Summer 8-10-2018

The Hilbert-Huang Transform: A Theoretical Framework and Applications to Leak Identification in Pressurized Space Modules

Kenneth R. Bundy
University of Maine, kenneth.bundy@maine.edu

THE HILBERT-HUANG TRANSFORM: A THEORETICAL FRAMEWORK AND APPLICATIONS TO LEAK IDENTIFICATION IN PRESSURIZED SPACE MODULES

By

Kenneth Richard Bundy

B.A., University of Maine, 2014

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

(in Mathematics)

The Graduate School

The University of Maine

August 2018

Advisory Committee:

Ali Abedi, Professor of Electrical and Computer Engineering, Co-Advisor

Andrew Knightly, Associate Professor of Mathematics, Co-Advisor

Nigel Pitt, Professor of Mathematics

THE HILBERT-HUANG TRANSFORM: A THEORETICAL FRAMEWORK AND APPLICATIONS TO LEAK IDENTIFICATION IN PRESSURIZED SPACE MODULES

By Kenneth Richard Bundy

Thesis Co-Advisors: Ali Abedi and Andrew Knightly

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Arts
(in Mathematics)
August 2018

Any manned space mission must provide breathable air to its crew. For this reason, air leaks in spacecraft pose a danger to the mission and any astronauts on board. The thesis examines the research in air leak detection, localization, and identification. Three methods for leak type identification are suggested. The methods are based on the Fourier Transform, Hilbert-Huang Transform, and Artificial Neural Networks, respectively.

In addition to the leak detection, open questions surrounding the Hilbert-Huang Transform are addressed. Specifically, it is shown that with an additional assumption, the transform's sifting process converges uniformly. From here, it is concluded that an appropriate stopping criterion yields a (non-unique) Intrinsic Mode Function (an oscillating function centered at zero), and that the sifting process always yields a decomposition of the input function into finitely many pieces. These results together provide more bricks in the transform's theoretical foundation.

ACKNOWLEDGEMENTS

I would like to thank Dr. Ali Abedi for his constant support, both financial and technical, through my graduate career. This would not have been possible without you.

I would like to thank Dr. Andrew Knightly for his many contributions, both in content and in funding, through this project. This would not have been possible without you.

I would like to thank Dr. Nigel Pitt for his numerous contributions to the work. His comments and guidance were invaluable. This would not have been possible without you.

I would like to thank Ayesha Maliwal for her constant emotional support. Thank you for standing by me when I am at my worst. This would not have been possible without you.

I would like to thank my parents for their support in all things, for as long as I can remember. This would not have been possible without you.

I would like to thank all of my friends who have supported me throughout this journey. This would not have been possible without you.

I would like to thank the Maine Space Grant Consortium and the National Aeronautics and Space Administration for funding parts of the research. This would not have been possible without you.

I would like to thank the people who often get missed, such as support staff and others who supported this work indirectly. It is easy to say that we have done things ourselves, but in reality, we all stand on the shoulders of invisible giants.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF THEOREMS

LIST OF DEFINITIONS

CHAPTER 1

INTRODUCTION

1.1 THE PROBLEM

Space travel poses certain inherent dangers. Due to the harsh environment, the spacecraft crew are dependent on life support systems for their survival. These systems must therefore have a high level of reliability. Among these systems is the maintenance of breathable air and cabin pressure. This system is at risk from air leaks; a means of detecting and localizing these leaks for repair is therefore necessary.

Air leaks in pressurized space vehicles can come from two main sources. The first is ordinary wear and tear. For example, a 2003 air leak was found in a section of rubber tubing that supplied air to the station; this was from wear and tear as astronauts held it to look out the window[1].

The second source of air leaks is Micrometeorites and Orbital Debris (MMOD). These are small objects that are orbiting earth; they can be natural or man-made, however most are from the latter category. Due to their high orbital velocities, they pose a significant threat to orbital vehicles. NASA tracks many pieces of orbital debris, but there are many that are too small to track, and may puncture the hull of a space vehicle [2].

The current leak detection systems rely on detecting a drop in cabin pressure. This process is slow and does nothing to determine the location of the air leak. On the ISS, modules may need to be closed in order to determine the location of the leak in this manner. This is certainly not an ideal system [1].

Other methods include the use of a hand held device for leak scanning [1]. These devices consist of an ultrasonic microphone, signal processing hardware, and a

headset. The input from the microphone is shifted into the audible range, which allows the operator to locate the leak by sweeping the station manually. This process is very time consuming and poses a significant drain on crew time and resources. For this reason, it has been deemed suboptimal by NASA [1].

1.2 PROPOSED SOLUTION

It was in developing methods for leak identification and localization that I first came across the Hilbert-Huang Transform. Since air leaks are comprised largely of turbulent flow, analyzing recordings of their acoustics can be difficult. While Fourier Transforms offered some promising results, the waveforms were very chaotic and we sought a better method. This lead to the Hilbert-Huang Transform, which is a signal processing methodology for poorly behaved signals.

While the Hilbert-Huang Transform is a relatively new methodology, and its theoretical framework in not yet established. This work addresses both the mathematical underpinnings of the Hilbert-Huang Transform, and the leak identification work that inspired them.

Leaks produce noise that is basically turbulent flow; it is a highly random pattern that is not necessarily 'stationary' in the sense of a stochastic process. For this reason, Fourier Transform and Wavelet based methods may not offer good results when used on audio recordings of air leaks. The Hilbert-Huang Transform is well suited to these kinds of signals.

However, there are some issues with the Hilbert Huang transform. For instance, there is no universally accepted theoretical framework for this transform. Furthermore, the transform requires that one choose upper and lower bounding functions, called envelopes, that are not uniquely defined. This leads to some debate about the best way to carry out the transform. Also, there has not yet been any proof that the iterative process defining the Hilbert-Huang Transform (HHT)

(a) Handheld Leak detection system.

(b) UMaine's WLD System, Dimensions 18cm by 13cm by 7cm.

Figure 1.1: A previously used leak detection system (a) and the University of Maine's Wireless Leak Detector (b) tested on the International Space Station in 2017.

converges to any limit, or that any given stopping condition will be met. In other words, there is no guarantee of getting a result from the HHT. In practice this is not an issue as there is strong empirical evidence of its effectiveness.

1.3 SUMMARY OF CONTRIBUTIONS

The work on leak detection has been published in three works at the Wireless Sensors in Extreme Environments Conference in 2015-2017.

The first presentation was a poster explaining the preliminary approach and results. It discussed the data collection methods and proposed approach to the problem. This work was presented before the data set was completely collected, so no final results were presented [3].

The second was a paper explaining the results obtained so far from the Fourier Transform and Hilbert-Huang Transform based approaches. It discussed the data collection methods in more depth and introduced the Fourier Transform and Hilbert-Huang Transform methods [4].

The final presentation was a poster outlining the Artificial Neural Network approach, and the results obtained [5].

Some new results on the Hilbert-Huang Transform are presented here as well. These results have not yet been published, but are presented here in detail. First, the process defining the transform converges with some additional assumptions. Second, the result of the process is the type of function that the inventors of the function specified. Third, the transform always yields a finite decomposition of the input function.

1.4 THESIS SUMMARY

Chapter 2 discusses some important background information on air leaks in pressurized space modules. Then the following sections of the same chapter provide some background on the three methods proposed in this thesis to solve the leak identification problem.

Chapter 3 presents mathematical results on the Hilbert-Huang Transform. While this is not directly related to leak identification, it provides new material including proofs of convergence and finiteness necessary to guarantee the good behavior of the transform.

Chapter 4 explains the experimental procedures and the results from the leak identification problem, and the results of the proposed solutions.

Chapter 5 concludes the work. It also suggests further work to be done in these areas, both the Hilbert-Huang Transform and the leak identification problem.

CHAPTER 2

BACKGROUND

2.1 INTRODUCTION

This section first presents the air leak problem in pressurized space modules and its history. It begins with some discussion of air leaks in spacecraft and past research addressing this concern. Several solutions to the problem are proposed. A Fourier Transform based method is developed first, then a Hilbert-Huang Transform based method, and finally an Artificial Neural Network based approach. We explain each of these in detail below.

2.2 HISTORY

One of the first and most tragic instances of an air leak in a space vehicle was on the USSR's Soyuz-11 craft in 1971. Three cosmonauts died as a result of their re-entry module rapidly losing air pressure after a hatch seal failed [6].

Further issues with air leaks happened in 1991 after an astronaut's suit was punctured during a space walk. Fortunately, the hole was small, and within the tolerance of the system. There was some bruising of the astronaut's skin near the leak opening, as the skin was drawn towards the opening [7].

