two indices

$$I_1 = \max\{k : p_k \neq 0, k = 0, \ldots, M\}, I_2 = \max\{k : q_k \neq 0, k = 0, \ldots, M\}.$$

This definition shows that the degrees of nonlinearity of the signal correspond to the highest order of the nonlinear terms. For example, $I_1 = 0$ and $I_2 = 1$ corresponds to a linear ODE. When $I_1 > 0$ or $I_2 > 1$, the IMF is nonlinear. The IMF becomes more nonlinear as the indices grow. As well as quantifying the degrees of nonlinearity of the IMF, then, we can also recover the coefficients associated with the leading order nonlinear terms. This information is very helpful in quantifying how nonlinear an IMF is and may have an important implication in engineering and biomedical applications.

If the signal is polluted with noise, the recovery of the coefficients will be influenced by these perturbations. This can be alleviated by setting up a small threshold ν_0 to enforce sparsity of the coefficients so that only coefficients larger than ν_0 are considered significant. This leads to the following modified definition of the degrees of nonlinearity:

(6.2.12)

$$I_1 = \max\{k : |p_k| > \nu_0, k = 0, \ldots, M\}, I_2 = \max\{k : |q_k| > \nu_0, k = 0, \ldots, M\}.$$

In the computations to be presented in the next section, we set $\nu_0 = 0.05$.

The method based on the l^1 regularized least squares performs very well in identifying those nonlinear terms with large coefficients. On the other hand, the l^1 regularization produces a sparse representation of the signal at the expense of the accuracy of the coefficients. In order to improve the accuracy of coefficients recovery, we first identify the dominant coefficients,

$$\Gamma_1 = \{k : |p_k| > \nu_1, k = 0, \ldots, M\}, \Gamma_2 = \{k : |q_k| > \nu_1, k = 0, \ldots, M\}, \tag{6.2.13}$$

(in our computations, ν_1 is chosen to be 0.05), and then solve a least squares problem without l^1 regularization to obtain more accurate coefficients for these dominant terms:

(6.2.14)

$$(p_{k_2}, q_{k_1}) = \arg\min_{\alpha_{k_2}, \beta_{k_1}} \sum_{i=1}^{N} \left| \left\langle x, \ddot{\phi}_i \right\rangle - \sum_{k_2 \in \Gamma_2} \alpha_{k_2} \left\langle x^{k_2+1}, \dot{\phi}_i \right\rangle + \sum_{k_1 \in \Gamma_1} \beta_{k_1} \left\langle x^{k_1}, \phi_i \right\rangle \right|^2.$$

For an accurate signal with no noise, this procedure does help to get more accurate coefficients. However, when the signal is polluted with noise, the IMF that we extract from the signal is not very accurate but the coefficients are very accurate even with the refinement procedure.

Algorithm 5 is used to implement this approach. In this algorithm, we first partition the entire physical domain into a number of sub-domains and localize the signal locally by multiplying a smooth cut-off function. Then we apply the optimizations

Algorithm 5 Degree of Non-Linearity Algorithm

- Calculate the phase function $\theta(t)$ of the signal. Choose K points t_j, $j = 1, \ldots, K$, such that the time variation of P and q is well resolved by the local resolution $(t_{j+1} - t_j)$.
- For $j = 1 : K$
- Extract the signal around the point t_j,

$$f_j(t) = f(t)\,\chi(\theta(t) - \theta(t_j)),$$

where $\chi(t)$ is a cutoff function. In our computations, it is chosen to be

$$\chi(t) = \begin{cases} \frac{1}{2}\left(1 + \cos\left(\frac{t}{\mu}\right)\right), & -\mu\pi < t < \mu\pi, \\ 0, & otherwise. \end{cases}$$

μ is a parameter to control the width of the cutoff function. In this dissertation, we choose $\mu = 3$, which means that for each point, we localize the signal within 3 periods to perform the degrees of nonlinearity analysis.

- Extract the IMF c_j for $f_j(t)$ using an appropriate STFR algorithm.
- Solve the optimization problem (6.2.11) with $x = c_j$ to get the coefficients of the polynomials, $P_j(x)$ and $q_j(x)$.
- (optional) Apply the refinement procedure to update the coefficients.
- End
- Calculate the degrees of nonlinearity of the signal according to (6.2.12).

(6.2.11) and (6.2.14) to the localized signal to extract the local degrees of nonlinearity of the signal; see Algorithm 5.

6.3. Numerical Results

To test its performance, we ran Algorithm 5 on a number of test cases, beginning with a signal generated from the solution of the Van der Pol equation.

Example 20. Consider the Van der Pol Equation

$$\ddot{x} + \left(x^2 - 1\right)\dot{x} + x = 0.$$

The equation is solved from $t = 0$ to $t = 100$ with the initial condition $x(0) = 1, \dot{x}(0) = 0$. We extracted coefficients from the original signal and recorded nonlinearity at different times (see Figure 6.1). In our computations, $M = 10$, which gives a total of 22 coefficients, of which only three are not zero: $p_1 = -1$, $p_2 = \frac{1}{3}$ and $q_1 = 1$. The algorithm achieved the exact recovery of all the coefficients when there was no

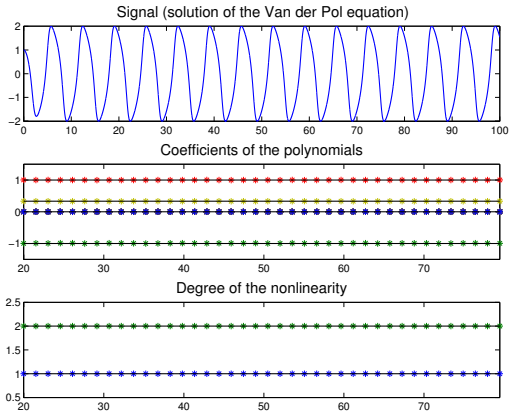

FIGURE 6.1. Top: The solution of the Van der Pol equation; Middle: Coefficients (q_k, p_k) recovered by our method, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

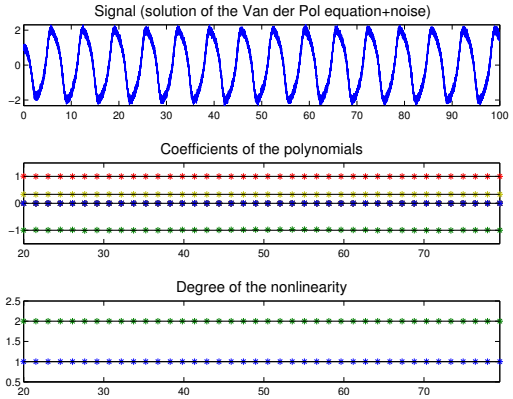

FIGURE 6.2. Top: The solution of the Van der Pol equation with noise $0.1X(t)$, where $X(t)$ is the white noise with standard derivation $\sigma^2 = 1$; Middle: Coefficients (q_k, p_k) recovered by our method, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

noise (see Figure 6.1). When the signal is polluted by noise, the method still gives reasonably accurate results (see Figure 6.2).

Example 21. Consider the Duffing equation

$$\ddot{x} + x + x^3 = 0,$$

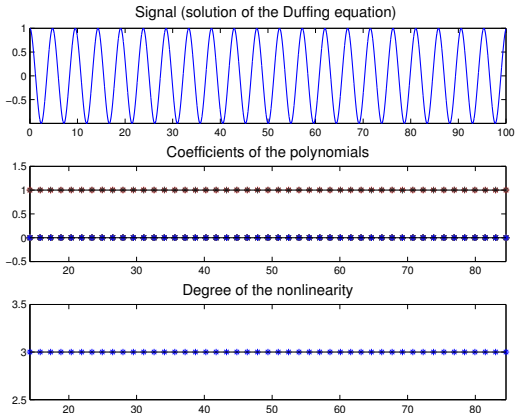

FIGURE 6.3. Top: The solution of the Duffing equation; Middle: Coefficients (q_k, p_k) recovered by our method, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

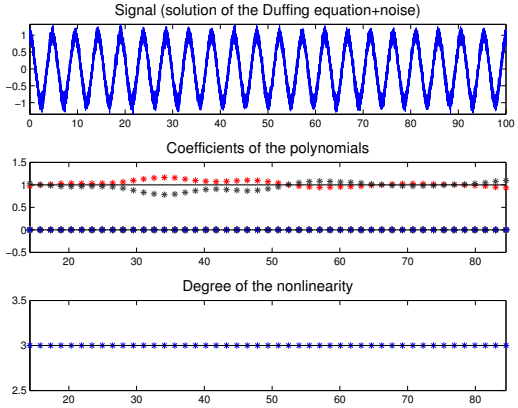

FIGURE 6.4. Top: The solution of the Duffing equation with noise $0.1X(t)$, where $X(t)$ is the white noise with standard derivation $\sigma^2 = 1$; Middle: Coefficients (q_k, p_k) recovered by our method, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

with initial conditions $x(0) = 1$ and $\dot{x}(0) = 0$. Again, the solution is solved from $t = 0$ to $t = 100$ and coefficients were extracted from the original signal and degrees of nonlinearity were recorded (see Figure 6.3). We again use $M = 10$ in our computations, which for this case results in only two coefficients that are not zero: $q_1 = 1$, and $q_3 = 1$. When the signal does not have noise, the recovery is very good for both of the coefficients and for the degrees of nonlinearity (see Figure 6.3). However, when

the signal is polluted by noise, the results for the Duffing equation are not as good as those for the Van der Pol equation: the coefficients are not accurate (see Figure 6.4). This is because the Duffing equation is closer to the linear sinusoidal wave with $q_3 = 0$. Nevertheless, even in this case, our method still gives the correct degrees of nonlinearity (see Figure 6.4).

Example 22. Since both the Van der Pol and Duffing equations are autonomous, the coefficients could be extracted globally, as they do not change over the whole time span. To show the locality of Algorithm 5, we consider an equation that is not autonomous:

$$\ddot{x} + a(t)\left(x^2 - 1\right)\dot{x} + (1 - a(t))\,x^3 + x = 0, \tag{6.3.1}$$

where

$$a(t) = \frac{1}{2}\left(1 - \frac{t - 100}{\sqrt{(t-100)^2 + 400}}\right).$$

The initial condition is that $\dot{x}(0) = 0$, $x(0) = 1$ and the equation is solved over $t \in [0, 200]$. This equation is essentially of the Van der Pol type when t is small ($t < 100$), but as t increases, the equation smoothly changes to the Duffing type equation. When the signal is not polluted by noise, the algorithm captures the time variation of the coefficients accurately (see Figure 6.5) and remain reasonably accurate even when the signal is polluted with noise (see Figure 6.6). The error of the coefficients is relatively large when $t > 100$ in the presence of noise. In this region, the equation is qualitatively of Duffing type, which is more sensitive to noise than the Van der Pol equation, as seen in Examples 21 and 22.

Example 23. To test the method on a more challenging equation in which the coefficients have a sharp change instead of a smooth transition we used the equation

$$\ddot{x} + \frac{1}{2}(1 - \operatorname{sgn}(t - 100))\,\dot{x} + \frac{1}{2}(1 + \operatorname{sgn}(t - 100))\,x^3 + x = 0, \tag{6.3.2}$$

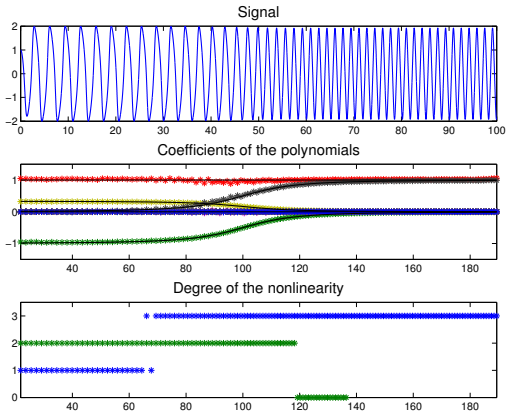

FIGURE 6.5. Top: The solution of the equation given in (6.3.1); Middle: Coefficients (q_k, p_k) recovered by our method, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

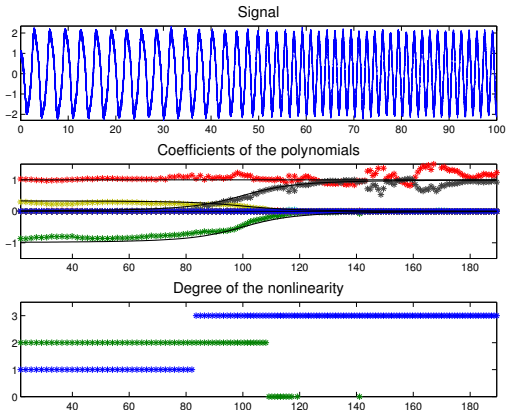

FIGURE 6.6. Top: The solution of the equation given in (6.3.1) with noise $0.1X(t)$, where $X(t)$ is the white noise with standard derivation $\sigma^2 = 1$; Middle: Coefficients (q_k, p_k) recovered by our method, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

where $\operatorname{sgn}(\cdot)$ is the sign function. There is a sharp transition from the Van der Pol equation to the Duffing equation at point $t = 100$ in (6.3.2). Not surprisingly, applying the algorithm to analyze the solution of this equation, generates a large error near $t = 100$, but the method still has reasonable accuracy in the region away from the transition point (see Figure 6.7). While poor accuracy near the transition point means the algorithm cannot locate the transition point faithfully, the method does identify

Algorithm 6 ENO STFR

- Calculate the phase function $\theta(t)$ of the signal. Choose K points t_j, $j = 1, \ldots, K$, such that the time variation of P and q is well resolved by the local resolution $(t_{j+1} - t_j)$.
- For $j = 1 : K$
  - **S1:** Extract the signal centered around the point t_j and also extract the signal to the left and to the right of t_j,

$$\begin{aligned} f_j^c(t) &= f(t)\,\chi^c(\theta(t) - \theta(t_j)), \\ f_j^l(t) &= f(t)\,\chi^l(\theta(t) - \theta(t_j)), \\ f_j^r(t) &= f(t)\,\chi^r(\theta(t) - \theta(t_j)), \end{aligned}$$

    where $\chi^c(t), \chi^l(t), \chi^r(t)$ are cutoff functions

$$\chi^c(t) = \begin{cases} \frac{1}{2}\left(1 + \cos\left(\frac{t}{\mu}\right)\right), & -\mu\pi < t < \mu\pi, \\ 0, & \text{otherwise.} \end{cases}$$

$$\chi^l(t) = \begin{cases} \frac{1}{2}\left(1 + \cos\left(\frac{t}{\mu} + \pi\right)\right), & -2\mu\pi < t < 0, \\ 0, & \text{otherwise.} \end{cases}$$

$$\chi^r(t) = \begin{cases} \frac{1}{2}\left(1 + \cos\left(\frac{t}{\mu} - \pi\right)\right), & 0 < t < 2\mu\pi, \\ 0, & \text{otherwise.} \end{cases}$$