In 2004, and air leak aboard the International Space Station (ISS) nearly lead to closing a section of the station. After multiple searches of the affected module over three days, the air leak was found in a rubber hose. The astronauts had been pulling to to look out the window without drifting away in zero gravity [8].

In 2009, there was a leak of coolant (ammonia) from the external cooling system. This was repaired by a space walk. The leak is thought to have been caused by an impact by a piece of orbital debris, but this was never verified [9].

Another happened in 2016; a spacewalk at the ISS was aborted after an astronaut's helmet began to leak air into space. Fortunately, the astronaut was able to re-enter the space station before permanent injury.

2.3 PROPERTIES OF LEAKS

Significant work has already been done to understand the properties of air leaks at the University of Iowa and NASA [10, 11]. NASA performed a thorough search of existing technologies and found the existing systems did not meet their requirements [1]. Furthermore, none of the leak detection schemes surveyed determined the leaking material. In addition to these commercial methods, other leak detection systems have been proposed. These include structure-borne acoustic leak localization [11], Laguerre Spectral Analysis [12], and Hilbert Spectral Analysis [13].

Spacecraft air leaks differ significantly from those in other applications. Air leaks in high pressure air lines, for example those used in pneumatic tools, produce sound from the jet of escaping air. The sounds are audible since the observer is on the low pressure side of the aperture. Systems on spacecraft must function on the high pressure side of the leak, as the vacuum or near-vacuum of space does not allow for the transmission of sound. While spacecraft leaks produce these sounds, they are not able to be heard inside the spacecraft. As noted in [11], the air leaving the spacecraft carries the sound with it, since the air escapes faster than the speed of sound. The leak acoustics which can be heard within the spacecraft come from the materials surrounding the leak aperture [11], as shown in Figure 2.1.

The frequency of spacecraft air leak acoustics have been shown to fall between 35 and 50 kHz [10]. This frequency band is the primary one for concern in the research here. The resonant frequency of the material in question can be used in structure-borne acoustic wave leak detection, as in [11], however this will not be the focus of this work.

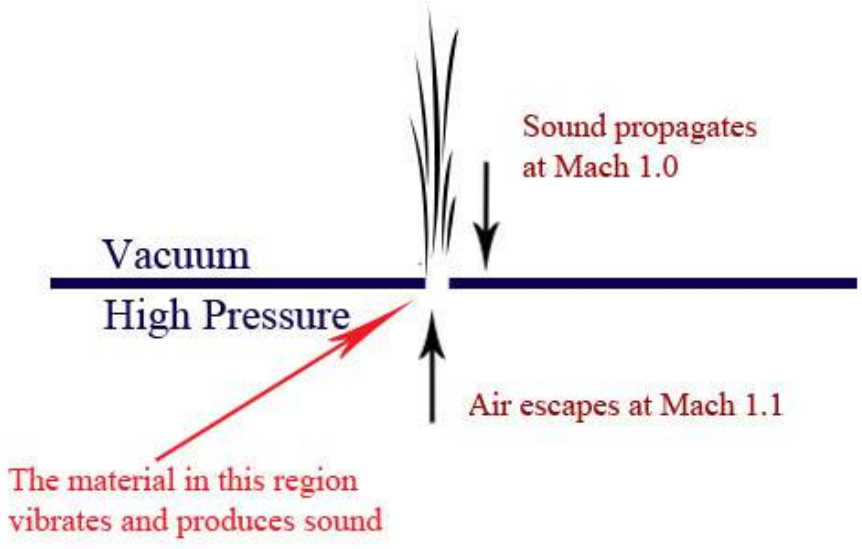


Figure 2.1: The physics of an air leak.

2.4 LEAK LOCALIZATION STRATEGIES

Current technology for the localization of air leaks in spacecraft are slow and demand significant amounts of crew time [1]. For this reason, the National Aeronautics and Space Administration (NASA) has identified this as one area of concern, especially for the International Space Station (ISS).

Various attempts at localization have already been made, using acoustics [10, 14, 12, 13, 15, 16, 17, 18], structure-borne vibrations [11], visible gas and camera-based methods [1, 19]. The latter two methods both rely on mounting cameras outside the pressurized vehicle, and using these to detect leaks. One suggestion is to use breathable, visible gases between inside layers of the ISS, and use the cameras to see the plumes of escaping gas. The other method uses the cameras to record impacts by MMODs.

Other leak detection and identification methods have been proposed in other contexts, including pneumatics and petroleum pipelines [12, 13, 20].

Another technique relies on a randomized array of sensors that use a Bartlett beam-former to focus in on a specific direction, and sweep for high energy points in the approximately 40 kHz range. The system uses many microphones spaced in an approximately random fashion, and by weighting varying the amplitudes of the different microphones, a composite signal can be created representing the sound in a particular direction. These sounds are then analyzed to determine the most likely leak location [14].

Other leak detection strategies have been developed here at the University of Maine's Wireless Sensor Networks (WiSe-Net) Laboratory. This system, the Wireless Leak Detector (WLD) was sent to the ISS for testing in December of 2016. This system uses three pairs of ultrasonic microphones spaced within one half-wavelength of each other to triangulate the leak location. The signals from each pair of microphones are cross-correlated to determine the time delay of arrival. This is then used to triangulate the leak. The final localization is done using a Bayesian Tree-search algorithm or Kalman Filter, as described in [2, 17, 21].

The Kalman Filter is a standard tool in estimation theory; interested readers can see [22] for more information. The Bayesian Tree Search algorithm uses Bayes' Theorem to express the probability that a leak falls in a given region. By using smaller and smaller subdivisions of the possible leak locations, one can estimate the true leak location. For more information, see [21].

2.4.1 LEAK IDENTIFICATION STRATEGIES

While these systems and others show promise in localizing leaks, little work has been done on the identification of the leaking material from its acoustics. To this end, the focus will be on analyzing the acoustics from a leak rather than its location.

A number of leak identification strategies are used in this work. The first was the Fourier Transform (using a Fast Fourier Transform, or FFT), Hilbert-Huang

Transform, Principle Component Analysis (PCA), Wavelet Transform Modulus Maxima (WTMM). The PCA and WTMM were not considered; the results of the transforms were too similar to one another to be of any practical use in identification of leaks. For example, all of the test conditions has principle components between 0.18-0.26. This makes determining the leak material very difficult. The WTMM was used too calculate the Hurst Exponents; those faced a similar problem.

The Fourier Transform and Hilbert-Huang transform were much better, but the FFT assumes that the signal is periodic. This is not exactly true of turbulent flow. This means that the HHT is better suited to these signals.

2.5 MATHEMATICAL NOTATION

In this section the notation used in this thesis is defined explicitly. This is meant to serve as a glossary of notation for the work, and make all of the symbols used in the discussions to follow explicit.

μ	Lebesgue Measure (see [23]).
$L^p(E)$	The Lebesgue space; Banach Space of $f : E \to \mathbb{R}$ with $\int_{\mathbb{E}} \lvert f\rvert^p d\mu < \infty$.
$\mathbb{R}$	The set of real numbers
$\mathbb{C}$	The set of complex numbers
$\overline{z}$	The complex conjugate of $z \in \mathbb{C}$
f'	The derivative of f.
$\hat{f}$	The Fourier transform of f
$F[f]$	The Fourier Transform of the function f
PV	The Cauchy Principal Value
$H[f]$	The Hilbert Transform of the function f

2.6 FOURIER TRANSFORM

Loosely put, the Fourier Transform is a method for changing a signal from a function of time, to a function of frequency. This often makes the signal easier to understand, or shows properties that would otherwise be hidden.

The Fourier Transform is usually computed through the Fast Fourier Transform (FFT); this is an efficient algorithm for computing the discrete Fourier Transform, or the analogue of the integral transform used on a data set.

The Fourier transform was developed by Fourier in 1822. It was first used to solve the heat equation in his work " Mémoire sur la propagation de la chaleur dans les corps solides". Fourier demonstrated that any sufficiently nice function can be written as a sum (or integral) of sines and cosines. Bernoulli and Euler had previously shown that certain functions can be written as a linear combination of sines and cosines, however it was Fourier who showed any (again, sufficiently nice) function can be written as such. Fourier then used this result to solve the heat equation [23, 24].

Since then, the Fourier Transform has grown to prominence in signal processing and mathematics. It has been generalized to other spaces beyond continuous, integrable functions to groups and distributions. Also, numerous computational methods have been developed to implement the Fourier Transform.

2.6.1 DEFINITION

The Fourier Transform, mathematically speaking, involves a writing a given functions in $L^p(E)$ in terms of the Fourier basis. The Fourier basis is comprised of sines and cosines with different frequencies, and the constant one. Expressing a function in this way lets us understand it better a difficult to analyze function can be written as a combination of well understood pieces.

To define the Fourier Transform formally, let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. We will denote the Lebesgue measure as μ. The Fourier transform of f is

$$\hat{f}(\eta) = \int_{\mathbb{R}} f(x) e^{-i\eta x} d\mu \tag{2.1}$$

where η is a real number. Conversely, the Fourier Inversion Formula states f can be written as

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\eta) e^{i\eta x} d\mu \tag{2.2}$$

This is the general definition of the Transform. While entire books have been written on the Fourier Transform, a only few of the pieces needed later on in the work are covered here.

A useful result regarding the Fourier Transform is Parseval's Theorem. It will be stated here without proof. Interested readers can see [25]. Continuing with the notation above, let $g \in L^2(\mathbb{R})$, then

$$\int_{\mathbb{R}} f(x)\overline{g(x)} d\mu = \int_{\mathbb{R}} \hat{f}(\eta)\overline{\hat{g}(\eta)} d\mu \tag{2.3}$$

Plancherel's Theorem is a corollary of this, which states that the L^2 norm of a function is equal to the L^2 norm of its Fourier Transform.

While there is a very rich mathematical theory surrounding the Fourier Transform, its most common modern application likely lies in signal processing. Due to the ease of calculating the discrete Fourier transform, it is often one of the first tools used on a time series to better understand it.

2.7 THE HILBERT HUANG TRANSFORM

In recent years, the Hilbert-Huang Transform has seen a lot of use in the area of signal processing. The transform is based on dividing a signal (time series) into separate varying components called Intrinsic Mode Functions (IMF). While there are many useful tools for looking at time series in this way, the HHT is unique in

that the transformation works well for many data sets where other transformations are not applicable.

The Fourier Transform, as implemented on most software systems, assumes that the data is periodic. One common issue with this assumption comes when there is an underlying trend to the data (increase or decrease in the 'center' of the signal over time). The Fourier Transform is theoretically defined for any square integrable function, there are computational issues in practice. For instance, a Fourier transform of a dataset with a trend may offer a poor reconstruction of the original signal when inverting the transform. This is due to truncating the Fourier Series (which should be infinite) too soon; that is a finite number of Fourier Coefficients my offer a poor representation of our data.

While the HHT has enjoyed solid empirical demonstrations of its effectiveness, it still lacks a solid theoretical framework on which to rest its case [26]. It has been applied to everything from climate data, global positioning systems, tidal models, and in this work, air leak detection in pressurized spacecraft. Part of the work in this thesis is to further develop the theoretical underpinnings of the transform, and to provide an example of its applications.

The HHT was developed by Norden Huang et al in 1992 [26, 27]. This transformation has been expanded and tweaked over a number of years since. After the transform gained some notoriety after demonstrations of its effectiveness. Hypothesis testing for the the results of the transform was developed in 1996 [26]. The Hilbert Transform, used in the HHT, was developed much earlier by David Hilbert in 1905 to answer a problem known today as the Riemann-Hilbert problem [28].

Work has also been done on the theoretical basis of the transformation, but as yet there is no widely accepted interpretation of the HHT [26]. Researchers at NASA have done some work in developing a theoretical background for the HHT.

They have succeeded in proving that the highest frequency components are extracted from the input signal first. They have proved that the function Empirical Mode Decomposition (EMD) sifting process converges for a modified version of the HHT; they rely on dividing the results by increasing powers of two while performing the algorithm. This and other results can be found in [27].

Nonetheless, we will take the theoretical framework in another direction be starting from the beginning of the transform, and also address notions of uniqueness of the limiting functions, and some practical considerations for implementing the transformation.

2.7.1 DEFINITION OF THE HILBERT-HUANG TRANSFORM

The Hilbert-Huang Transform is a data analysis methodology with two main parts [26]:

- *Empirical Mode Decomposition*: Also called the *sifting process*, this is a tool for breaking the input data or input function into a special class of functions for which the Hilbert Transform (step two) is physically meaningful.

- *Hilbert Spectral Analysis*: This step uses the Hilbert Transform to extract the frequency information from the signal.

We will proceed to describe these two parts in depth. The goal of the transform is to extract information about the frequency content of the signal as it changes over time. The HHT serves the same purpose as the Fourier Transform, but it may be applied to a wider range of functions and data sets.

2.7.1.1 EMPIRICAL MODE DECOMPOSITION

The goal of the EMD process is to break the data into a series of *Intrinsic Mode Functions*. IMFs can be thought of as generalized oscillatory modes of a system, except with non-constant frequency and amplitude. These are a class of functions

for which the Hilbert Transform is physically meaningful. Norden Huang, the inventor of the HHT, wrote [26]:

> "An IMF represents a simple oscillatory mode as a counterpart to the simple harmonic function, but it is much more general: instead of constant amplitude and frequency, as in a simple harmonic component, the IMF can have a variable amplitude and frequency as functions of time."

Here it is necessary to establish some useful notation and running assumptions. First, assume that $f : E \to \mathbb{R}$ is a continuously differentiable function n extreme values for $2 < n < \infty$. This will allow us to define the upper and lower envelopes of f, called $U(x)$ and $L(x)$ respectively. These are needed to formally define an IMF.

Definition 1 (Upper Envelope)**.** *The upper envelope of $f(x)$, denoted in this work as $U(x) : E \to \mathbb{R}$, is defined as a function which satisfies:*

- *$U(x) \geq f(x)$ for all x in the domain of f,*
- *$U(x) = f(x)$ if and only if x is a local or absolute maximum of f,*
- *$U(x) \in C^1(E)$, and*
- *if $|f| \leq M$ for some $M \in \mathbb{R}$, then $|U(x)| \leq M$. It is worth noting that not all authors use this final piece of the definition.*

A lower envelope is defined in an analogous way, as follows:

Definition 2 (Lower Envelope)**.** *The lower envelope of $f(x)$, denoted in this work as $L(x) : E \to \mathbb{R}$, is defined as a function which satisfies:*

- *$L(x) \leq f(x)$ for all x in the domain of f,*
- *$L(x) = f(x)$ if and only if x is a local or absolute minimum of f,*

- $L(x) \in C^1(E)$ *is smooth, and finally*
- *if* $|f| \leq M$ *for some* $M \in \mathbb{R}$, *then* $|L(x)| \leq M$.

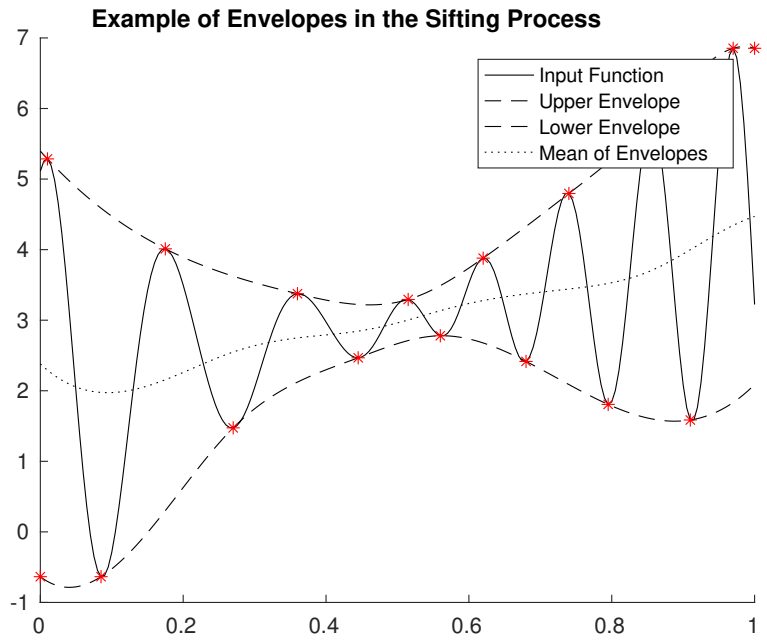


Figure 2.2: An example of envelope functions and their mean.

While this defines of the envelopes, this tells us nothing of how to create such a function explicitly. Typically, this is done using cubic splines as they are convenient to implement on a computer. It is also worth noting that there are many suitable envelopes of a given function, and that relaxing the requirements of smoothness of respecting bounds on f can increase the number of suitable functions further. An example of these functions can be found in Figure 2.2.

With some understanding of an envelope of a function, it is now possible to formally define an IMF.

Definition 3 (Intrinsic Mode Function). *An IMF is a bounded, continuously differentiable function satisfying:*

1. *The number of local extrema (minimum or maximum) and the number of zero crossings are equal or differ by at most one, and*

2. *The mean of the upper and lower envelopes of the function are zero.*

The first criterion requires that the number of minimums and maximums together must be close to the number of times the function changes from negative to positive. The second criterion in the definition is intended to insure the function be centered around the zero; its graph should fall on both sides of the x-axis.