    As before, we choose $\mu = 3$.
  - **S2:** Extract the IMFs c_j^c, c_j^l, c_j^r for $f_j^c(t), f_j^l(t), f_j^r(t)$ respectively.
  - **S3:** Pick up the IMF c_j^* such that the residual $\left\|c_j^\alpha - f_j^\alpha\right\|_2$ is minimized over the choices $\alpha = c, l, r$, i.e.

$$c_j^* = \arg\min_{\alpha \in \{c,l,r\}} \left\|c_j^\alpha - f_j^\alpha\right\|_2.$$

  - **S4:** Solve the optimization problem (6.2.11) with $x = c_j^*$ to get the coefficients of the polynomials, $P_j(x)$ and $q_j(x)$.
  - **S5:** (optional) Apply the refinement procedure to update the coefficients.
- End
- Calculate the degrees of nonlinearity of the signal according to (6.2.12).

that the nonlinearity of the signal changes from the Van der Pol type to the Duffing type. Performance remains qualitatively the same, even in the presence of noise (see Figure 6.8). To improve the accuracy of our method near the transition points, we combine the idea of the ENO method in computing shock waves in fluid dynamics [37] with Algorithm 5. This gives rise to Algorithm 6 which accurately extracts the coefficients and locates of the transition point (see Figure 6.9) and can accurately

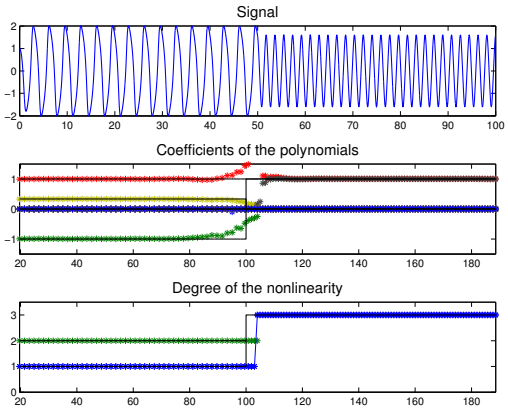

FIGURE 6.7. Top: The solution of the equation given in (6.3.2); Middle: Coefficients (q_k, p_k) recovered by our method, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

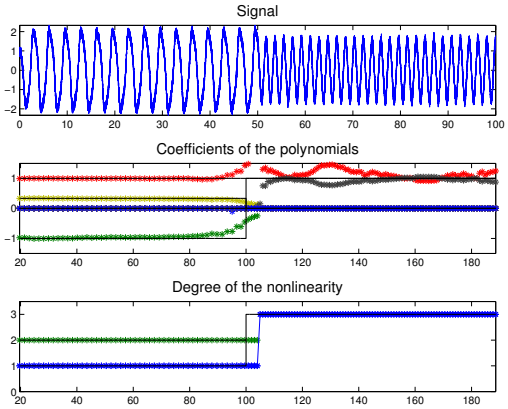

FIGURE 6.8. Top: The solution of the equation given in (6.3.2) with noise $0.1X(t)$, where $X(t)$ is the white noise with standard derivation $\sigma^2 = 1$; Middle: Coefficients (q_k, p_k) recovered by our method, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

approximate degrees of nonlinearity and the transition point even in presence of noise perturbation (see Figure 6.10).

Example 24. Here, we test the method on a signal, $f(t)$, which consists of several components:

$$f(t) = s(t) + \cos\left(\frac{16\pi t}{200}\right) + \frac{t}{100} + 0.1X(t), t \in [0, 200],$$

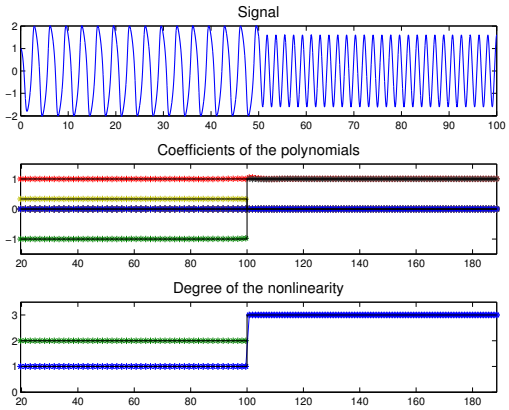

FIGURE 6.9. Top: The solution of the equation given in (6.3.2); Middle: Coefficients (q_k, p_k) recovered by our method together with the trick in ENO method, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

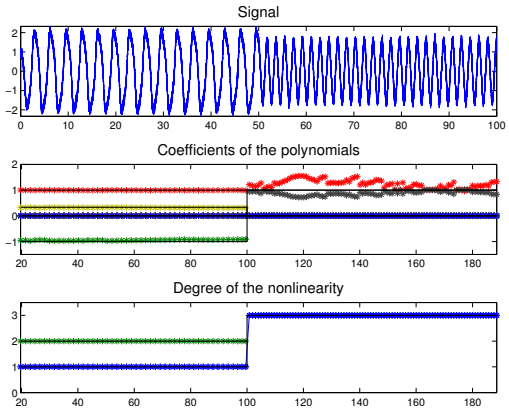

FIGURE 6.10. Top: The solution of the equation given in (6.3.2) with noise $0.1X(t)$, where $X(t)$ is the white noise with standard derivation $\sigma^2 = 1$; Middle: Coefficients (q_k, p_k) recovered by our method together with the ENO type method, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

where $s(t)$ is the solution of the Van der Pol equation with the initial condition $\dot{x}(0) = 0$, $x(0) = 2$ and $X(t)$ is the white noise with standard derivation $\sigma^2 = 1$. After decomposing the signal to tow IMFs (see Figure 6.11) we analyze each IMF to obtain its degrees of nonlinearity (see Figures 6.12 and 6.13). Even for this complex signal, the performance of our method is still reasonably good.

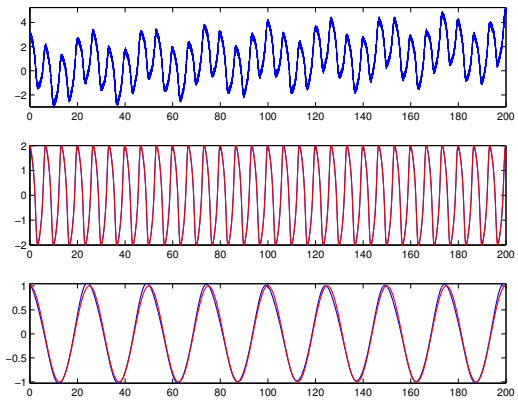

FIGURE 6.11. Top: The signal consists of the solution of the Van der Pol equation and a cosine function and a linear trend and noise $0.1X$; Middle: The IMF extracted from the signal corresponding to the solution of the Van der Pol equation, blue: numerical result; red: exact solution; Bottom: The IMF extracted from the signal corresponding to the cosine function, blue: numerical result; red: exact solution.

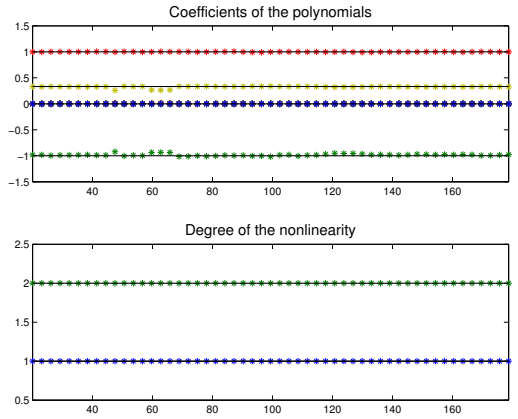

FIGURE 6.12. Top: Coefficients (q_k, p_k) recovered by our method for the first IMF in Fig. 6.11, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

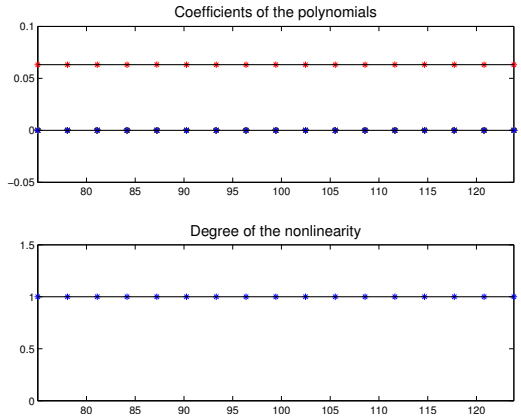

FIGURE 6.13. Top: Coefficients (q_k, p_k) recovered by our method for the second IMF in Fig. 6.11, star points $*$ represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points $*$ represent the numerical results, black line is the exact one.

CHAPTER 7

Bioengineering Applications of an Approximation of Sparse Time-Frequency Method

Cardiovascular Diseases (CVD) are one of the main causes of death in the United States every year [39]. With an increasing number of deaths every year, there is a need to develop new CVD diagnosis methods. Researchers [45] have recently shown that at a given left ventricle (LV) and vascular condition, there exists an optimum Heart-Arterial system coupling that minimizes the LV external pulsatile power. Using the Periodic STFR method (see Figure 7.1), we observed that the IF changes its trend before and after the closure of heart aortic valve (i.e. dicrotic notch). This observation was consistent in different aortic rigidities. Since we observed the IF trend shift in these examples, we proposed a modified version of STFR. We call this modified version as the Intrinsic Frequency (InF) method. The InF method assumes that there are two constant dominant frequencies before and after the dicrotic notch. These frequencies are called Intrinsic Frequencies. Using this modified version of the STFR method, we can identify the optimum heart rate from the aortic pressure wave alone (see Figure 7.2).

This preliminary work [44] raised the question of whether this approach could predict heart performance. Here, we will show that the InF method, can predict heart performance with good medical accuracy.

7.1. Problem Formulation

Ignoring the effect of breathing, we can assume that the pressure waveform at the entrance of aorta is almost periodic, an observation that is supported by the fact that both heart and aorta can be seen as dynamical systems that act at different

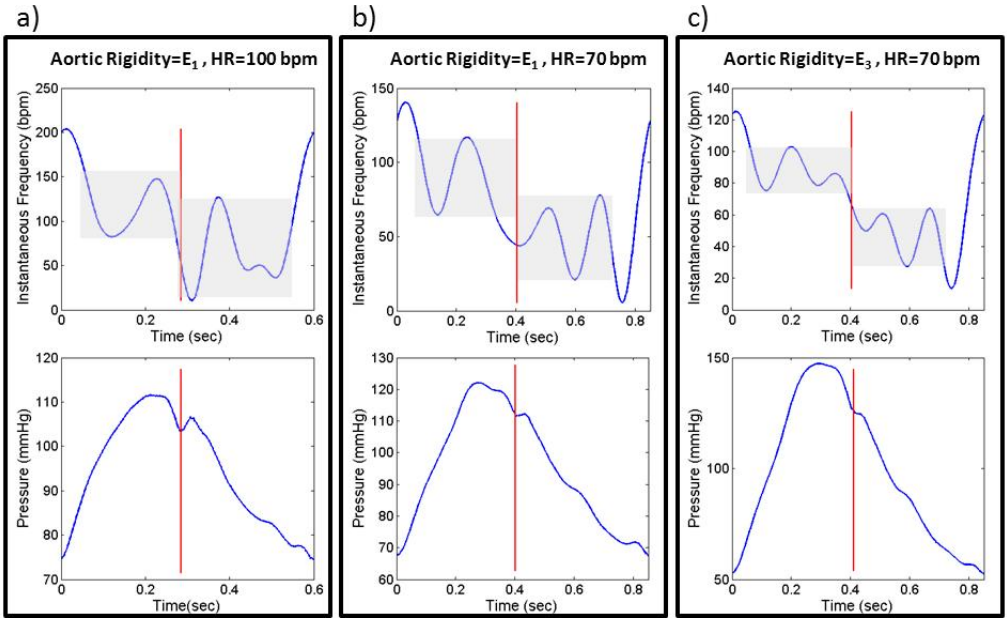

FIGURE 7.1. Instantaneous frequency of the first IMF. The range of instantaneous frequency oscillation (gray band) changes after the dicrotic notch (marked by the red line). a) Instantaneous frequency (top) of the aortic input pressure (bottom) for an aorta with rigidity E1 at HR=100 bpm. b) Instantaneous frequency (top) of the aortic input pressure (bottom) for an aorta with rigidity E1 at HR=70 bpm. c) Instantaneous frequency (top) of the aortic input pressure (bottom) for an aorta with rigidity E3 at HR=70 bpm (E3 = 1.5E1) .

characteristic frequencies. Since we know that reflected waves in the cardiovascular system will distort the trend and introduce subharmonic distortion [45], we need to be able to extract the trend of the pressure waveform with some adaptive method.

Using STFR terminology, in this chapter we are trying to extract a single Intrinsic Mode Function (IMF) from the pressure wave signal. Physically and biomedically, we believe that this IMF conveys the coupling characteristics of the heart-aorta system. However, the IMF that we are trying to extract might have a sharp edge at the location of the dicrotic notch (a sudden drop in pressure that occurs at the instant of aortic valve closure). Hence, in this chapter, the definition of an IMF is slightly abused. However, we still call it an IMF. Attempting to extract this IMF using EMD or STFR would fail or produce a blur extraction, primarily because the change from one frequency regime before the dicrotic notch into another after the closure

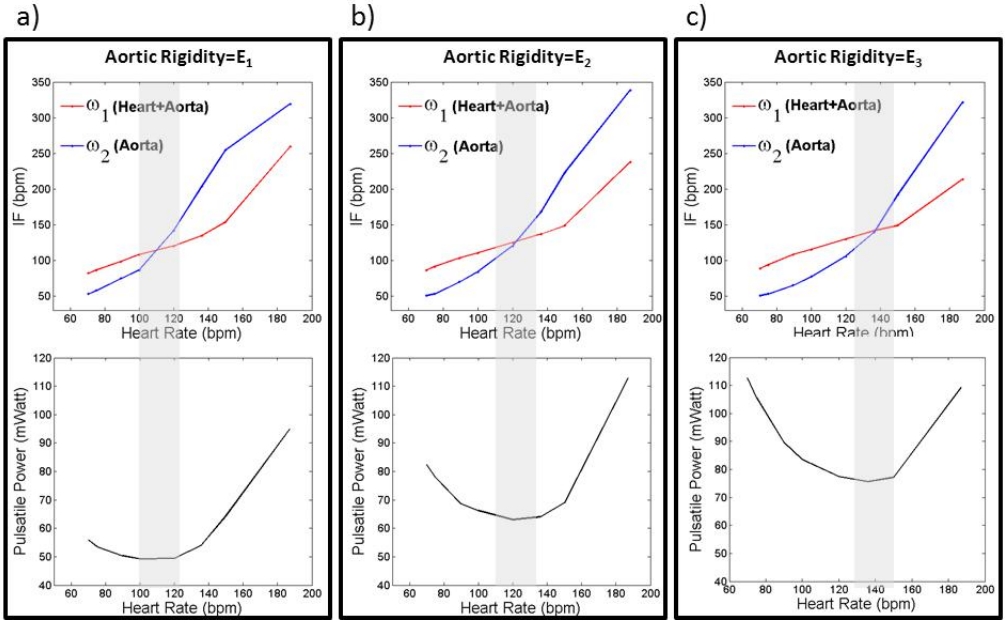

FIGURE 7.2. Intrinsic frequencies Vs. HR (top graphs) with the corresponding pulsatile power vs. HR (bottom graphs). ω_1 (red) is the InF for coupled heart+aorta and ω_2 (blue) is the InF for the decoupled aorta. a) Aortic rigidity is E1, the gray band shows that the two IF curves cross each other at the optimum HR (≈110 bpm). b) Aortic rigidity is E2, the two InF curves cross each other at the optimum HR (≈120 bpm). c) Aortic rigidity is E3, two InF curves cross each other at the optimum HR (≈140 bpm).

of the heart valve is accompanied by an abrupt change in frequency of the whole cardiovascular system or by a discontinuity in the first time derivative of the pressure waveform at the dicrotic notch. In the best case, using STFR methods would just capture a vague picture of the instantaneous frequency response of the system, which is good solely for qualitative interpretation and an initial guess for the possible InFs. As a result, we use a modified version of the STFR method that has less mathematical regularity and focuses on the more basic Intrinsic Frequency (InF) rather than on the instantaneous frequency.

It is assumed that before and after the dicrotic notch, we have the following simple waveforms for the general IMF of the aortic pressure wave at time t:

$$s_i = a_i \cos \omega_i t + b_i \sin \omega_i t + \bar{p}, \quad i = 1, 2. \tag{7.1.1}$$

This assumption has shown its credibility as an index to characterize the heart and cardiovascular diseases [44]. In this formula, $i = 1$ corresponds to the behavior of the IMF before the valve closure, and $i = 2$ to the behavior of the IMF after that. Here, a_i, b_i are constants and correspond to the envelopes of the IMF. The constants ω_1, ω_2 correspond to the InFs of the IMF. $\bar{p}$ is the mean pressure during the heart beat period. Equations (7.1.1) do not clearly show the physics of the problem, however. As the IMF is composed of two different sinusoids, continuous at the dicrotic notch, we can write (7.1.1) in a more compact form. Take $[0, T]$ to be the whole period of the pressure wave and T_0 as the location of the dicrotic notch: $0 < T_0 < T$. Also, define the *indicator function* as

$$\mathbf{1}_{[x,y)}(t) = \begin{cases} 1, & x \leq t < y, \\ 0, & else. \end{cases}$$

Now, the main IMF of the pressure waveform can be expressed as

$$\begin{aligned} S(a_i, b_i, \bar{p}, \omega_i; t) &= (a_1 \cos \omega_1 t + b_1 \sin \omega_1 t + \bar{p})\, \mathbf{1}_{[0,T_0)}(t) + \\ &\quad (a_2 \cos \omega_2 t + b_2 \sin \omega_2 t + \bar{p})\, \mathbf{1}_{[T_0,T)}(t). \end{aligned} \tag{7.1.2}$$

If $0 \leq t < T_0$, one would get the part of the IMF corresponding to the action of heart and aorta before the valve closure, i.e. $s_1 = a_1 \cos \omega_1 t + b_1 \sin \omega_1 t + \bar{p}$. On the other hand, if $T_0 \leq t < T$, the part of the IMF that reflects the behavior of the aorta after the valve closure is depicted by $s_2 = a_2 \cos \omega_2 t + b_2 \sin \omega_2 t + \bar{p}$. In general, we are interested in extracting the IMF (7.1.2) and the corresponding InFs ω_1, ω_2.

At this stage the goal is to extract the IMF carrying most of the energy and consequently the InFs, ω_1, ω_2, from the observed aortic pressure waveform $f(t)$. Taking t

as a continuous variable, one can use least squares minimization to find the unknowns.

$$\underset{a_i,b_i,\omega_i,\bar{p}}{minimize} \qquad \|f(t) - S(a_i, b_i, \bar{p}, \omega_i; t)\|_2^2$$

(7.1.3)

$$subject\ to \qquad \begin{aligned} a_1 \cos\omega_1 T_0 + b_1 \sin\omega_1 T_0 &= a_2 \cos\omega_2 T_0 + b_2 \sin\omega_2 T_0, \\ a_1 &= a_2 \cos\omega_2 T + b_2 \sin\omega_2 T. \end{aligned}$$

In this optimization problem, $\|g(t)\|_2^2 = \int_0^T |g(t)|^2 dt$. The first linear condition in this optimization enforces the continuity of the extracted IMF at the dicrotic notch. The second one imposes the periodicity. In practice, as the data is sampled on discrete temporal points, one must solve the discrete version of (7.1.3). Assume that the data is sampled on time instances $0 = t_1 < t_2 < \ldots < t_n = T$, then one can convert problem (7.1.3) into a discrete least squares of the form

$$\underset{a_i,b_i,\omega_i,\bar{p}}{minimize} \qquad \sum_{j=1}^{n} \left(f(t_j) - S(a_i, b_i, \bar{p}, \omega_i; t_j)\right)^2$$

(7.1.4)

$$subject\ to \qquad \begin{aligned} a_1 \cos\omega_1 T_0 + b_1 \sin\omega_1 T_0 &= a_2 \cos\omega_2 T_0 + b_2 \sin\omega_2 T_0, \\ a_1 &= a_2 \cos\omega_2 T + b_2 \sin\omega_2 T. \end{aligned}$$

Unfortunately, problem (7.1.4) is not convex. In the next subsection we will introduce a brute-force algorithm to solve it.

7.2. Algorithm

Tackling a non-convex problem like (7.1.4) is not easy. In Algorithm 7, we break down the problem into a convex part and a global domain search [35]. Optimization software handles the convex part. Initially, we used Matlab CVX package [22] to solve the convex part. However, later we implemented a C++ code that was based on QR decomposition [18]. The domain search part is nothing but the brute-force part of the algorithm. For this algorithm, the frequency domain is

$$\mathcal{D}_{fr} = \{(\omega_1, \omega_2) \,|0 < \omega_1 \leq C, 0 < \omega_2 \leq C\}, \tag{7.2.1}$$

Algorithm 7 Intrinsic Frequency Algorithm

(1) Discretize $\mathcal{D}_{fr}$ in to a uniform $r \times r$ mesh $\overline{\mathcal{D}}_{fr}$, $r \in \mathbb{N}$,

$$\overline{\mathcal{D}}_{fr} = \left\{ \left(\omega_1^l, \omega_2^m \right) | \omega_1 = \frac{l}{r} C, \omega_2 = \frac{m}{r} C; l, m \in \{0, 1, \ldots, r\} \right\}.$$

(2) For all $l, m \in \{0, 1, \ldots, r\}$ solve

$$\begin{array}{ll} \underset{a_i, b_i, \bar{p}}{minimize} & \sum_{j=1}^{n} \left(f(t_j) - S\left(a_i, b_i, \bar{p}, \omega_1^l, \omega_2^m; t_j \right) \right)^2 \\ subject\,to & \begin{array}{rcl} a_1 \cos \omega_1 T_0 + b_1 \sin \omega_1 T_0 & = & a_2 \cos \omega_2 T_0 + b_2 \sin \omega_2 T_0, \\ a_1 & = & a_2 \cos \omega_2 T + b_2 \sin \omega_2 T. \end{array} \end{array} \tag{7.2.2}$$

and store $P\left(\omega_1^l, \omega_2^m \right) = \sum_{j=1}^{n} \left(f(t_j) - S\left(a_i^*, b_i^*, \bar{p}^*, \omega_1^l, \omega_2^m; t_j \right) \right)^2$ for minimizers $a_i^*, b_i^*, \bar{p}^*$.

(3) Find the intrinsic frequencies (InFs)

$$(\omega_1^*, \omega_2^*) = \underset{l,m}{argmin} \left(P\left(\omega_1^l, \omega_2^m \right) \right).$$

which is characterized by a constant parameter, C, that is set from the physics of the problem. The basic idea behind the Algorithm 7 is to freeze (ω_1, ω_2), solve (7.1.4), and find the minimum of the function

$$P(\omega_1, \omega_2) = \sum_{j=1}^{n} \left(f(t_j) - S(a_i^*, b_i^*, \bar{p}^*, \omega_i; t_j) \right)^2,$$

where $a_i^*, b_i^*, \bar{p}^*$ are the values upon which (7.1.4) is minimized for a fixed vector (ω_1, ω_2). We collect all possible values of $P(\omega_1, \omega_2)$ and find the minimum of them among (ω_1, ω_2). The minimizer of P over (ω_1, ω_2) would then be the InFs that we are looking for.

The second step of Algorithm 7 is just solving a linearly constrained least squares algorithm. This brute-force algorithm can also be made parallel since step 2 can be solved separately for different (l, m) pairs.

7.3. Convergence Analysis

In this section, we analyze the convergence properties of Algorithm 7. In order to discuss the algorithm's convergence and accuracy, we need the following lemma [8] and theorem .

Lemma 5. *The minimum of a function can first be found over a few variables and then over the remaining ones:*

$$\inf_{x,y} f(x,y) = \inf_{x} \left(\inf_{y} f(x,y) \right).$$

This lemma allows us to first find the minimization (7.1.4) on $a_i, b_i, \bar{p}$ and then on ω_1, ω_2.

Further, we need to make sure that the second step of the algorithm has a unique minimizer, which can be provided by this theorem [18]:

Theorem 11. *If $A \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{p\times n}$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $d \in \mathbb{R}^p$ where $p \leq n$, $n \leq m+p$ and $\begin{bmatrix} A \\ C \end{bmatrix}$ is of full column rank, then the optimization problem*

$$\begin{aligned} \underset{x}{minimize} \quad & \|Ax-b\|_2^2 \\ subject\,to \quad & Cx = d \end{aligned}$$

has a unique solution.

As the algorithm freezes the frequency parameters (ω_1, ω_2) and then solves a least-squares problem on other variables, we can form a notation similar to that used in Theorem 11. Take the matrix A to be

$$A = \begin{pmatrix} \cos\omega_1 t_1 & \sin\omega_1 t_1 & 0 & 0 & 1 \\ \cos\omega_1 t_2 & \sin\omega_1 t_2 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cos\omega_1 t_{n_0} & \sin\omega_1 t_{n_0} & 0 & 0 & 1 \\ 0 & 0 & \cos\omega_2 t_{n_0+1} & \sin\omega_2 t_{n_0+1} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cos\omega_2 t_n & \sin\omega_2 t_n & 1 \end{pmatrix},$$

and the matrix C and vector x to be

$$C = \begin{pmatrix} \cos\omega_1 t_{n_0} & \sin\omega_1 t_{n_0} & -\cos\omega_2 t_{n_0} & -\sin\omega_2 t_{n_0} & 0 \\ 1 & 0 & -\cos\omega_2 t_n & -\sin\omega_2 t_n & 0 \end{pmatrix}, \; x = \begin{pmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ \bar{p} \end{pmatrix}.$$

Sample points $t_1, \ldots, t_{n_0}$ correspond to the trend before the dicrotic notch and $t_{n_0+1}, \ldots, t_n$ to the points after it. Finally, the vector b would be the observed sampled signal, $\{b\}_j = f(t_j), j = 1, 2, \ldots, n$ and $d = \mathbf{0}$. These matrices and vectors satisfy the conditions of Theorem 11. Hence, the second step of the algorithm always has a unique solution. This fact, combined with Lemma 5, guarantee that Algorithm 7 always has at least one unique solution for the minimization problem (7.1.4).

7.3.1. No Noise. Assume there is no noise in the observation and the observed signal is exactly of type (7.1.2). As a result, the signal can be expressed as $f = \bar{A}\bar{x}, \bar{C}\bar{x} = 0$ for some $(\bar{\omega}_1, \bar{\omega}_2)$ and $\bar{x}$. If there is a well resolved $\overline{\mathcal{D}}_{fr}$, then for some l, m, one would obtain $\left(\omega_1^l, \omega_2^m\right) = (\bar{\omega}_1, \bar{\omega}_2)$. At these specific frozen frequencies, the solution of the second step of the algorithm is nothing but $\bar{x}$, based on Theorem 11. In detail, at this step of the algorithm, one is in fact solving

$$\begin{aligned} \underset{x}{minimize} \quad & \left\| \bar{A}x - \bar{A}\bar{x} \right\|_2^2 \\ subject\,to \quad & \bar{C}x = 0. \end{aligned}$$

These facts combined with the definition of $P(\omega_1, \omega_2)$ and Lemma 5 guarantee the existence of at least one minimizer.