Definition 4 (Empirical Mode Decomposition). *The Empirical Mode Decomposition process is an iterative algorithm. The result of the k-th iteration of the EMD process is denoted f_k, with $f_1 = f$. The algorithm is as follows:*

1. *Draw the envelope functions of f_k, $U_k(x)$ and $L_k(x)$,*

2. *Find the mean of the envelope functions, denoted $m_k(x)$. This is calculated as:*

$$m_k(x) = \frac{U_k(x) + L_k(x)}{2} \tag{2.4}$$

3. *Define $f_{k+1} = f_k - m_k = f_1 - \sum_{j=1}^{k} m_j$*

4. *Repeat the process from step 1 with f_{k+1} until the desired stopping criterion is met.*

5. *If the stopping criteria is met, then set*

$$c_j = f_{k+1} \qquad \text{and} \qquad r_j = f - \sum_{i=1}^{j} c_i \tag{2.5}$$

where c_j is the jth IMF, and r_j is the remainder. Now, if r_j has at least two extreme values, start the process again from step 1 with r_j. Otherwise, output $c_1, c_2, ..., c_j$ and r_j.

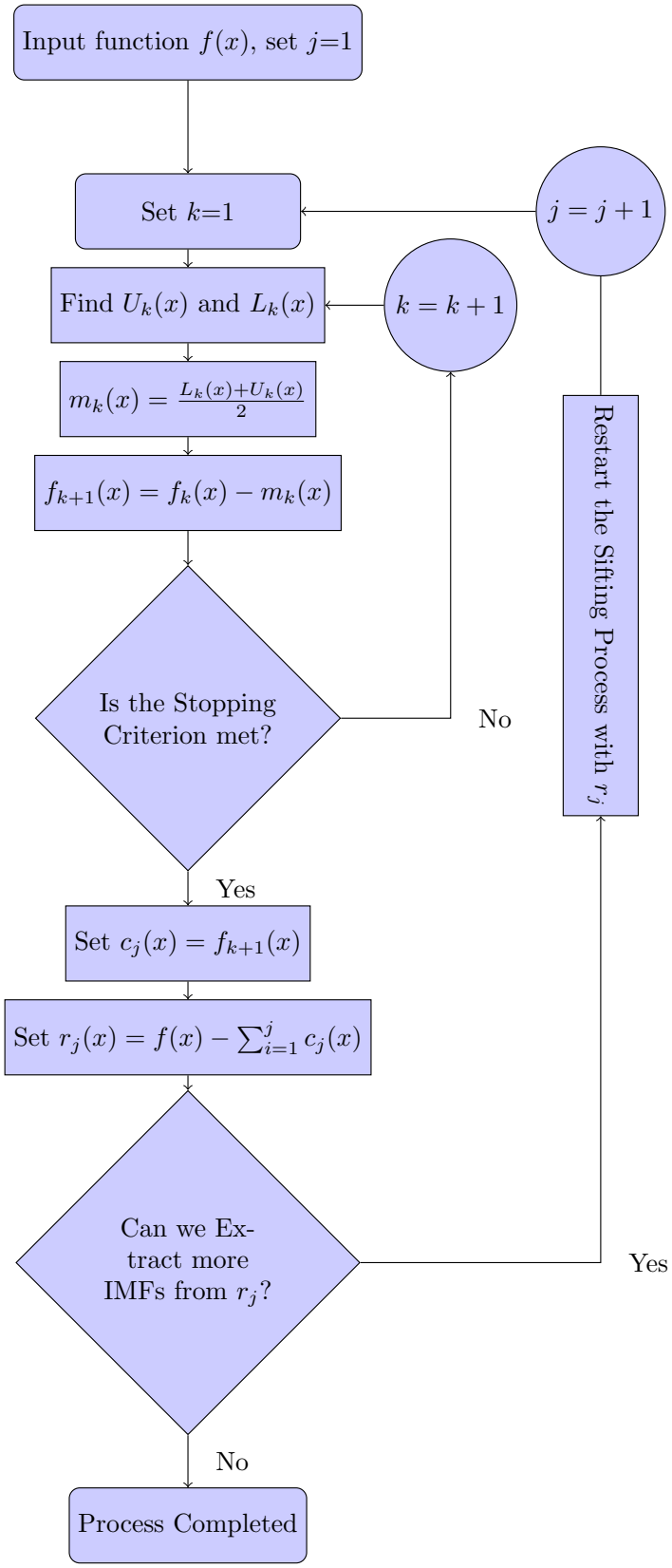


Figure 2.3: The block diagram of the Empirical Mode Decomposition process.

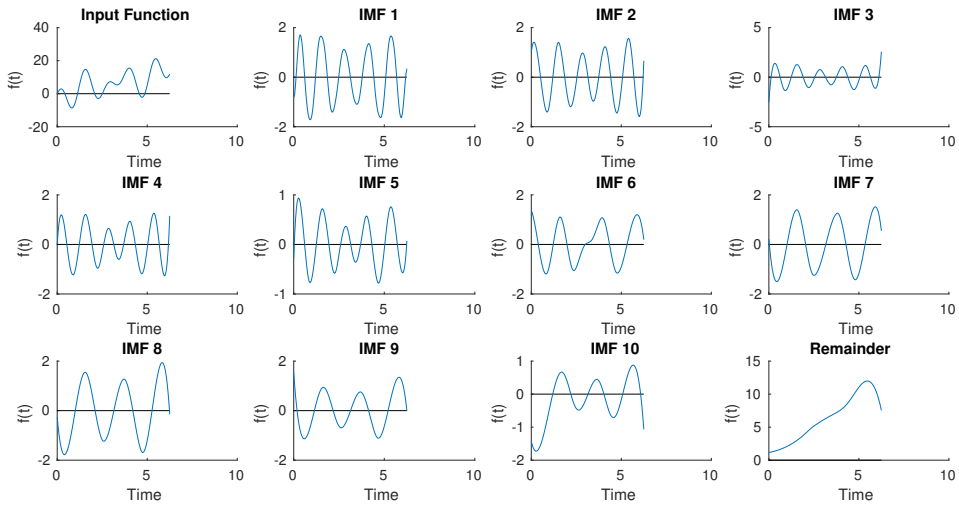


Figure 2.4: Example output of the EMD process.

This process yields a series of IMFs and a single remainder function, which will have at most one extreme value. A flow chart of the EMD process can be found in Figure 2.3.

Now, it is important to define the stopping criterion for the sifting process. It will be shown later that the m_k converge to zero uniformly in $\mathbb{R}$; it is necessary to stop this sifting process in order to extract an IMF. For a discrete-time time series with domain $[0, T]$, the original criterion proposed by Huang et al. suggests stopping when $SD_k < K$, for

$$SD_k = \frac{\sum_{t=0}^{T} |m_{k-1}(t) - m_k(t)|}{\sum_{t=0}^{T} m_{k-1}(t)^2} \tag{2.6}$$

where K is a positive real number. A good value for K has been shown empirically to be 0.2-0.4 [26]. This means that the m_k are sufficiently small to allow f_k to be an IMF or very near one. However, this does not include the definition of IMF in the stopping criterion, so other methods have been introduced.

Another possible stopping criterion is the S-number: If the number of zero crossings and extrema do not change for S successive siftings, stop the process. A good range for S has been shown to be 4 to 8, with 6 showing the best results for

most datasets [26]. However there has been some empirical evidence of different S numbers offering different decompositions of the input dataset.

2.7.1.2 HILBERT SPECTRAL ANALYSIS

The name 'Hilbert Spectral Analysis' hints at the purpose of this step of the HHT: to analyze the frequency content of the data. If the dependent variable is time, then the transformed data is the frequency spectrum. In other cases, the spectrum my have another meaning, but this depends on the application of the transform.

The goal here is to extract the frequency content of the signal. While there are multiple ways of doing this, the HHT is unique in that it does it in a way that is as tailored to the function as possible. We therefore move our function in to the complex plane to extract the argument. After this, the resulting function for the argument is differentiated, and a frequency representation is obtained.

We would like to find a complex analytic function whose real part, evaluated along the real axis, corresponds with our real valued function. If such a function exists, one can extract the argument, and hence, a frequency representation of our input signal. This can be done with the Hilbert Transform [25, 29].