Furthermore, this minimizer is unique. In fact, if there exists another set of x and (ω_1, ω_2) as the solution of problem (7.1.4), namely $\bar{x}'$ and $\left(\bar{\omega}_1', \bar{\omega}_2'\right)$, then the

two trends, $\bar{A}\bar{x}$ and $\bar{A}'\bar{x}'$, arising from these parameters must be equal. For a finely-sampled observation f, equality of these trends implies the equality of the parameters $\bar{x}' = \bar{x}$, $\left(\bar{\omega}_1', \bar{\omega}_2'\right) = (\bar{\omega}_1, \bar{\omega}_2)$. In short, we can state the following theorem:

Theorem 12. *In the absence of noise, if the observed signal is of the form* (7.1.2)*, for a well resolved $\overline{\mathcal{D}}_{fr}$, Algorithm 7 finds the unique minimizer of*

$$\underset{a_i, b_i, \omega_i, \bar{p}}{minimize} \qquad \sum_{j=1}^{n} \left(f\left(t_j\right) - S\left(a_i, b_i, \bar{p}, \omega_i; t_j\right)\right)^2$$

$$subject\ to \qquad \begin{aligned} a_1 \cos \omega_1 T_0 + b_1 \sin \omega_1 T_0 &= a_2 \cos \omega_2 T_0 + b_2 \sin \omega_2 T_0, \\ a_1 &= a_2 \cos \omega_2 T + b_2 \sin \omega_2 T. \end{aligned}$$

7.3.2. Noisy measurements. Assume here, that the IMF (7.1.2) is polluted with noise that is independent of the IMF itself. In other words, taking the noise to be ε, $f = \bar{A}\bar{x} + \varepsilon, \bar{C}\bar{x} = 0$. Here, the algorithm will not extract the exact values of $(\bar{\omega}_1, \bar{\omega}_2)$ and $\bar{x}$, but it is possible to find an error bound on the distance between the extracted and the true IMF. If x^* and (ω_1^*, ω_2^*) are the extracted values by the algorithm, one can write

$$\left\|A^* x^* - \bar{A}\bar{x} - \varepsilon\right\|_2 \leq \left\|Ax - \bar{A}\bar{x} - \varepsilon\right\|_2 \leq \|\varepsilon\|_2 .$$

The first inequality comes from the fact that any set of x and (ω_1, ω_2), where $Cx = 0$, is a feasible point; consequently, the second inequality follows if x and (ω_1, ω_2) are assigned the values of $\bar{x}$ and $(\bar{\omega}_1, \bar{\omega}_2)$. Now, it is not hard to show by triangle inequality that

$$\left\|A^* x^* - \bar{A}\bar{x}\right\|_2 \leq \left\|A^* x^* - \bar{A}\bar{x} - \varepsilon\right\|_2 + \|\varepsilon\|_2 \leq 2\|\varepsilon\|_2 .$$

Using this we can state the following theorem:

Theorem 13. *In the presence of noise that is independent from the trend* (7.1.2)*, for a well resolved $\overline{\mathcal{D}}_{fr}$, Algorithm 7 finds the minimizer of* (7.1.4) *with an error having at most the same order as the noise.*

How much the solution of the noisy problem differs from the real solution depends on the noise level. If the noise level $\|\varepsilon\|$ is sufficiently small, then the distance between x^*, (ω_1^*, ω_2^*) and $\bar{x}$, $(\bar{\omega}_1, \bar{\omega}_2)$ is also of $O(\|\varepsilon\|)$ [4]. In practice, since the 2-norm of the trend is large compared to the noise level, the relative error in finding the trend is extremely low. In mathematical terms

$$\frac{\left\|A^*x^* - \bar{A}\bar{x}\right\|_2}{\left\|\bar{A}\bar{x}\right\|_2} \leq 2\frac{\|\varepsilon\|_2}{\left\|\bar{A}\bar{x}\right\|_2}.$$

So, in general, Algorithm 7 can extract the needed IFs with good accuracy, even in the presence of noise perturbation. In real data, where a lot of reflected waves are superposed with the heart-aorta wave system, a signal could be a combination of many IMFs. Usually these waves have higher frequencies compared to the main IMF. In the next subsection, we answer this question in detail.

7.3.3. Other IMFs. Here, we investigate this scenario in detail. For the sake of simplicity, we assume that the added IMFs are of high frequency and that time is a continuous variable and the signal is not sampled on discrete points (it is a continuous variable). Take the recorded signal to be

$$g(t) = \bar{S}(t) + D_M(t), \tag{7.3.1}$$

where $\bar{S}(t)$ is the IMF of form (7.1.2), and $D_M(t)$ is a combination of IMFs with higher frequencies compared to $\bar{S}(t)$. Without loss of generality, take $D_M(t)$ to have a Fourier series of the form

$$D_M(t) = \sum_{n>M} \left(A_n \cos\frac{2\pi nt}{T} + B_n \sin\frac{2\pi nt}{T}\right). \tag{7.3.2}$$

Implicitly, we have assumed that the added IMFs are of high-frequency nature. Having this terminology in mind we can state the following theorem.

Theorem 14. *The optimum curve $S^*(t)$, which is the solution of the minimization problem*

$$\underset{S(t)}{\textit{minimize}} \quad \left\| S(t) - \bar{S}(t) - D_M(t) \right\|_{L^2}^2$$

(7.3.3)

$$\textit{subject to} \quad \begin{array}{c} S(t) \textit{ is continuous at } T_0, \\ S(t) \textit{ is periodic}, \end{array}$$

satisfies

$$\left\| S^*(t) - \bar{S}(t) \right\|_{L^2} \leq \frac{8 \left\| D_M^{(m)}(t) \right\|_{L^1}}{(m-1)\sqrt{T}\,(M+1)^{m-1}}, \tag{7.3.4}$$

provided that $D_M(t) \in C^m$, and $m \geq 2$.[1]

Proof. As $\bar{S}(t)$ is a feasible point, the minimizer $S^*(t)$ of (7.3.3) must satisfy

$$\left\| S^*(t) - \bar{S}(t) - D_M(t) \right\|_{L^2}^2 \leq \left\| \bar{S}(t) - \bar{S}(t) - D_M(t) \right\|_{L^2}^2 = \left\| D_M(t) \right\|_{L^2}^2. \tag{7.3.5}$$

Define the Fourier series of $S^*(t) - \bar{S}(t)$ as

$$S^*(t) - \bar{S}(t) = \frac{1}{2}\Delta a_0 + \sum_{n=1}^{\infty} \left(\Delta a_n \cos \frac{2\pi n t}{T} + \Delta b_n \sin \frac{2\pi n t}{T} \right).$$

Hence, we have

$$\begin{aligned} S^*(t) - \bar{S}(t) - D_M(t) &= \frac{1}{2}\Delta a_0 + \sum_{n=1}^{M} \left(\Delta a_n \cos \frac{2\pi n t}{T} + \Delta b_n \sin \frac{2\pi n t}{T} \right) \\ &\quad + \sum_{n=M+1}^{\infty} \left((\Delta a_n - A_n) \cos \frac{2\pi n t}{T} + (\Delta b_n - B_n) \sin \frac{2\pi n t}{T} \right). \end{aligned}$$

[1] This bound can be made sharper if $D_M(t) \in C^m$ and $D_M^{(m+1)}(t) \in L^2_{(0,T)}$. Here the p-norm is defined as $\|\vartheta(t)\|_{L^p} = \left(\int_0^T |\vartheta(t)|^p \, dt \right)^{\frac{1}{p}}$ for $p \geq 1$.

Using Parseval's identity and (7.3.5) gives

$$\begin{aligned}\left\|S^*(t)-\bar{S}(t)-D_M(t)\right\|_{L^2}^2 &= \left\|S^*(t)-\bar{S}(t)\right\|_{L^2}^2+\|D_M(t)\|_{L^2}^2 \\ &\quad -2T\sum_{n=M+1}^{\infty}\left(\Delta a_n A_n+\Delta b_n B_n\right)\leq\|D_M(t)\|_{L^2}^2.\end{aligned}$$

Simplifying and using triangle inequality results in

$$\left\|S^*(t)-\bar{S}(t)\right\|_{L^2}^2\leq 2T\sum_{n=M+1}^{\infty}\left(|\Delta a_n|\,|A_n|+|\Delta b_n|\,|B_n|\right). \tag{7.3.6}$$

Since $S^*(t)-\bar{S}(t)$ is continuous, $|\Delta a_n|,|\Delta b_n|\leq\frac{\left\|S^*(t)-\bar{S}(t)\right\|_{L^1}}{T}$. Using Cauchy-Schwartz inequality gives $|\Delta a_n|,|\Delta b_n|\leq\frac{\left\|S^*(t)-\bar{S}(t)\right\|_{L^2}}{\sqrt{T}}$. On the other hand, as $D_M(t)\in C^m$, then $|A_n|,|B_n|\leq 2\frac{\left\|D_M^{(m)}(t)\right\|_{L^1}}{Tn^m}$. Using these estimates and bounding the sum by an integral, (7.3.6) will result in

$$\left\|S^*(t)-\bar{S}(t)\right\|_{L^2}\leq\frac{8\left\|D_M^{(m)}(t)\right\|_{L^1}}{(m-1)\sqrt{T}\,(M+1)^{m-1}}.$$

□

Remark 2. The bound (7.3.4) shows that as the minimum frequency M of the IMFs increases, then the optimum curve $S^*(t)$ and the true curve $\bar{S}(t)$ get closer. This bound is in fact a measure of the "Scale Separation Condition" mentioned in [25, 24, 26]. In simple words, if the IMFs are orthogonal to the original IMF, then the extracted optimum curve $S^*(t)$ is almost the true IMF $\bar{S}(t)$. Hence, the frequencies found in $S^*(t)$ are close to true InF values. Interestingly, in deriving this bound, we have not used the structure of the main IMF. This bound is in general an orthogonality condition. In practice, we find shorter bounds than (7.3.4). In other words, the algorithm works much better than the bound error estimate.

7.4. Synthetic Examples

In this section we work on synthetic examples to show the performance of the proposed Algorithm 7.

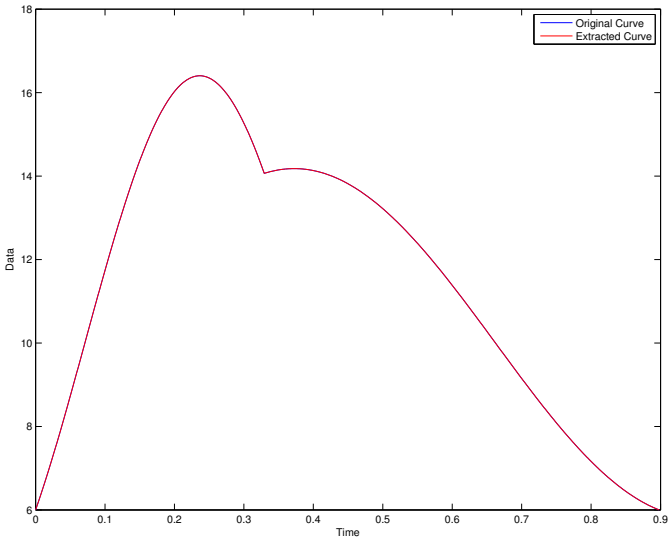

FIGURE 7.3. Synthetic Data, with no Noise and Well-Resolved Domain

Example 25. Assume that the trend has intrinsic frequencies of $\omega_1 = 9.5$ and $\omega_2 = 5.42$. The envelope of the first part of the trend is taken to be $\sqrt{41}$, and the envelope of the second sinusoid is taken in a way that it matches a signal of period 0.898 and a dicrotic notch at $T_0 = 0.3293$. The average pressure is $\bar{p} = 10$. $\overline{\mathcal{D}}_{fr}$ is defined in a way that the resolution of the frequency domain, $\min_{l \neq m} \left(\left| \omega_1^l - \omega_1^m \right|, \left| \omega_2^l - \omega_2^m \right| \right)$, is 0.08. For this well resolved domain, the extraction is accurate up to machine error precision: the curves are indistinguishable (see Figure 7.3). The IFs are found with no error.

Example 26. To test the algorithm for a case in which $\overline{\mathcal{D}}_{fr}$ is not well-resolved, we use the signal from Example 25 and define $\overline{\mathcal{D}}_{fr}$ so that the resolution of the frequency domain is 0.1. This resulted in a faithful extraction of the curve, with less than 0.4% relative error[2] (see Figure 7.4). The InFs are $\omega_1^* = 9.4$ and $\omega_2^* = 5.2$, which have less than 5% relative error.

In the next example, we will investigate the effect of noise on the extracted IFs. Here, $\overline{\mathcal{D}}_{fr}$ is well-resolved.

[2]Relative error for the signal is defined as $\frac{\left\| S^*(t) - \bar{S}(t) \right\|_{L^2}}{\left\| \bar{S}(t) \right\|_{L^2}}$

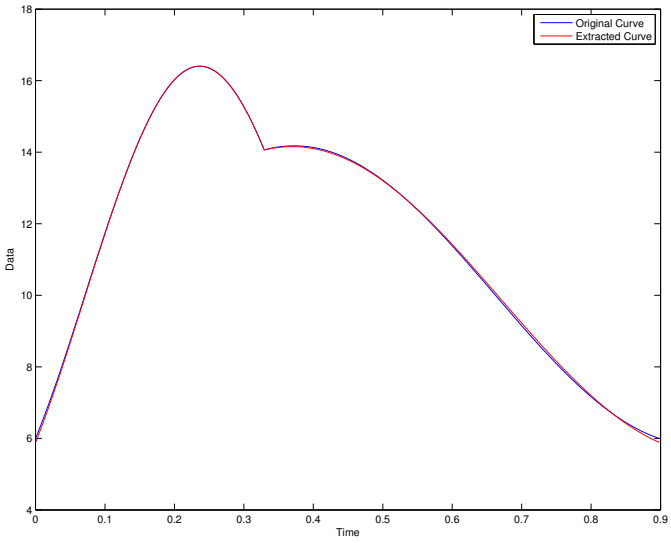

FIGURE 7.4. Synthetic Data, with no Noise and not Well-Resolved Domain

Example 27. To test the effect of noise on a signal with well-resolved $\overline{\mathcal{D}}_{fr}$ we use the same signal and define $\overline{\mathcal{D}}_{fr}$ so that the resolution of the frequency domain is 0.027. The signal is polluted with random Gaussian noise of mean zero and variance one (see Figure 7.5) and the relative error in extraction is less than 0.7% (see Figure 7.6). The InFs that were found are $\omega_1^* = 9.61$ and $\omega_2^* = 5.74$, which have relative error of less than 6%. This example shows the stability of the algorithm.

Example 28. (a) To synthetically examine the effect of adding an IMF to the main IMF, a simple sine wave of form $\sin\left(\frac{20\pi t}{T}\right)$ was added was added to the same signal used in previous examples. The extracted InFs are again accurate. In fact, if $\overline{\mathcal{D}}_{fr}$ has a resolution of 0.027, the error in extracted InFs is zero up to 3 digits of accuracy. With a $\overline{\mathcal{D}}_{fr}$ resolution of 0.25, the relative error is at most 8%. This small example shows that the algorithm works better than the bound provided by (7.3.4).

(b) To test whether a noisy observation with an added low frequency IMF would be still a tractable problem for the algorithm, we take the IMF from Example 25 and add $\sin\left(\frac{4\pi t}{T}\right) + \mathcal{N}(0,1)$ (see Figure 7.7). Here, $\mathcal{N}(0,1)$ is the white (Gaussian)

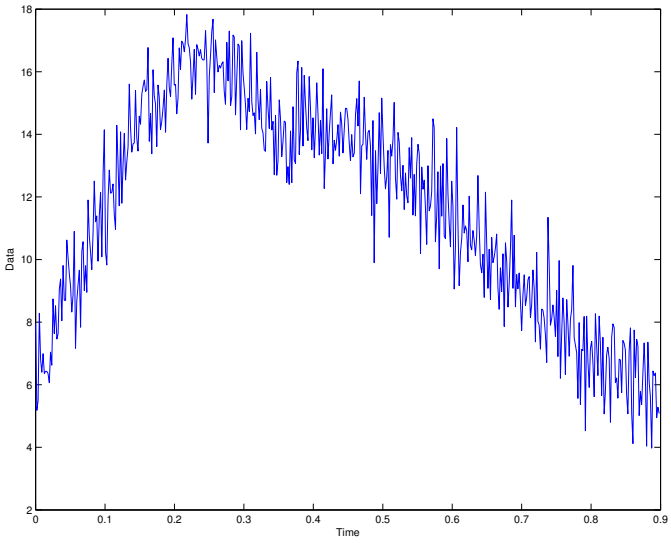

FIGURE 7.5. Original Noisy Data

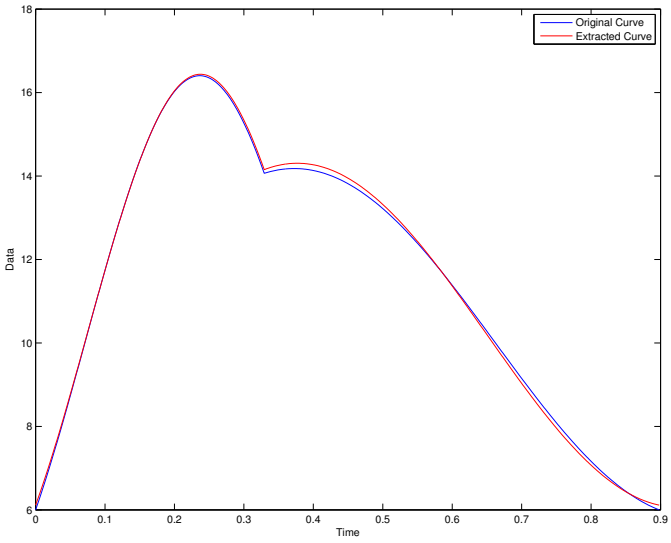

FIGURE 7.6. Extracted Curve vs the Original Curve in a Noisy environment

noise with mean zero and variance one. For a $\overline{\mathcal{D}}_{fr}$ resolution of 0.08, the extracted InFs are $\omega_1^* = 9.66$ and $\omega_2^* = 6.14$ the maximum relative error in extracted IFs is less than 14%. If a higher resolution is used, the results are much better. For example,

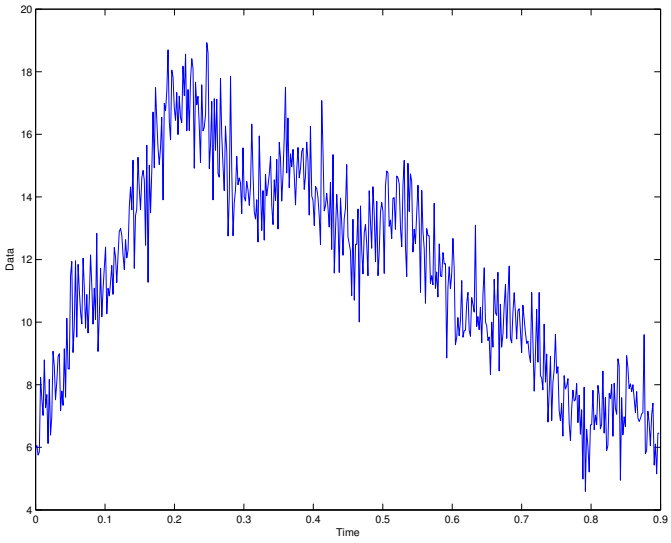

FIGURE 7.7. Synthetic trend plus IMF and Noise

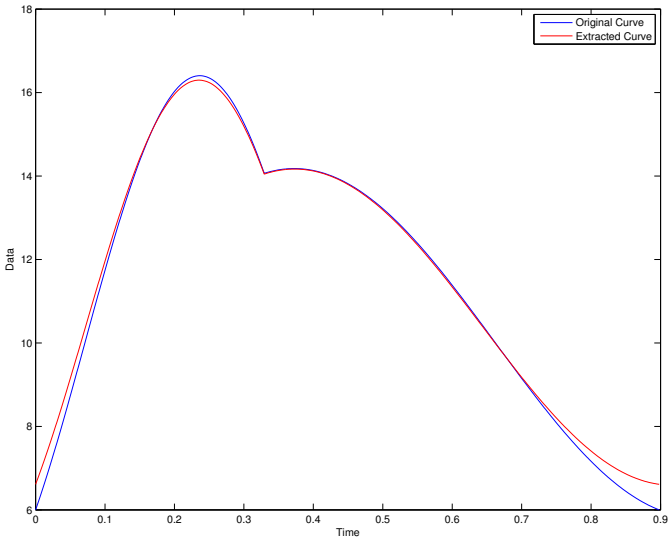

FIGURE 7.8. Extracted trend for the IMF-Noise case

for a resolution of 0.027, the extracted InFs are $\omega_1^* = 9.61$ and $\omega_2^* = 5.79$, having a maximum relative error of just 7% (see Figure 7.8). The curve extraction has a relative error of 2%.

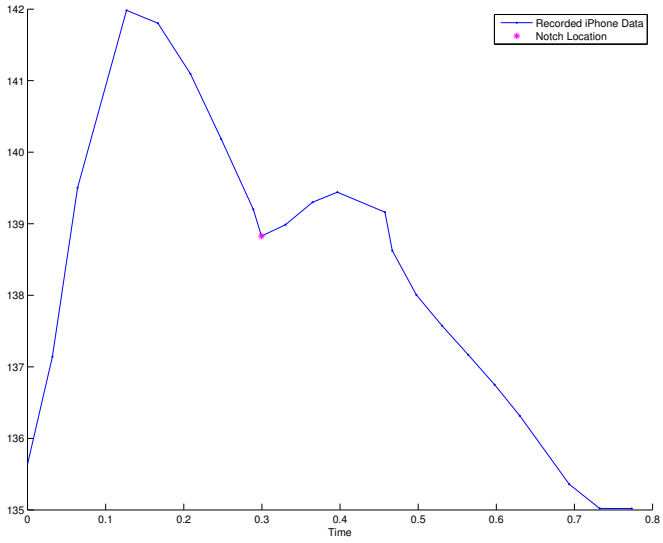

FIGURE 7.9. Recorded Data From iPhone: The data is recorded using iPhone camera.

7.5. Clinical Data

We have tested the performance of the InF algorithm on both invasive and non-invasive clinical data. The performance of our method has been beyond expectations. Data that had been collected from the neck artery of a subject solely using an iPhone camera (see Figure 7.9) was analyzed using the InF algorithm. Although the data has been collected from an iPhone, the algorithm faithfully extracted a pertinent IMF (see Figure 7.10).

We employed the InF algorithm on pressure wave signals collected from human beings (both invasively using a catheter and non-invasively using an iPhone camera) and dogs (invasively using a catheter). We found that the performance of the heart (the Ejection Fraction[3] (EF)) can be predicted using the normalized values of the InFs and the normalized value of $\bar{p}$. In order to be more specific, we need to define the following variables:

[3]Ejection Fraction is essentially a measure of the percentage of blood leaving the heart in each contraction. Ejection Fraction is a good measure of the performance of the heart. A very low Ejection Fraction corresponds to some cases of Heart Failure. A very traditional way of measuring the Ejection Fraction is through Echocardiography.

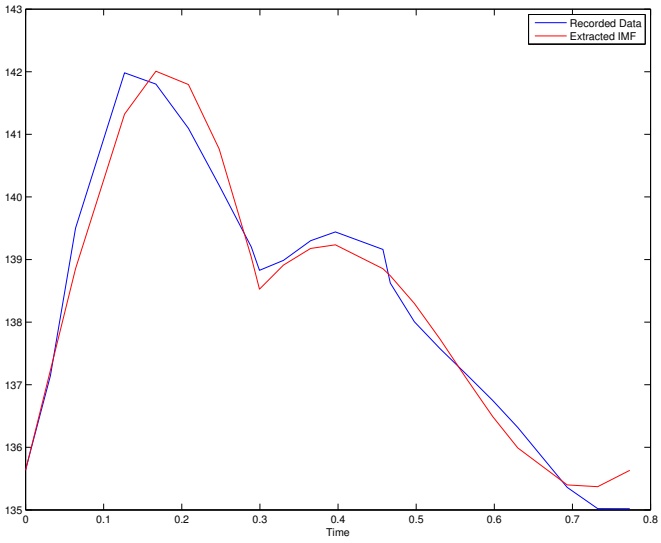

FIGURE 7.10. Extracted IMF from the Recorded iPhone Data

- $\omega_{1n} = \frac{\omega_1}{HR}$, where ω_1 is in bpm and HR is the heart rate in bpm.
- $\omega_{2n} = \frac{\omega_2}{HR}$, where ω_2 is in bpm and HR is the heart rate in bpm.
- $\overline{\omega_1} = \omega_1 T_0$, where ω_1 is in Hz and T_0 is in seconds.
- $\overline{\omega_2} = \omega_2 \left(T - T_0\right)$, where ω_2 is in Hz and $\left(T - T_0\right)$ is in seconds.
- $c_R = \frac{\bar{p} - \min(f(t))}{\max(f(t)) - \min(f(t))}$.

Irrespective of the data collection method (invasive or non-invasive), the EF could be predicted as follows:

(1) If $\omega_{1n} > 1.45$, then

$$EF = -1.1862 + 0.9661\omega_{1n}, \tag{7.5.1}$$

(2) If $\omega_{2n} > 1.5$, then

$$EF = -0.16907 + 0.83189\omega_{1n}, \tag{7.5.2}$$

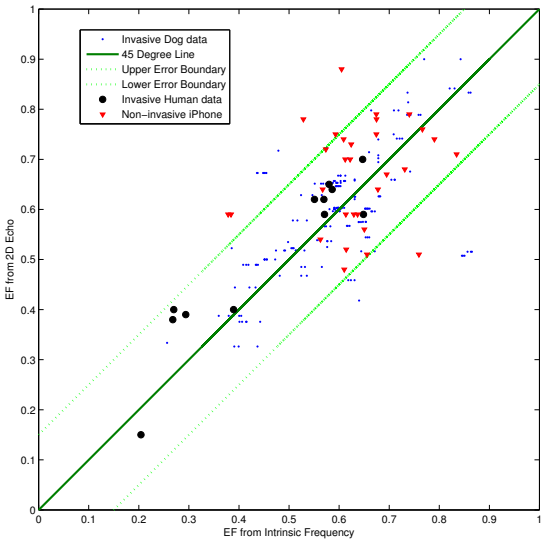

FIGURE 7.11. EF Comparison: The vertical axis shows the EF found by 2D echocardiography. The horizontal axis shows the EF calculated by the InF algorithm. The two dotted green lines show ±15% error offset from the expected 45 Degree line.

(3) Else

$$EF = 1 - c_R \frac{\overline{\omega_1}}{\overline{\omega_2}}. \tag{7.5.3}$$

The predictions of formulas (7.5.1), (7.5.2), and (7.5.3) are within the range of EF that is found by echocardiography (see Figure 7.11).

The performance of the InF algorithm clearly shows potential for use in health care systems. Hopefully by implementing this idea, many lives will be saved in the near future.

CHAPTER 8

Future Works: Unsolved Issues of IMF and IF Uniqueness

In this chapter, we will open the door to more fundamental issues regarding the definition of IMFs, IF and the issue of uniqueness. This chapter is solely an introduction to some future work in STFR methods. The sections in this chapter should not be considered as deep research results. However, they can be considered to be future research directions.

8.1. IMFs, Frequency, and Uniqueness

Here, we try to shed more light on the definitions of IMF, IF and uniqueness. Our tests showed that when IMFs are extracted from an IS, there are cases in which the IMF is uniquely extracted but the IF and the envelope are not unique.

8.1.1. Introduction. The idea of the IF was first proposed by Van der Pol in 1930 [51]. Later, Carson and Fry [11] were among the first to generalize the definition of the IF in the context of FM signals as

$$s\left(t\right) = e^{i\left(\omega_0 t + \lambda \int_0^t m(t)dt\right)},$$

and Van der Pol [52] validated this result with his own very similar definition. They defined the IF as the derivative of the phase function of an FM signal.

There have been attempts to define IF in the context of time-frequency methods. However, these methods have inherent shortcomings in data analysis, let alone helping to define IF. As mentioned in [17], for linear time-frequency methods such as wavelet transforms and windowed Fourier transforms the signal is analyzed by its inner product with an *a priori* dictionary of basis functions. The main problem with these methods, specifically the windowed Fourier transform, is the Heisenberg

uncertainty principle. Quadratic methods, on the other hand, compare a signal with a family of templates, which is useful unless a multi-scale signal is present, in which case interference becomes a problem. The main problem with spectrum analysis in general is that the resolution[1] is low [47]. For "non-stationary" signals, which have jumps or changes in behavior [17], both linear and quadratic methods fail. Hence, no definite definition of IF can be specified.

Boashash [6] was among the very first who raised the issue of the definition of the IF in depth. Boashash defines a "non-stationary" signal as one whose spectral characteristics, e.g. peaks, vary with time. In the stationary case, the definition of IF is clear. It can be related to the Fourier Transform. Boashash believes that the IF is an ambiguous object, stating "Since frequency usually defines the number of cycles or vibrations undergone during one unit of time by a body in periodic motion, there is an apparent paradox in associating the words instantaneous and frequency. For this reason the definition of IF is controversial, application-related, and empirically assessed." Citing Shekel [48], Boashash explains that the analytic signal approach cannot have a unique physical representation, although the complex representation could be unique. Boashash mentions that the problem of IF definition comes from the ambiguity in IF, envelope, and oscillation intuition. Boashash finally adopts the analytic signal approach to construct his theory. However, his approach breaks down for the case of multicomponent signals.