In order to do this, we need to first define the Cauchy Principal Value of an integral. Let $f(x)$ a functions with a singularity at the point $x = b$, with $a < b < c$. Then

$$PV \int_a^c f(x)dx = \lim_{\epsilon \to 0+} \int_a^{b-\epsilon} f(x)dx + \int_{b+\epsilon}^c f(x)dx \tag{2.7}$$

The Cauchy Principal value is used for getting values of integrals that would otherwise be out of reach. For instance, the the function $f(x) = x^{-1}$ is not integrable at $x = 0$, but the principle can be used to get a value of some integrals involving this function over intervals containing zero.

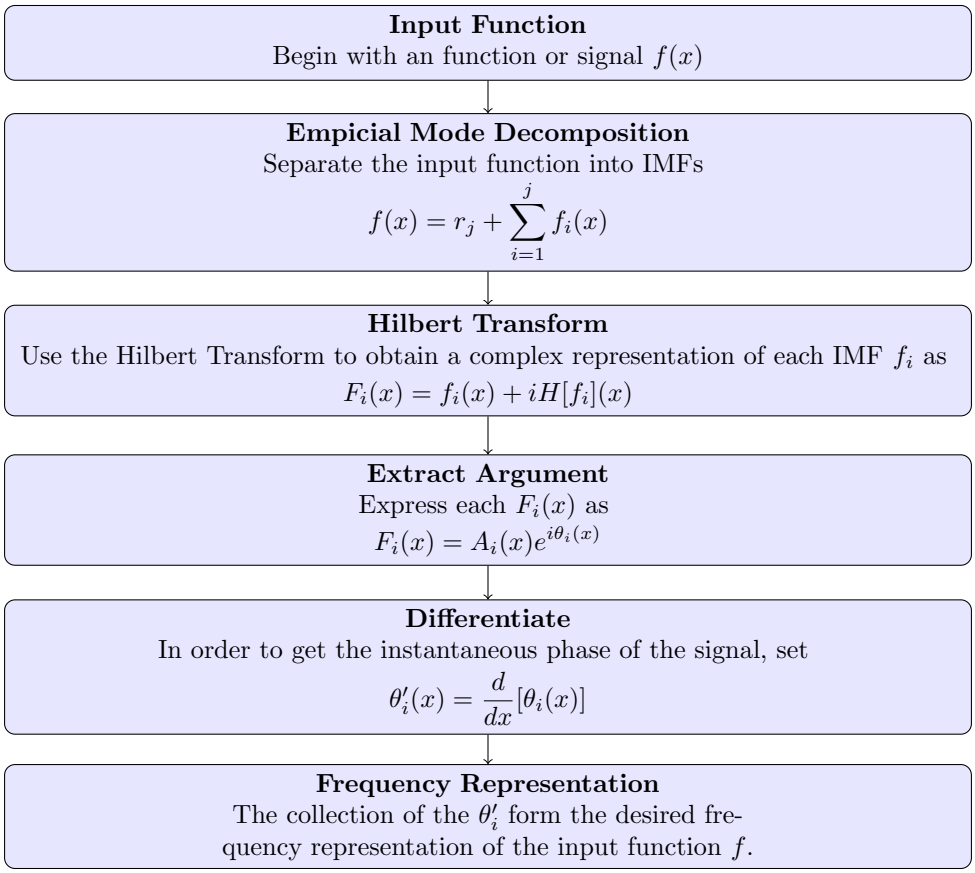


Figure 2.5: The process of Hilbert Spectral Analysis.

Now that the Cauchy principal value is defined, we can move on to the Hilbert Transform. The Hilbert Transform is defined as [26, 25]

$$H[f(t)] = \frac{1}{\pi}\ \mathrm{PV}\int_{-\infty}^{\infty}\frac{f(t)}{t-\tau}d\tau = \frac{1}{\pi}\lim_{\epsilon\to 0+}\int_{\epsilon}^{\infty}\frac{f(t+x)-f(t-x)}{x}dx \tag{2.8}$$

where PV denotes the Cauchy Principle value. The two forms of the transform above are equivalent after expanding the Principle Value and setting $x = t - \tau$. The transform is defined for any $f \in L^2$ [25]. This transform has been shown to map L^2 functions to L^2 functions. Interested readers are invited to see [25] for more information.

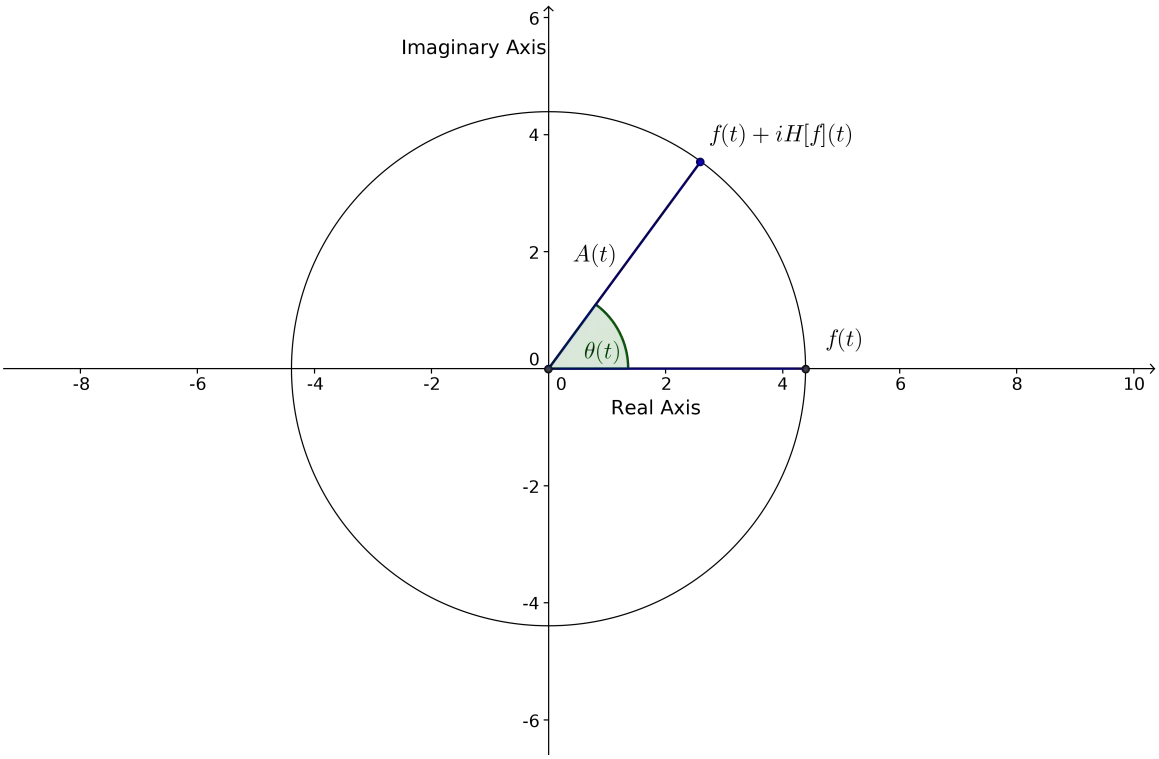


Figure 2.6: The Hilbert Transform, mapping a real valued function into the complex plane.

The Hilbert Transform is used in the HHT to extend the input function to the complex plane. This mean one can write

$$f(t) \to f(t) + iH[f(t)] = A(t)e^{i\theta(t)} \tag{2.9}$$

Here $A(t)$ is the amplitude function and $\theta(t)$ is the phase or angle function. From here, the argument of the transformed function, $\theta(t)$, is of particular interest. In most real world applications, this corresponds to the frequency-domain content of the signal. As shown in Figure 2.6, the Hilbert Transform maps the real function to a complex valued one. We can represent this in polar coordinates, and extract the arguments of the signal.

The figure shows the point $f(t)+iH[f](t)$ on a circle of radius $a(t)$; we do not get a neat circle for an arbitrary function. The path traced by $f(t) + iH[f](t)$ is a circle for $f(t) = a_0 \cos(\theta(t))$ for $\theta(t)$ a monotone function and a_0 a real constant. However,

it is possible for $f(t) + iH[f](t)$ to trace a path far from the origin, which will give a $\theta(t)$ that falls within a narrow band or obscures some useful frequency information. For this reason, we use the Empirical Mode Decomposition to center to create IMFs centered a zero, whose Hilbert Transform will give good frequency information.

One important theorem related to the Hilbert Transform is Bedrosian's Theorem. This theorem, speaking loosely, states that the Hilbert Transform of the product of two band limited signals is the product of the lower frequency signal and the Hilbert Transform of the high frequency one. Formally, the theorem states [30]

Theorem 2.1. *Let $f(x)$ and $g(x)$ be denote generally complex functions in $L^2(-\infty, \infty)$ of the variable x. If*

1. *The Fourier Transform, $F(u)$, of $f(x)$ vanishes for $|u| > a$ and the Fourier Transform, $G(u)$, of $g(x)$ vanishes for $|u| < a$ for some constant a*

2. *f and g are analytic (i.e. their real and imaginary parts are Hilbert pairs)*

then the Hilbert transform of the product of $f(x)$ and $g(x)$ is given by

$$H[f(x)g(x)] = f(x)H[g(x)] \tag{2.10}$$

This basically says that we can 'factor out' a slow varying component from the Hilbert Transform.