Daubechies [17] mentions that for an IMF $a(t)\cos\theta(t)$, the changes in time of the envelope $a(t)$ and frequency $\theta'(t)$ should be much slower than the change of $\theta(t)$ itself. This means that for $[t-\delta, t+\delta]$, where $\delta \approx \frac{2\pi}{\theta'(t)}$, the IMF is essentially a harmonic signal with amplitude $a(t)$ and frequency $\theta'(t)$. In order to define the IF and an IMF in practice, Daubechies uses Synchrosqueezing Wavelet Transforms. An in depth analysis of the method can be found in [50]. The wavelet transform W_s of

[1] i.e. the capability of distinguishing between more or less closely spaced neighboring spectral components

the signal s is defined by

$$W_s(a,b) = \int \frac{s(t)}{\sqrt{a}} \overline{\psi\left(\frac{t-b}{a}\right)} dt,$$

where ψ is a chosen wavelet. Later, in the same paper, a *candidate for IF* is defined as

$$\omega_s(a,b) = \frac{\frac{\partial}{\partial b} W_s(a,b)}{i W_s(a,b)}.$$

They show that this candidate would be a very good approximation of the true frequency $\phi'(t)$ if the following strict conditions are satisfied for $s(t) = A(t)\, e^{i\phi(t)}$:

$$A \in C^1(\mathbb{R}) \bigcap L_\infty, \phi \in C^2(\mathbb{R}),$$

$$\inf_{t\in\mathbb{R}} \phi'(t) > 0, \sup_{t\in\mathbb{R}} \phi'(t) < \infty,$$

$$|A'(t)|, |\phi''(t)| \leq \epsilon |\phi'(t)|, \forall t \in \mathbb{R}, \tag{8.1.1}$$

$$M'' := \sup_{t\in\mathbb{R}} |\phi''(t)| < \infty.$$

8.1.1 is the strictest condition. When considering an IS, other conditions are imposed on components. Nevertheless, this method does not address uniqueness concerns and these conditions cannot address intrawave signals. They also do not discuss cases with mode mixture. In that paper, a toy example is mentioned by which the problem of uniqueness is explained:

$$\begin{aligned} s(t) &= 0.25\cos((\Omega-\gamma)t) + 2.5\cos(\Omega t) + 0.25\cos((\Omega+\gamma)t) \\ &= \left(2 + \cos^2\left(\frac{\gamma}{2}t\right)\right)\cos(\Omega t). \end{aligned}$$

The definition of the frequency of an IMF was originally based on the Hilbert Transform (HT) of the signal [32]. Huang's definition was based on the number of extrema and zero crossings. However recently it was modified by a sifting process [33], and the normalization scheme, which is based on normalizing the extracted IMF

in an iterative way so that at the final step, the signal is in $[-1, 1]$. In order to find the IF, HT or a simple quadrature like $\sqrt{1-\cos^2\varphi(t)}$ is used.

The problem of non-uniqueness is mentioned in detail in [47] for two-tone signals. Under some conditions, the EMD method provides the physical or perceived extraction: "... EMD allows one to address in a fully data-driven way the question whether a given signal is better represented as a sum of two separate, unmodulated tones, or rather as a single, modulated waveform, with an answer that turns out to be in good agreement with intuition." Unfortunately, the study considers a very simplified example and is therefore not general.

Hou and Shi [25, 24, 26] use different approaches to define an IMF. Their initial approach [24] was to define an IMF $a(t)\cos\theta(t)$ that has a smooth envelope. The smoothness of the envelope was defined based on the Total Variation of the envelope. Later [25], they changed their approach to define an IMF $a(t)\cos\theta(t)$ as $a(t) \in V(\theta(t))\ s.t.\ a(t) > 0,\ \theta'(t) > 0$ where

$$V(\theta) = span\left\{1, \cos\left(\frac{\theta}{\lambda}\right), \sin\left(\frac{\theta}{\lambda}\right) \mid \lambda \geqslant 2\right\}. \tag{8.1.2}$$

This definition is the most rigorous of all. Nevertheless, it has the drawback that it is not in accord with the traditional definition of an IMF defined by Huang [32, 33]. For example, in $f = \left(1 + 0.9\cos\frac{t}{2}\right)\cos t$, there are three extrema rather than one around $t = 0.5$, between two zeros of the function. In other words, the IMF is bumpy.

Wu [54] has suggested that the definition of the IF should be extended to shape functions. He proposes to have $a(t)\,\mathcal{S}(\theta(t))$, where $\mathcal{S}$ is a shape function, which itself could be a summation of many $A(t)\cos\psi(t)$s. He puts the periodicity burden on the shape function; essentially, the phase function $\theta(t)$ is just a one-to-one map from the t-coordinate to the θ-coordinate. However, this approach has two problems here: the shape function must be known in advance, (but no method is provided for extracting it), and if the physical properties of the signal are embedded in $\psi(t)$, taking the whole

signal as $a(t)\mathcal{S}(\theta(t))$ would dilute the physical interpretation. Furthermore, lack of uniqueness is also a big problem for their method.

Cohen [14] defines the instantaneous frequency of a signal of form $f(t) = A(t)e^{i\varphi(t)}$ as the time derivative of $\varphi(t)$ but does not put any constraint on the envelope or phase function. This is a very similar definition to that in [11]. Here, IF is an empirical phenomenon experienced daily as changing colors or pitch. Based on his definitions, one interesting way of connecting time to frequency is to find the quantity

$$\langle t\varphi'(t)\rangle = \int t\varphi'(t)\left|f(t)\right|^2 dt.$$

Cohen emphasizes that the IF concept is not mature enough and tries to introduce ways by which an intuitive definition of IF can be found. The first way is to have an analytic signal (see [21] and [53]) using the Hilbert transformation of the signal

$$z(t) = s(t) + \frac{j}{\pi}\int_{PV}\frac{s(t')}{t-t'}dt'.$$

Here, $s(t)$ is the observed signal. Huang [32] mentions that this way of defining the instantaneous frequency would be flawed and counter intuitive if a mono component signal (IMF) is not used. The other way that Cohen defines the IF is based on the quadrature method. It is essentially not useful:

$$s_q(t) = A(t)e^{i\varphi(t)}\ for\ s(t) = A(t)\cos\varphi(t).$$

Cohen explains that the "exact mathematical description and understanding of the concept of changing frequency is far from obvious and it is fair to say that it is not a settled question." He himself explains the problems of the analytic signal definition in his work.

Lin et al. in [38] have introduced a new EMD method that is based on "moving averages". Unfortunately, this method does not necessarily extract the same IMFs that are extracted by the traditional EMD method. This can also be seen as a non-uniqueness issue. Nevertheless, the method has a strong mathematical background

(see [28, 38]). They also emphasize that the notion of IF is controversial in nature. However, they continue to use the analytic signal approach definition to define the IF.

There are also other definitions of IF that we just enumerate:

(1) Periodic Oscillatory motion: $\omega = \frac{1}{T}$, where T is the period of the motion [14].

(2) Dynamical System: $\omega(I) = \frac{\partial \mathcal{H}(I)}{\partial I}$, where $\mathcal{H}(\mathbf{p}, \mathbf{q})$ is the Hamiltonian of the system. For a conservative system, we have $\mathcal{H} = E$ the energy of the system, $I = \oint \mathbf{p} d\mathbf{q}/2\pi$ and

$$\frac{\partial \mathcal{H}}{\partial q_i} = -\dot{p}_i,$$
$$\frac{\partial \mathcal{H}}{\partial p_i} = \dot{q}_i.$$

This definition is good for low dimensional dynamical systems (see [36]).

(3) Wigner Ville Distribution (WVD) [14]:

$$V(t, \omega) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) \overline{x\left(t - \frac{\tau}{2}\right)} e^{-i\omega\tau} d\tau,$$

$$\langle \omega \rangle_t = \frac{\int_{-\infty}^{\infty} \omega V(t, \omega)\, d\omega}{\int_{-\infty}^{\infty} V(t, \omega)\, d\omega}.$$

Clearly, then the literature mentioned has no consensus on the definition of an IMF and IF.

8.1.2. A New Look into the Definitions of IMF and IFs. The classical definition of an IMF is based on the fact that the phase function $\theta(t)$ and the envelope function $a(t)$ are defined and then the IMF is defined as $a(t)\cos\theta(t)$. This approach encompass the idea that the envelope belongs to a certain dictionary [25, 24, 26]. The other approach, mainly depicted by Huang [32], considers an IMF to be an oscillatory function in which the difference between the number of extrema and zero crossings could be at most one, meaning that the defined IMF must have an $a(t)\cos\theta(t)$ representation.

Here, we show that both definitions result in the fundamental observation that an IMF does not have a unique representation. We will start with Huang's claim, and prove it. Then we will combine it with Hou's definition.

Definition 9. An n-zero-C^m IMF is a real function $f(t)$ in C^m, having n first-order zeros in $[t_0, t_{n-1}]$ and $(n-1)$ extrema. t_0 and t_{n-1} are two of the zeros. $m \in \mathbb{N}$ and $n \in \mathbb{N}$, $n \geq 3$.

8.1.2.1. C^0 *Phase functions.* Between any two zeros of an n-zero-C^m IMF, one can define a linear phase function $\theta(t)$. This linear phase function would help to construct a representation of the form $a(t)\sin\theta(t)$. However, as the phase function has a discontinuous first derivative on zeros of the IMF, the envelope function $a(t)$ would be undetermined on all t_i, $i = 0, \ldots, n$. This kind of representation can be useful in numerical cases.

8.1.2.2. C^1 *Phase functions.* In order to alleviate the discontinuity problem in the previous part, we propose the next lemma, by which we can find a smoother representation for an n-zero-C^m IMF.

Lemma 6. *There exists a* C^1 *map* $\theta(t) : [t_0 = 0, t_2 = 1] \to [0, \pi]$ *such that,* $\theta(0) = 0, \theta(1) = \pi$, $\frac{d\theta}{dt} > 0$ *and* $\frac{d\theta}{dt}|_{t_0=0} = M > 0$.

Proof. Take the function,

$$\theta(t) = \begin{cases} \pi t + kte^{\frac{1}{t^q-1}} & t \in [0,1) \\ \pi & t = 1 \end{cases}$$

for $k = e(M-\pi), 0 < q < 1$. The only free parameter so far is q, which would determine if $\frac{d\theta}{dt} > 0$. Taking the first derivative of $\theta(t)$ gives

$$\frac{d\theta}{dt}(t) = \begin{cases} \pi + k\mathrm{e}^{\frac{1}{t^q-1}} - \frac{kqt^q\mathrm{e}^{\frac{1}{t^q-1}}}{(t^q-1)^2} & t \in [0,1), \\ \pi & t = 1. \end{cases}$$

Now, we need to make sure that q guarantees a positive derivative. Taking the second derivative of the phase function and equating that to zero to find the location of the minimum of the first derivative would tell us that the relevant zero would occur at

$$t = \left(-\frac{\left(q + 2 - \sqrt{q\left(5q+4\right)} \right)}{2\left(q-1\right)} \right)^{1/q}.$$

The value of the first derivative of $\theta\left(t\right)$ at this point is $\pi + f\left(q\right)k$, where $f\left(q\right)$ is a monotone decreasing function for $0 < q < 1$. The bounds of this function are 0 and $-\frac{5}{e^3}$. Hence, for any value of k, one can always find an appropriate value of $q \in (0,1)$ such that $\frac{d\theta}{dt} > 0$. □

The next lemma will now pave the way to express smoother representation of an n-zero-C^m IMF.

Lemma 7. *There exists a C^1 map $\theta\left(t\right) : [t_0 = 0, t_{n-1} = 1] \rightarrow [0, \left(n-1\right)\pi]$ such that, $\theta\left(t_i\right) = i\pi$, for $i = 0, \ldots, n-1$, $\frac{d\theta}{dt} > 0$.*

Proof. For $t \in [t_0 = 0, t_1]$, set $\theta\left(t\right) = \pi\frac{t}{t_1}$. For $t \in [t_1, t_2]$, use Lemma 6 to construct $\theta\left(t\right)$ having $\dot{\theta}\left(t_1\right) = \frac{\pi}{t_1}$. In other words, in $t \in [t_1, t_2]$ take

$$\theta\left(t\right) = \begin{cases} \pi + \pi\left(\frac{t-t_1}{t_2-t_1}\right) + e\pi\left(\frac{t_2-t_1}{t_1} - 1\right)\left(\frac{t-t_1}{t_2-t_1}\right) e^{\frac{1}{\left(\frac{t-t_1}{t_2-t_1}\right)^q - 1}} & t \in [t_1, t_2), \\ 2\pi & t = t_2. \end{cases}$$

In general, for the i^{th} piece, $i = 2, \ldots, n-1$, take

$$\theta\left(t\right) = \begin{cases} \left(i-1\right)\pi + \pi\left(\frac{t-t_{i-1}}{t_i-t_{i-1}}\right) + e\pi\left(\frac{t_i-t_{i-1}}{t_{i-1}-t_{i-2}} - 1\right)\left(\frac{t-t_{i-1}}{t_i-t_{i-1}}\right) e^{\frac{1}{\left(\frac{t-t_{i-1}}{t_i-t_{i-1}}\right)^q - 1}} & t \in [t_{i-1}, t_i), \\ i\pi & t = t_i. \end{cases}$$

The way that we have defined the map satisfies $\theta\left(t\right) \in C^1$, where $\frac{d\theta}{dt} > 0$. □

Lemma 8. *If there is a map $\theta\left(t\right) \in C^1$, where $\frac{d\theta}{dt} > 0$, having fixed points $\left(t_0 = 0, 0\right)$, $\left(t_1, \pi\right), \left(t_2, 2\pi\right), \ldots, \left(t_{n-1} = 1, \left(n-1\right)\pi\right)$, then an n-zero-C^1 IMF $f\left(t\right)$ on $[t_0 = 0, t_{n-1} = 1]$ has a representation of $a\left(t\right)\sin\theta\left(t\right)$ form, for $\left(a\left(t\right) > 0\right) \in C$.*

PROOF. Without loss of generality, assume that between the first two zeros of $f(t)$, the IMF is positive. Take $a(t) = \frac{f(t)}{\sin\theta(t)}$ for all $t \in [0,1]$ except the nodes at t_i for $i = 0, \ldots, n-1$. For those nodes, one should take $a(t_i) = \frac{f'(t_i)}{-\theta'(t_i)}$, then $a(t) \in C$. The positivity of $a(t)$ follows from the way it is defined. □

These lemmas show that an n-zero-C^1 IMF, whether defined in Hou's or Huang's way, has a representation of $a(t)\sin\theta(t)$. The following Lemma and Theorem show that uniqueness does not exist.