Once $\theta(t)$ has been determined, the instantaneous frequency is its derivative [26]:

$$\theta'(t) = \frac{d}{dt}[\theta(t)] \tag{2.11}$$

This is one of the main advantages of the HHT over the Fourier Transform. It allows this type of decomposition into instantaneous frequencies. This type of representation is not possible with the Fourier transform.

The other advantage of the HHT is the assumptions about the input function. The Fourier Transform assumes that the frequency content of a signal is constant

[26]. The HHT does not make this assumption, and allows varying frequency components. This means the HHT is suitable for virtually any signal. For example, the Fourier Transform assumes that the energy at 40kHz is approximately constant throughout a data set. The Hilbert Huang Transform allows the energy at 40kHz to vary.

2.8 ARTIFICIAL NEURAL NETWORKS

Artificial neural networks (ANN) are computer programs that attempt to solve complex problems by mimicking the way human brains work. They are formed from layers of 'neurons' which work together to recognize patterns or make predictions.

The earliest known artificial neural network is the Perceptron, due to Frank Rosenblatt in 1958 [31, 32]. ANNs saw some scattered development in the late 50s and 60s. The next architectural change came in 1988 with the invention of Madaline, a network of linear sum-and-forward nodes. While the stage was set for the development of ANNs, it was not until later that they gained prominence as a problem-solving methodology [31, 32]. In their early days, the computation required to train and use the networks was prohibitive. Today, modern computers are fast enough to run a trained network without any noticeable delay. Training the networks takes hours for large data sets, but this is typically not prohibitive as this step must typically be completed only once.

2.8.1 DEFINITION

Neural networks are divided into a series of layers. The first is the input layer; this is simply the input data. From here, the data is processed by one or more hidden layers, and then the data is passed to the output layer, which gives the user the final results of the processing. This is shown in Figure 2.7

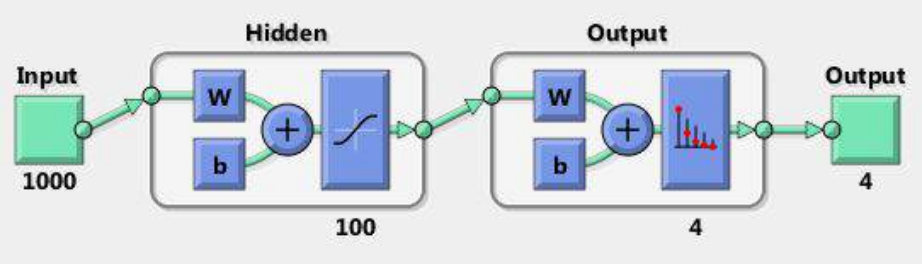


Figure 2.7: A diagram of the structure of an ANN for pattern recognition.

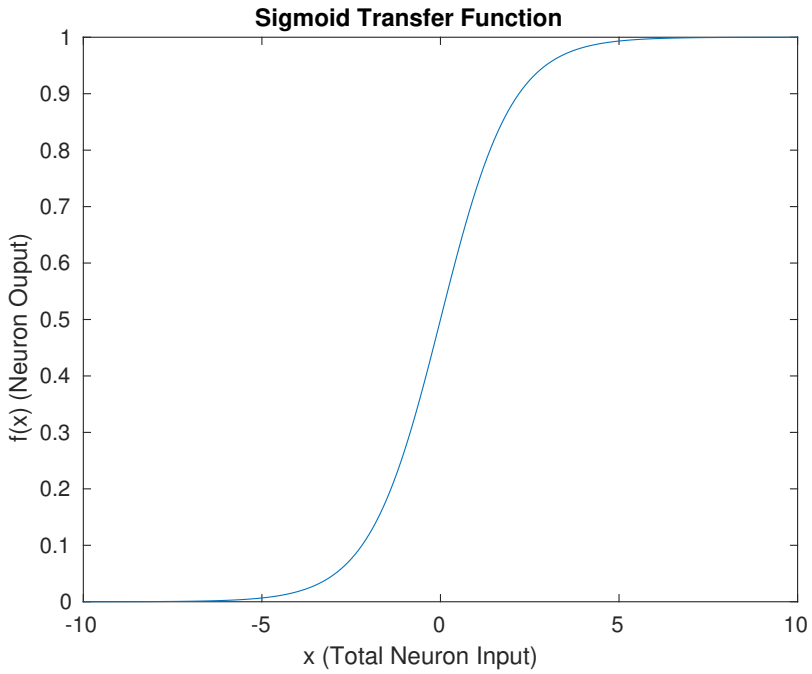


Figure 2.8: The most common type of transfer function for neural networks.

The hidden layers do the real processing in the system. While they are a powerful in solving problems where a large data set can be obtained, they have drawbacks. For instance, they are black boxes, with very little transparency. They also require significant computer resources compared to other methods, especially in the training phase.

The nodes or neurons for a highly connected system, and the data passes through the layers. The system is called feed-forward if the data only passes in one direction. Most neural networks work this way, but not all of them [31, 32].

Each node has a transfer function $f : \mathbb{R} \to \mathbb{R}$ and multiple inputs, x_i. During the training process, each input is assigned a weight, $w_i \in \mathbb{R}$. The node outputs $f(\sum w_i x_i)$. A graph of the transfer function can be see in figure 2.8. These functions are typically sigmoid and soft-max functions.

While training the network, the weights w_i are adjusted for each node. This is done using a gradient decent scheme for high dimensional optimization. Specifically, the neural networks were trained with a scaled conjugate gradient descent method included in the MATLAB neural networking toolbox (trainscg). We write the error function of the system as a hyper-surface, and choose a point. From here, we follow the path where the gradient points down the steepest. Now, viewing the weights of all inputs as a transformation, we choose a path that is orthogonal to the previous directions traveled, and scale the gradient accordingly. This process, it can be shown, converges (with exact arithmetic) in n iterations, where n is the dimension of the matrix of weights. The process is similar to Newton's method from elementary calculus, only generalized to include all directional derivatives. This means that the gradient decent algorithm cannot find saddle points, only minima, unlike Newton's method. For this reason it cannot be used in applications like Lagrangian Optimization (at least not without a little creativity). However, since the purpose of the training process is to minimize the error, we can use gradient decent here. In fact, it was advances in gradient decent algorithms, combined with the increase in computational power of modern PC, that have lead to the rise of neural network applications today.

Neural networks have been used extensively in speaker recognition in audio and image recognition in image processing. For this reason it is logical to extend the approach to recognize a leak from its sound, rather than a human from their voice. The approach used here is similar to [33]. The training methods and exact algorithms used will be details in the methods section.

CHAPTER 3

THEORETICAL FRAMEWORK FOR THE HILBERT-HUANG TRANSFORM

3.1 INTRODUCTION

There is currently no known theoretic framework on which the HHT can rest its case. In order to give the transform such a framework, this chapter begins with refining the definition of IMF. Then come some important open questions that must be answered: Does the EMD sifting process converge, and if so, is that limit unique? Does the HHT always admit a decomposition of the input function into finitely many IMFs?

In order to answer these questions, the following theoretical framework for the Hilbert-Huang Transform is proposed. For sufficiently nice envelope functions:

1. Given the correct choice of envelopes, m_k converges uniformly everywhere to the zero function.

2. The sifting process, if stopped at the appropriate time, gives an IMF. In other words, there exists a stopping criterion which allows one to stop the sifting process and obtain something as close as one likes to an IMF.

3. The Hilbert transform extracts the frequency content of the IMFs.

This is part of a theoretical framework for the Hilbert-Huang Transform. However, this is not without a key assumption that is not true of the Fourier Transform: The domain of the function we wish to transform must be compact. In most applications, this corresponds to assuming that the signal is bounded in time. While this offers no practical drawbacks in most applications, it is worth noting the mathematical necessity of this assumption for the proofs to follow.

3.2 ALTERNATIVE DEFINITION OF INTRINSIC MODE FUNCTION

This chapter begins with proving an equivalent definition of IMF which will be much easier to use. The following definition is equivalent to the previous one given in Chapter 2. Recall that an IMF is defined as follows:

Definition 5 (Intrinsic Mode Function(1)). *An IMF is a bounded, continuously differentiable function satisfying:*

1. *The number of local extrema (minimum or maximum) and the number of zero crossings are equal or differ by at most one, and*

2. *The mean of the upper and lower envelopes of the function are zero.*

We will show that the following definition is equivalent:

Definition 6 (Intrinsic Mode Function (2)). *A continuously differentiable function $f : E \to \mathbb{R}$ is an* Intrinsic Mode Function *if it has at least two extreme values, with all local maxima positive and all local minima negative.*

Proof. The proof of the equivalence of the definitions has two parts: First, it is shown that this definition is sufficient to meet the requirements of an IMF by the previous definition. The proof of the converse will proceed by contrapositive; we will show that violating the above definition violates the previous definition of an IMF. This will show the logical equivalence of the two.

Assume f is a function with at least two extreme values, with all maximum values positive and all minimum values negative.

Since f is continuous, it must have a zero crossing between each successive minimum and maximum by the intermediate value theorem. Therefore, the number of roots of f must be within one of the number of extrema of f. This guarantees that f meets the first part of the definition of an IMF.

We will now show the existence of envelope functions whose sum is zero. Let U be any upper envelope on $|f|$. Then U is an upper envelope for f also, as $U = f$ for each maximum only, and $f \leq |f| \leq U$. Furthermore, $L = -U$ is a lower envelope on f. Clearly $U + L = 0$, completing the first half of the proof.

Assume, to the contrary, that f is a function with a minimum that is positive or a maximum that is negative at a point c in its domain. Assume that there exist envelopes U and L such that $U + L = 0$.

If c is a local maximum, then for any envelope U, $U(c) = f(c)$. Then $L(c) = -f(c) > f(c)$ as $U(c) + L(c) = 0$. This is a contradiction.

Alternatively, if c is a local minimum, $L(c) = f(c)$, so $U(c) = -f(c) < f(c)$, again a contradiction. Therefore, f is not an IMF, completing the proof. □

This definition has a significant advantage over the previous one: it is easy to verify, and is without ambiguity. The criterion regarding envelopes in other definition is not clear cut, so we offer this replacement. Furthermore, this makes checking whether the sifting process has given an IMF much simpler.

As nice as this definition is, it still lacks a degree of specificity that may be desired. It is not well understood exactly what functions are allowed in this class, and so one may wish to limit the class of IMFs further. One way of doing this is to require the mean of two specific envelopes be zero, rather than simply requiring the existence of envelopes that sum to zero. Therefore,

Definition 7 (Intrinsic Mode Function (3))**.** *An IMF is a bounded, continuously differentiable function satisfying:*

1. *The number of local extrema (minimum or maximum) and the number of zero crossings are equal or differ by at most one, and*

2. *The mean of the upper and lower envelopes ($U(x)$ and $L(x)$, respectively) of the function are zero,*

where U is defined segment-wise as follows. Let a_i and b_i be the minima and maxima of the input function. On the interval $[a_i, b_i]$,

$$U(x)|_{[a_i,b_i]} = f(a_i) + \frac{P_k - f(a_i)}{f(b_i) - f(a_i)} \int_{a_i}^{x} f'(t)dt \tag{3.1}$$

where P_k is defined as before as

$$P_k = \begin{cases} |f(b_i)| & f(b_i) < 0 \\ \min(f(a_i), f(a_{i+1})) & f(b_i) \geq 0 \end{cases} \tag{3.2}$$

And L is defined analogously.

This definition is of course not equivalent to the previous one, but may be superior.

3.3 CONVERGENCE OF THE SIFTING PROCESS

In this section, it is demonstrated that the mean of the function envelopes, $m_k \to 0$ uniformly as $k \to \infty$. This convergence comes in several parts. The sifting process is shown to be bounded and monotone at the extreme values of the input function. This will require some discussion of the envelopes. Furthermore it is shown that $f_k \to 0$ as $k \to \infty$ as well.

Some work has already been done in this area[26, 27, 27, 34, 35, 36, 37]. Most notably, there has been some work using cubic splines or basis splines to show convergence. This angle looks promising, but it is not the approach used here. Instead, it is shown that the sifting process converges for the most general envelope functions possible. In order to do this, it is necessary to define what exactly is a sufficiently nice envelope.

The definition of an upper and lower envelope must be refined. The generally accepted definition from signal processing allows for a few pathological examples that would not converge. For example, if there was a large bump between two

maximums, that would distort the mean of the envelopes, and create unwanted distortions, possibly introducing new modes to the resulting function.

Furthermore, very rapid changes in the envelope are undesirable. It should somehow be less variable than the input functions. One nice way to avoid these pathological examples is to make a requirement that the envelopes take extreme values only at the points where the input function does.

Proving the convergence of the EMD process will require the following steps:

1. Discussion of envelope functions and the implications to the Sifting Process.
2. Show that the sifting process produces bounded functions, with the same upper bound as the input function,
3. Show that the partial sums of $m_k(c)$ is monotone if c is a local maximum or minimum of the input function.
4. Finish by showing that the sifting process converges to zero for tight envelopes.

Let us first define a criterion for what makes an envelope desirable.

Definition 8 (Tight Envelope)**.** *An envelope ($L \leq f \leq U$) is* tight *if*

1. $\operatorname{sgn}(f'(x)) = \operatorname{sgn}(U'(x)) = \operatorname{sgn}(L'(x))$, *and*
2. $|U'(x)| < |f'(x)|$ *and* $|L'(x)| < |f'(x)|$ *for all x in the domain of f which are not critical points of f, and*
3. *If c is a minimum of f with $f(c) < 0$, then $|U(c)| \leq |f(c)|$ and*
4. *If c is a maximum of f with $f(c) > 0$, then $|L(c)| \leq |f(c_2)|$.*

Specifically, it will be shown that as long as the envelopes of f_k are tight, the sifting process converges to an IMF. However, before this, it is useful to know that such envelopes exist. Better yet, the next proposition offers a method of constructing such an envelope:

Proposition 3.1 (Existence of Tight Envelopes)**.** *Let $f : E \to \mathbb{R}$ be a continuously differentiable function whose first derivative is Riemann Integrable, and let f have n local extreme values, for $2 < n < \infty$. Then there exists tight envelopes for f.*

Proof. Let $[a, b]$ be a subinterval of E such that a is a maximum of f and b is a minimum of f. let c be the maximum of f following the minimum at b. The general method of this construction defines $U(x)$ on $[a, b]$ to be

$$U_{[a,b]}(x) = C + p \int_a^x f'(t)dt \tag{3.3}$$

where C and p are constants, and $-1 < p < 1$. In this way, we notice $|U'(x)| = |pf'(x)| \leq |f'(x)|$ with equality happening only at the critical points of f. Furthermore, the sign of f' and U' will always be the same. This shows that U meets the first two requirements of a tight envelope. The last requirements of a tight envelope is that U not be too far from $f(b)$ as b is a maximum. We will construct a point P_b which will satisfies the last requirement when $f(b) < 0$ and is a reasonable choice of value when $f(b) > 0$.

$$P_b = \begin{cases} q|f(b)| & f(b) < 0 \\ \frac{f(b)+f(c)}{2} & f(b) \geq 0 \end{cases} \tag{3.4}$$

where q is a real number between zero and one. Any envelope with $U(a) = f(a)$ and $U(b) = P_b$ will satisfy the last requirement of a tight envelope. Then we can construct U as follows:

$$p = \frac{P_b - f(a)}{f(b) - f(a)} \qquad \text{and} \qquad C = f(a) \tag{3.5}$$

This can be shown to satisfy the requirements of the endpoints:

$$U(a) = f(a) + p \int_a^a f'(t)dt = f(a) \tag{3.6}$$

as needed. At the other end,

$$U(b) = f(a) + \frac{P_b - f(a)}{f(b) - f(a)} \int_a^b f'(t)dt = f(a) + \frac{P_b - f(a)}{f(b) - f(a)}(f(b) - f(a)) \tag{3.7}$$
$$= f(a) + P_b - f(a) = P_b$$

again as needed. All that remains is that we show $-1 < p < 1$. Since $f(b) < 0$ and $f(a) > f(b)$

$$f(b) - f(a) > 0 \qquad \text{and} \qquad -(\frac{f(b)}{2} + f(a)) > 0 \tag{3.8}$$

And so $p > 0$. Since $f(a) > f(b)$ there exists a q such that:

$$f(a) > qf(b) > f(b) \tag{3.9}$$
$$\therefore 0 > qf(b) - f(a) < f(b) - f(a)$$
$$\therefore 0 < \frac{qf(b) - f(a)}{f(b) - f(a)} < 1$$

where the last line holds as $f(b) - f(a) < 0$.