Lemma 9. *If there exists a strictly increasing map $\theta(t)$ on $t \in [0,1]$, having fixed points $0, \pi, 2\pi, \ldots, n\pi$, then there are infinitely many strictly increasing maps having the same fixed points.*

PROOF. Since the map $t \to \theta(t)$ is strictly increasing, then it is also invertible on the same domain. Take the map $\chi(\theta) = \theta + c\sin\theta$. This map has the same fixed points; if $|c| < 1$, then χ would be strictly increasing in θ and consequently in t. $t \to \theta(t)$ is invertible, and $\theta \to \chi(\theta)$ is also invertible, hence $t \to \chi(t)$ is invertible. Since there are infinitely many values satisfying $|c| < 1$, the proof is complete. □

Theorem 15. *An n-zero-C^1 IMF has infinitely many $a(t)\sin\psi(t)$ representations.*

PROOF. Using Lemmas 8 and 7, one can construct one $a(t)\sin\theta(t)$. Using Lemma 9 on $\theta(t) \in C^1$, where $\frac{d\theta}{dt} > 0$, one can find $\psi(t) \in C^1$, where $\frac{d\psi}{dt} > 0$. Again using Lemma 8 one can find another representation for the same IMF, $A(t)\sin\psi(t)$. □

8.1.2.3. *C^∞ Phase functions.* In this subsection, we present very smooth phase functions, which helps to produce an $a(t)\sin\theta(t)$ representation for an n-zero-C^m IMF that has a C^m envelope $a(t)$. We first construct piecewise linear phase functions and then connect them using mollifiers in a way that the resulting phase function is a C^∞ smooth function.

Definition 10. A compact support mollifier is a C^∞ smooth function that has a compact support in $\mathbb{R}$.

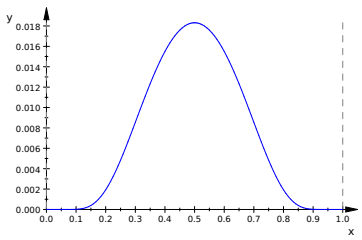

FIGURE 8.1. C^∞ Compact Support Mollifier

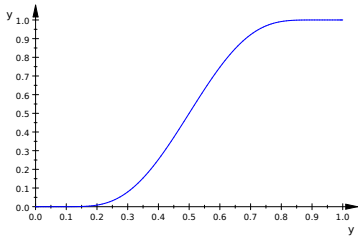

FIGURE 8.2. C^∞ Connector Mollifier

An example of a mollifier is

$$\chi(x) = \begin{cases} e^{\frac{-1}{x}} & x > 0, \\ 0 & x \leq 0. \end{cases}$$

This function is a C^∞ smooth function. One can easily convert this into a compact support mollifier by introducing the new function

$$\mu(x) = \begin{cases} \chi(x)\chi(1-x) & 0 < x < 1, \\ 0 & otherwise. \end{cases}$$

This function is symmetric (see Figure 8.1). Now, we need a mollifier to connect the two frequencies:

Definition 11. We define a connector mollifier (see Figure 8.2) as follows:

$$\eta(x) = \frac{\int_0^x \mu(\xi)\, d\xi}{\int_0^1 \mu(\xi)\, d\xi}.$$

Lemma 10. *For a connector mollifier defined in Definition 11, we have*

$$\int_0^1 \eta(x)\,dx = \frac{1}{2}.$$

PROOF. We have

$$\begin{aligned}
\int_0^1 \eta(x)\,dx &= \int_0^1 \frac{\int_0^x \mu(\xi)d\xi}{\int_0^1 \mu(\xi)d\xi}dx \\
&= \frac{1}{\int_0^1 \mu(\xi)d\xi}\int_0^1\int_0^x \frac{e^{\frac{1}{y-1}}}{e^{\frac{1}{y}}}dydx \\
&= \frac{1}{\int_0^1 \mu(\xi)d\xi}\int_0^1\int_y^1 e^{\frac{1}{y(y-1)}}dydx \\
&= \frac{1}{\int_0^1 \mu(\xi)d\xi}\int_0^1 (1-y)\,e^{\frac{1}{y(y-1)}}dy \\
&= \frac{\int_0^1 (1-y)e^{\frac{1}{y(y-1)}}dy}{\int_0^1 e^{\frac{1}{y(y-1)}}dy} \\
&= 1 - \frac{\int_0^1 ye^{\frac{1}{y(y-1)}}dy}{\int_0^1 e^{\frac{1}{y(y-1)}}dy} = 1-\tfrac{1}{2} = \tfrac{1}{2}.
\end{aligned}$$

Remember that $\frac{\int_0^1 ye^{\frac{1}{y(y-1)}}dy}{\int_0^1 e^{\frac{1}{y(y-1)}}dy} = \frac{1}{2}$ comes from the symmetric behavior of $e^{\frac{1}{y(y-1)}}$ around $y = 0.5$. □

The connector mollifier defined in Definition 11 can connect two discrete values a, b on an interval $(\varepsilon_1, \varepsilon_2)$ by

$$\eta_{\varepsilon_1}^{\varepsilon_2}(x; a, b) = (b-a)\,\eta\left(\frac{x-\varepsilon_1}{\varepsilon_2-\varepsilon_1}\right) + a.$$

We will use this later to construct the C^∞ phase function.

Lemma 11. *There exists $0 < \varepsilon_1 < \varepsilon_2 < 1$ and $f_{\varepsilon_1} > 0$, such that*

$$\int_0^{\varepsilon_1} \eta_0^{\varepsilon_1}(x; a, f_{\varepsilon_1})\,dx + \int_{\varepsilon_1}^{\varepsilon_2} \eta_{\varepsilon_1}^{\varepsilon_2}(x; f_{\varepsilon_1}, b)\,dx = b\varepsilon_2, \tag{8.1.3}$$

for any $a > 0$, $b > 0$ and $a \neq b$.

PROOF. Case 1 ($b > a$):

For sufficiently small ε_1 ($0 < \varepsilon_1 \ll 1$) one can always find $f_{\varepsilon_1} > 0$ satisfying the inequality

$$b + (b-a)\,\varepsilon_1 < f_{\varepsilon_1} < b + (b-a). \tag{8.1.4}$$

Now set $\varepsilon_2 = \left(\frac{b-a}{f_{\varepsilon_1}-b}\right)\varepsilon_1$. This condition, based on inequality (8.1.4), will satisfy the condition $0 < \varepsilon_1 < \varepsilon_2 < 1$:

i) $f_{\varepsilon_1} < b + (b-a) \Rightarrow f_{\varepsilon_1} - b < (b-a) \Rightarrow 1 < \frac{(b-a)}{f_{\varepsilon_1}-b} \Rightarrow \varepsilon_1 < \frac{(b-a)}{f_{\varepsilon_1}-b}\varepsilon_1 \Rightarrow \varepsilon_1 < \varepsilon_2$.

ii) $b + (b-a)\,\varepsilon_1 < f_{\varepsilon_1} \Rightarrow (b-a)\,\varepsilon_1 < f_{\varepsilon_1} - b \Rightarrow \frac{(b-a)}{f_{\varepsilon_1}-b}\varepsilon_1 < 1 \Rightarrow \varepsilon_2 < 1$.

Using this value of ε_2 we will prove that Equation (8.1.3) is satisfied:

$$\begin{aligned}\int_0^{\varepsilon_1} \eta_0^{\varepsilon_1}(x; a, f_{\varepsilon_1})\,dx + \int_{\varepsilon_1}^{\varepsilon_2} \eta_{\varepsilon_1}^{\varepsilon_2}(x; f_{\varepsilon_1}, b)\,dx &= \int_0^{\varepsilon_1} \left((f_{\varepsilon_1}-a)\,\eta\left(\tfrac{x}{\varepsilon_1}\right) + a\right)dx \\ &+ \int_{\varepsilon_1}^{\varepsilon_2} \left((b-f_{\varepsilon_1})\,\eta\left(\tfrac{x-\varepsilon_1}{\varepsilon_2-\varepsilon_1}\right) + f_{\varepsilon_1}\right)dx \\ &= \varepsilon_1 \int_0^1 \left((f_{\varepsilon_1}-a)\,\eta(x) + a\right)dx \\ &+ (\varepsilon_2-\varepsilon_1)\int_0^1 (b-f_{\varepsilon_1})\,\eta(x)\,dx \\ &+ (\varepsilon_2-\varepsilon_1)\int_0^1 f_{\varepsilon_1}dx\end{aligned}$$

Now, using the fact that $\int_0^1 \eta(x)\,dx = \frac{1}{2}$, we have

$$\begin{aligned}\int_0^{\varepsilon_1} \eta_0^{\varepsilon_1}(x; a, f_{\varepsilon_1})\,dx + \int_{\varepsilon_1}^{\varepsilon_2} \eta_{\varepsilon_1}^{\varepsilon_2}(x; f_{\varepsilon_1}, b)\,dx &= \varepsilon_1\tfrac{a-b}{2} + \varepsilon_2\tfrac{b+f_{\varepsilon_1}}{2} \\ &= \varepsilon_2\left(\left(\tfrac{b-f_{\varepsilon_1}}{a-b}\right)\tfrac{a-b}{2} + \tfrac{b+f_{\varepsilon_1}}{2}\right) \\ &= b\varepsilon_2.\end{aligned}$$

Case 2 ($b < a$):

For sufficiently small ε_1 ($0 < \varepsilon_1 \ll 1$) one can always find $f_{\varepsilon_1} > 0$ satisfying the inequality

$$b - (a-b) < f_{\varepsilon_1} < b - (a-b)\,\varepsilon_1. \tag{8.1.5}$$

Now set $\varepsilon_2 = \left(\frac{b-a}{f_{\varepsilon_1}-b}\right)\varepsilon_1$. This condition, based on inequality (8.1.5), will satisfy the condition $0 < \varepsilon_1 < \varepsilon_2 < 1$:

i) $f_{\varepsilon_1} < b - (a-b)\,\varepsilon_1 \Rightarrow (a-b)\,\varepsilon_1 < b - f_{\varepsilon_1} \Rightarrow \frac{(a-b)}{b-f_{\varepsilon_1}}\varepsilon_1 < 1 \Rightarrow \varepsilon_2 < 1$.

ii) $b - (a-b) < f_{\varepsilon_1} \Rightarrow b - f_{\varepsilon_1} < a - b \Rightarrow \frac{(a-b)}{b-f_{\varepsilon_1}}\varepsilon_1 > \varepsilon_1 \Rightarrow \varepsilon_2 > \varepsilon_1$.

Again, using this value of ε_2 we will prove that Equation (8.1.3) is satisfied. □

Lemma 11 will help us to match together the piecewise continuous instantaneous frequencies and consequently make a C^∞ phase function. However, before doing so, we need more prerequisites.

Lemma 12. *If the function $f(\theta)$ is C_θ^m and $\theta(t)$ is C_t^m, then $f(\theta(t))$ is C_t^m, for $m \geq 0$.*

PROOF. For $m = 0$ the proof is trivial, and for $m > 0$ one can use the Chain Rule. □

Lemma 13. *If the function $f(u)$ is C_u^m and $\theta(u)$ is C_u^m and $\frac{d\theta}{du} > 0$, then $f(\theta)$ is C_θ^m, for $m \geq 1$.*

PROOF. We show the proof for one step, which can be extended by induction. Define $y = f(u)$ and $x = \theta(u)$. As a result, $y = f(\theta^{-1}(x))$. Now apply the Chain Rule:

$$\begin{aligned} \tfrac{dy}{dx}|_{x=p} &= \tfrac{dy}{du}|_{u=\theta^{-1}(p)} \tfrac{du}{dx}|_{x=p} \\ &= f'(\theta^{-1}(p)) \tfrac{du}{dx}|_{x=p} \\ &= f'(\theta^{-1}(p)) \tfrac{1}{\theta'(\theta^{-1}(p))} \end{aligned}$$

In short, $\frac{dy}{dx}(x) = \left(\frac{f'}{\theta'}\right)(\theta^{-1}(x))$. This procedure can be continued for higher derivatives as well. As $\theta' > 0$, the conclusion follows that $y = f(\theta)$ is in C_θ^m. □

Lemma 14. *There exists a non-unique C^∞ map $\theta(t) : [t_0 = 0, t_{n-1} = 1] \to [0, (n-1)\pi]$ such that, $\theta(t_i) = i\pi$, for $i = 0, \ldots, n-1$, $\frac{d\theta}{dt} > 0$.*

PROOF. At the first step, fit a piecewise map between the nodes $\theta(t_i) = i\pi$. This piecewise map has constant frequencies in (t_i, t_{i+1}) namely ω_i for $i = 0, \ldots, n-2$. In order to make this piecewise constant frequency map into a C^∞ map, use Lemma 11: For $[t_0, t_1]$, $\omega_0 = \frac{\pi}{t_1 - t_0}$. For $[t_1, t_1 + \varepsilon]$, where $t_1 + \varepsilon < t_2$, use Lemma 11 with $a = \omega_0$ and $b = \omega_1 = \frac{\pi}{t_2 - t_1}$; hence, for $[t_1 + \varepsilon, t_2]$, we set $\omega_1 = \frac{\pi}{t_2 - t_1}$. Doing this in $[t_0 = 0, t_{n-1} = 1]$ one can construct a C^∞ instantaneous frequency$\omega(t)$. Integrating such a function would produce our desired map $\theta(t) = \int_0^t \omega(\xi)\, d\xi$.

There are three subtle points that we need to address to finalize the proof:

1- When using Lemma 11, we implicitly use a linear map between $[t_i, t_{i+1}]$ and $[0, 1]$. This map preserves the C^∞ properties of $\omega(t)$ due to the way we have defined $\eta(x)$ in Lemma 11.

2- The non-uniqueness property of the map comes from the free parameters in Lemma 11.

3- The way we have defined $\eta(x)$ and the essential integral in Definition 11 and Lemma 11 would guarantee $\pi = \int_{t_i}^{t_{i+1}} \omega(\xi)\, d\xi$. □

Theorem 16. *If there is a map $\theta(t) \in C^\infty$, where $\frac{d\theta}{dt} > 0$, having fixed points $(t_0 = 0, 0), (t_1, \pi), (t_2, 2\pi), \ldots, (t_{n-1} = 1, (n-1)\pi)$, then an n-zero-$C^m$ IMF $f(t)$ on $[t_0 = 0, t_{n-1} = 1]$, for $m \geq 2$, has a representation of $a(t)\sin\theta(t)$ form, for $a(t) > 0$ at least in C^{m-2}.*

Proof. Define $a(t) = \frac{f(t)}{\sin(\theta(t))}$. If we prove that $a(\theta) = \frac{f(\theta)}{\sin\theta}$ is in C_θ^{m-2}, using Lemmas 12 and 13 we can say that $a(t)$ at least in C_t^{m-2}. Without loss of generality, assume that between the first two zeros of $f(t)$, the IMF is positive.

Take $a(\theta) = \frac{f(\theta)}{\sin\theta}$. For all $\theta \in [0, (n-1)\pi]$, $a \in C_\theta^m$ except the nodes at $\theta_i = i\pi$ for $i = 0, \ldots, n-1$. At those nodes, we define $a^{(q)}(\theta_i) = \lim\limits_{\theta\to\theta_i} \frac{d^q}{d\theta^q}\left(\frac{\frac{f(\theta)}{\theta-\theta_i}}{\frac{\sin(\theta-\theta_i)}{\theta-\theta_i}}\right)$. The function $\frac{\sin\theta}{\theta}$ is in C_θ^∞. The function $\frac{f(\theta)}{\theta}$ will have a representation like

$$\frac{1}{\theta-\theta_i}\left(\sum_{j=1}^{m-1}\frac{1}{j!}\frac{d^j f}{d\theta^j}(\theta_i)(\theta-\theta_i)^j + \frac{1}{m!}\frac{d^m f}{d\theta^m}(\xi)(\theta-\theta_i)^m\right),$$

based on Taylor's Theorem, for some ξ. In fact, $\lim\limits_{\theta\to\theta_i} \frac{d^q}{d\theta^q}\left(\frac{\frac{f(\theta)}{\theta-\theta_i}}{\frac{\sin(\theta-\theta_i)}{\theta-\theta_i}}\right)$ exists up to order $q = m-2$ without depending on $\frac{d^m f}{d\theta^m}(\xi)$. This proves that $a \in C_\theta^{m-2}$ at least. The positivity of a follows from the way it is defined. This concludes the proof. □

Corollary 1. *An n-zero-C^m IMF, for $m \geq 2$, has infinitely many $a(t)\sin\psi(t)$ representations for $\psi(t) \in C^\infty$ and $a(t) \in C^{m-2}$.*

8.1.3. Best Representations. Since an IMF might have many representations with different regularities in phase and envelope, we must identify the best representation.

Assume that there is a representation of an IMF of the form $a(t)\cos\theta(t)$ for some $\theta(t) \in C^\infty$ and $a(t) \in C^k$, $k \geq 0$. From visual point of view, the best representation

is the one that has the smoothest envelope compared to the oscillatory part $\cos\theta(t)$. This smoothness must be enforced in the θ-space and not the t-space. So, in the continuous case, the best representation of such an IMF is the solution of the following optimization problem.

$$\begin{array}{ll} \underset{\omega(t)}{minimize} & \|a\|_{W_\theta^{k,p}} \\ subject\,to & \int_{t_i}^{t_{i+1}} \omega(t)\,dt = \pi,\ i = 0\ldots n-2, \\ & \theta(t) \in C^\infty, \\ & \omega(t) = \frac{d\theta}{dt}(t). \end{array} \tag{8.1.6}$$

Here, we have abused the notation a bit. In fact, $W_\theta^{k,p}$ is the symbol of Sobolev Spaces [19] with a wider range of functions that have weak derivatives, but here we use $\|a\|_{W_\theta^{k,p}}$ to refer to

$$\|a\|_{W_\theta^{k,p}} = \left(\sum_{j=0}^{k} \left\| \frac{d^j a}{d\theta^j}(\theta) \right\|_{L_\theta^p}^p \right)^{\frac{1}{p}}.$$

This problem is not necessarily convex and in fact it is not always a tractable problem. However, we can have a relaxed discrete version of it to help us chose between two different representations of the same IMF. We propose that if an IMF has two different representations, we pick the one that has the following norm as the minimum:

$$\|a\|_{W_\theta^{1,2}} = \left(\|a(\theta)\|_{L_\theta^2}^2 + \left\| \frac{da}{d\theta}(\theta) \right\|_{L_\theta^2}^2 \right)^{\frac{1}{2}}.$$

CHAPTER 9

Concluding Remarks

In this thesis, we have further developed the STFR methods initially developed by Hou and Shi [25, 24, 26]. The main areas of the concentration for this development have been

- The development the non-periodic STFR method
- Modifying the STFR method to extract intrawave signals, sharp signals, and rare events
- Analysis of the convergence of the periodic STFR method with a variable filter length
- Applications of the STFR method in dynamical systems
- Application of the modified version of the STFR method, namely the Intrinsic Frequency (InF) method in cardiovascular health analysis.

Since the periodic STFR method was not able to handle non-periodic signals, we introduced an l_1-regularized-l_2 sparse optimizer that could handle this problem readily. Our results show that the method is not only able to handle non-periodic data, but is also fairly stable to noise and is a good trend detector. When compared with other STFR methods, the only shortcoming of the Non-Periodic STFR method is the speed of the algorithm.

The implementation of non-periodic STFR with the adjustment of the parameters of the algorithm so that it is capable of handling intrawave signals paved the way to properly analyze such signals. After minor adjustments, many signals which were hard to analyze using any other adaptive method can now be easily extracted and analyzed. The EMD method can extract one intrawave IMF in the absence of noise. However, neither EMD nor EEMD can extract even one intrawave IMF in the presence

of noise. Even intrawave signals that contain a mode mixture can be easily tackled by extracting the mode mixed signals simultaneously using the non-periodic STFR method with adjusted the filter parameter. Not only is the algorithm that we use to extract the intrawave IMFs stable, it is also robust to noise perturbation. The EMD/EEMD methods fail to extract the two IMFs properly. In fact all other adaptive methods fail to extract one IMF with intrawave frequency modulation in the presence of noise, let alone two IMFs with intrawave mode mixture characteristics.

In this thesis, we also showed the convergence of the periodic STFR method with a variable filter length. This convergence analysis, based on a contraction and harmonic analysis, is essentially viable when one is trying to extract an ordinary IMF or an intrawave IMF. We proved that for any signal, whether intrawave or not, increasing the filter span reduces the extraction error. We showed that the STFR method will converge to an IMF that is close to one of the IMF representations, but with an error associated with the width (span) of the filter.

By applying our method in dynamical systems, we showed that many IMFs are simply byproducts of the dynamical systems. We presented a method to quantify the nonlinearity of the IMFs given by the STFR method. The main idea is to establish a connection between the IMFs and classical second order differential equations. This observation enables us to discover the second order ODE of a specified IMF. Furthermore, this approach enables us to define the degree of nonlinearity of an IMF.

The InF method was also used to measure heart performance in humans and dogs, raising the possibility of using the InF to diagnose cardiovascular diseases based on the analysis of the aortic pressure wave. We proved the convergence of the InF method, and hope that this modified version of STFR can be used in real life diagnosis of cardiovascular diseases.

All in all, in this work, we successfully expanded the realm of the STFR method both in theory and application.

Bibliography

[1] FV Atkinson. *Discrete and Continuous Boundary Problems. 1964.* Academic Press, New York.

[2] J.H. Barrett. Oscillation theory of ordinary linear differential equations. *Advances in Mathematics*, 3(4):415–509, 1969.

[3] C.M. Bender and S.A. Orszag. *Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory*, volume 1. Springer, 1999.

[4] D.P. Bertsekas. Nonlinear programming. 1999.

[5] C. Blatter. *Wavelets: a primer.* Universities Press, 2003.

[6] B. Boashash. Estimating and interpreting the instantaneous frequency of a signal. i. fundamentals. *Proceedings of the IEEE*, 80(4):520–538, 1992.

[7] A. Boggess. *A first course in wavelets with Fourier analysis.* Wiley. com, 2009.

[8] S. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge university press, 2004.

[9] E. Brevdo. Synchrosqueezing wavelet transform, https://web.math.princeton.edu/ ebrevdo/synsq/, 2012.

[10] E.J. Candes and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *Information Theory, IEEE Transactions on*, 52(12):5406–5425, 2006.

[11] J.R. Carson and T.C. Fry. Variable frequency electric circuit theory with application to the theory of frequency modulation. *Bell Syst. Tech. J*, 16(4):513–540, 1937.

[12] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *SIAM journal on scientific computing*, 20(1):33–61, 1998.

[13] E.A. Coddington and N. Levinson. *Theory of ordinary differential equations*. Tata McGraw-Hill Education, 1955.

[14] L. Cohen. *Time-frequency analysis*, volume 778. Prentice Hall PTR New Jersey, 1995.

[15] J.W Cooley and J.W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301, 1965.

[16] I Daubechies. The wavelet transform, time-frequency localization and signal analysis. *Information Theory, IEEE Transactions on*, 36(5):961–1005, 1990.

[17] I. Daubechies, J. Lu, and H. Wu. Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool. *Applied and computational harmonic analysis*, 30(2):243–261, 2011.

[18] J.W. Demmel. *Applied numerical linear algebra*. Society for Industrial Mathematics, 1997.

[19] L.C. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics. American Mathematical Society, 2010.

[20] P. Flandrin, G. Rilling, and P. Goncalves. Empirical mode decomposition as a filter bank. *Signal Processing Letters, IEEE*, 11(2):112–114, 2004.

[21] D. Gabor. Theory of communication. part 1: The analysis of information. *Electrical Engineers-Part III: Radio and Communication Engineering, Journal of the Institution of*, 93(26):429–441, 1946.

[22] M. Grant, S. Boyd, and Y. Ye. Cvx: Matlab software for disciplined convex programming. *Online accessiable: http://stanford. edu/˜ boyd/cvx*, 2008.

[23] D. Guignard. Adaptive data analysis methods for nonlinear and nonstationary data. Master's thesis, Caltech, EPFL, 2012.

[24] T.Y. Hou and Z. Shi. Adaptive data analysis via sparse time-frequency representation. *Advances in Adaptive Data Analysis*, 3(01n02):1–28, 2011.

[25] T.Y. Hou and Z. Shi. Data-driven time-frequency analysis. *Applied and Computational Harmonic Analysis*, 2012.

[26] T.Y. Hou, Z. Shi, and P. Tavallali. Convergence of a data-driven time-frequency analysis method. *Applied and Computational Harmonic Analysis*, (0):–, 2014.

[27] B. Huang and A. Kunoth. An optimization based empirical mode decomposition scheme. *Journal of Computational and Applied Mathematics*, 2012.

[28] C. Huang, L. Yang, and Y. Wang. Convergence of a convolution-filtering-based algorithm for empirical mode decomposition. *Advances in Adaptive Data Analysis*, 1(04):561–571, 2009.

[29] N.E. Huang, M. Lo, Z. Wu, and X. Chen. Method for quantifying and modeling degree of nonlinearity, combined nonlinearity, and nonstationarity, March 28 2013. US Patent 20,130,080,378.

[30] N.E. Huang and S.S. Shen. *Hilbert-Huang transform and its applications*, volume 5. World Scientific Pub Co Inc, 2005.

[31] N.E. Huang, Z. Shen, and S.R. Long. A new view of nonlinear water waves: The hilbert spectrum. *Annual review of fluid mechanics*, 31(1):417–457, 1999.

[32] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971):903, 1998.

[33] N.E. Huang, Z. Wu, S.R. Long, K.C. Arnold, X. Chen, and K. Blank. On instantaneous frequency. *Advances in Adaptive Data Analysis*, 1(02):177–229, 2009.

[34] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. An interior-point method for large-scale l1-regularized least squares. *Selected Topics in Signal Processing, IEEE Journal of*, 1(4):606–617, 2007.

[35] T.G. Kolda, R.M. Lewis, and V. Torczon. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM review*, 45(3):385–482, 2003.

[36] L.L.D. Landau. *Course of Theoretical Physics.-1: Mechanics*. Pergamon Press, 1976.

[37] R.J. LeVeque. *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*. Siam, 2007.

[38] L. Lin, Y. Wang, and H. Zhou. Iterative filtering as an alternative algorithm for empirical mode decomposition. *Advances in Adaptive Data Analysis*, 1(04):543–560, 2009.

[39] D. Lloyd-Jones, R.J. Adams, T.M. Brown, M. Carnethon, S. Dai, G. De Simone, T.B. Ferguson, E. Ford, K. Furie, C. Gillespie, et al. Heart disease and stroke statistics 2010 update. *Circulation*, 121(7):e46–e215, 2010.

[40] W. Magnus and S. Winkler. *Hill's equation*. Dover publications, 2004.

[41] S.G. Mallat. *A wavelet tour of signal processing*. Academic Pr, 1999.

[42] S.G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *Signal Processing, IEEE Transactions on*, 41(12):3397–3415, 1993.

[43] N.W. McLachlan. *Ordinary non-linear differential equations in engineering and physical sciences*. Clarendon Press, 1958.

[44] N. Pahlevan, P. Tavallali, T.Y. Hou, and M. Gharib. Intrinsic frequency hemodynamic waveform analysis, December 21 2012. US Patent App. 13/725,039.

[45] N.M. Pahlevan and M. Gharib. Aortic wave dynamics and its influence on left ventricular workload. *PloS one*, 6(8):e23106, 2011.

[46] H. Prüfer. Neue herleitung der sturm-liouvilleschen reihenentwicklung stetiger funktionen. *Mathematische Annalen*, 95(1):499–518, 1926.

[47] G. Rilling and P. Flandrin. One or two frequencies? the empirical mode decomposition answers. *Signal Processing, IEEE Transactions on*, 56(1):85–95, 2008.

[48] J. Shekel. Instantaneous frequency. *Proc. IRE*, 41(548):426, 1953.

[49] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*, volume 140. American Mathematical Society, 2012.

[50] G. Thakur, E. Brevdo, N.S. Fučkar, and H. Wu. The synchrosqueezing algorithm for time-varying spectral analysis: robustness properties and new paleoclimate applications. *Signal Processing*, 2012.

[51] B. Van Der Pol. Frequency modulation. *Proceedings of the Institute of Radio Engineers*, 18(7):1194–1205, 1930.

[52] B. Van der Pol. The fundamental principles of frequency modulation. *Electrical Engineers-Part III: Radio and Communication Engineering, Journal of the Institution of*, 93(23):153–158, 1946.

[53] J.d. Ville. Théorie et applications de la notion de signal analytique. *Cables et transmission*, 2(1):61–74, 1948.

[54] H. Wu. Instantaneous frequency and wave shape functions (i). *Applied and Computational Harmonic Analysis*, 2012.

[55] Z. Wu and N.E. Huang. A study of the characteristics of white noise using the empirical mode decomposition method. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 460(2046):1597–1611, 2004.

[56] Z. Wu and N.E. Huang. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1(01):1–41, 2009.

[57] Z. Wu and N.E. Huang. On the filtering properties of the empirical mode decomposition. *Advances in Adaptive Data Analysis*, 2(04):397–414, 2010.

[58] Z. Wu, N.E. Huang, S.R. Long, and C.K. Peng. On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proceedings of the National Academy of Sciences*, 104(38):14889–14894, 2007.