Now, between the b and the next extreme value of f, denoted c, we have

$$U_{[b,c]}(x) = P_b + \frac{f(c) - P_b}{f(c) - f(b)} \int_b^x f'(t)dt \tag{3.10}$$

Now, we note that the function is continuous at b. Furthermore, $U'_{[a,b]}(b) = U'_{[b,c]}(b) = pf'(b) = 0$, so we have a continuous first derivative as well.

Thus function meets the requirements of being a tight envelope. □

This construction is not a useful one in most applied work, as numerical differentiation is not ideal. However, it shows the existence of tight envelopes. Also, there is some choice of the value of q, so there are multiple options for tight envelopes as well. However, this is not an issue in practice.

Let us begin demonstrating the convergence of the sifting process by proving the following proposition:

Proposition 3.2 (Extreme Values of Tight Envelopes)**.** *If the envelopes* (U, L) *of* $f : E \to \mathbb{R}$*,* U *and* L*, are tight, then* f*,* $m = (U + L)/2$*, and* $f - m$ *have extreme values at the same points in the domain of* f*.*

Remark: this follows directly from the fact if that U and L have maximums at a given place, so does their sum. Nonetheless a proof is included next.

Proof. Fix $\epsilon > 0$. Let c be a point in the domain for which $f(c)$ is a local maximum of f.

First, note that all of the functions involved are continuously differentiable. Then we may see that $\operatorname{sgn}(m'(x)) = \operatorname{sgn}(U') = \operatorname{sgn}(f')$. Furthermore, we see This means

$$|m'(x)| = \frac{|L'(x) + U'(x)|}{2} \leq |f'(x)| \tag{3.11}$$

Note that equality occurs only when $f'(x) = 0$. Since $f(c)$ is a maximum, $f'(c) = U'(c) = L'(c) = 0$. Therefore, consider the derivative of m, m' at c with:

$$m'(c) = \frac{U'(c) + L'(c)}{2} = 0 \tag{3.12}$$

Therefore, m has a critical point at c. By the first condition of tight envelopes, U' and L' are positive before the maximum, and U' and L' are negative after the maximum. By the first derivative test from introductory calculus, m' also has a maximum at c.

Now notice the derivative of $f - m$ is zero at c. We must show that $f' - m'$ has the same sign as f'. Since U and L are tight, $|f'(x)| > |U'(x)|$ and $|f'(x)| > |L'(x)|$. Therefore $|f'(x)| > |m'(x)|$. Since $U'(x)$ and $L'(x)$ have the same sign as f', so does m'.

We note that $m' < f'$ before the maximum of f, and $f' < m'$ after the maximum of f. This follows from the fact that f' and m' have the same sign and $|m'| < |f'|$. Thus $f - m$ has a maximum, as needed. □

Proposition 3.3 (Stationary Maxima)**.** *If c is a maximum of f, then* $f(c) - m(c) > 0$.

Remark: This proposition guarantees that the local maximum values of f_k are positive. This will be necessary for further results.

Proof. Notice that $L(x) < m(x) < U(x)$ for all x. Let c be a local maximum of f, so $f(c) = U(c)$. Then $m(c) < U(c) = f(c)$. Therefore $f(c) - m(c) > 0$. □

Therefore after sifting all maximum values of $f - m$ are positive. Similarly, all minimum values are negative. The proof is analogous to the one above, but it is not included here for the sake of brevity.

Now, there are two propositions left to prove regarding the sifting process. For the remainder of this section, f_k will denote the output of the kth iteration of the sifting process.

Proposition 3.4 (Nonnegativity)**.** *Let c be a maximum of f_k for each $k \geq N$, $N \in \mathbb{N}$. If f is sifted using tight envelopes, the sequence $m_k(c)$ is nonnegative for $k > N$.*

Proof. First, notice that $U_k(c) = f_k(c)$. Then

$$m_k(c) = \frac{f_k(c) + L_k(c)}{2} \qquad \therefore \qquad f_{k+1} = f_k(c) - m_k(c) = \frac{f_k(c) - L_k(c)}{2} \tag{3.13}$$

This expression relates f_k to f_{k+1} at c. From proposition 3.1, $f_k(c)$ is nonnegative for $k > 1$. Furthermore, $L_k(c) < f_k(c)$, and so $f_k(c) - L_k(c) > 0$. Therefore $f_{k+1}(c) \geq 0$.

Since L_k is tight, $|L_k(c)| \leq f_k(c)$ whenever $f_k(c) > 0$. However we know from the previous proposition that the maximums of $f_k(c)$ are positive for k sufficiently large. In this case, we may use the last requirement for tight envelopes to conclude $f_k(c) + L_k(c) > 0$, and so $m_k(c) > 0$. This completes the proof. □

Now the partial sums of the series $\sum_{j=1}^{\infty} m_j(c)$ is a monotone sequence. Next it will be shown that each $f_k(x)$ is bounded. This brings us to the following proposition:

Proposition 3.5 (Convergence and Summability)**.** *Let $f : E \to \mathbb{R}$ be a function such that $|f(x)| < M$ for all $x \in E$ and some $M \in \mathbb{R}$. Then $|f_k(x)| < M$ for all x and k, and*

$$\sum_{k=1}^{\infty} m_k(x) < 2M. \tag{3.14}$$

.

Proof. Note that $|f_1(x)| = |f(x)| < M$. From here we will proceed by induction. Assume $|f_{k-1}(x)| < M$ for all x. Then

$$\begin{aligned} f_k(x) = f_{k-1}(x) - m_{k-1}(x) &= f_{k-1}(x) - \frac{U_{k-1}(x) + L_{k-1}(x)}{2} \\ &\leq U_{k-1}(x) - \frac{U_{k-1}(x) + L_{k-1}(x)}{2} = \frac{U_{k-1}(x)}{2} - \frac{L_{k-1}(x)}{2} \\ &\leq \frac{1}{2}|U_{k-1}(x)| + \frac{1}{2}|L_{k-1}(x)| \leq M \end{aligned} \tag{3.15}$$

This yields an upper bound on the function. However, it is necessary to check that the same bound works for $-M$.

$$\begin{aligned} f_k(x) = f_{k-1}(x) - m_{k-1}(x) &= f_{k-1}(x) - \frac{U_{k-1}(x) + L_{k-1}(x)}{2} \\ &\geq L_{k-1}(x) - \frac{U_{k-1}(x) + L_{k-1}(x)}{2} = \frac{L_{k-1}(x)}{2} - \frac{U_{k-1}(x)}{2} \\ &\geq -\frac{1}{2}|L_{k-1}(x)| - \frac{1}{2}|U_{k-1}(x)| \geq -M \end{aligned} \tag{3.16}$$

This completes the proof of the first part. Now, for all $k \geq 1$,

$$\begin{aligned} -M < f_k(x) = f(x) - \sum_{j=1}^{k} m_j(x) < M \\ \implies -M - f_k(x) < -\sum_{j=1}^{k} m_j(x) < M - f_k(x) \\ \implies -2M < \sum_{j=1}^{k} m_j(x) < 2M \end{aligned} \tag{3.17}$$

□

Building on the results above,the main theorem is presented next:

Theorem 3.6 (Convergence)**.** *For any continuously differentiable function on a compact domain, the Empirical Mode Decomposition Sifting Process converges, in the sense that the sequence* $\{m_k\}$ *is uniformly Cauchy, provided that the sifting is done with tight envelopes.*

Proof. Let $f : E \to \mathbb{R}$ be a continuously differentiable function and let c_i, $i = 1, 2, 3, ..., r$ be the extreme values of f including the endpoints. First, for $k > 1$, notice that $f_k(c_i) > 0$ if c_i is a maximum, and $f_k(c_i) < 0$ if c_i is a minimum, by 3.3. This allows us to apply 3.4, which tells us that the m_k are monotone at these points. However, they are also bounded at these points. Therefore,

$$\sum_{j=1}^{\infty} |m_j(c_i)| < 2M \qquad \therefore \qquad \lim_{j\to\infty} m_j(c_i) = 0 \tag{3.18}$$

This gives us that the extreme values of m_k go to zero as k increases to infinity. Note that since m_k is defined on a compact domain, it has an absolute maximum and an absolute minimum, each of which converges to zero. Therefore $m_k(x) \to 0$ uniformly on E as $k \to \infty$.

Now, we show that for all $x \in E$, the sequence $f_k(x)$ is Cauchy, and hence convergent. Fix $\epsilon > 0$. By 3.14, there exists $N \in \mathbb{N}$ such that

$$\sum_{j=N}^{\infty} m_k(x) < \epsilon \tag{3.19}$$

Now for $n, q \geq N$

$$|f_q(x) - f_n(x)| = |f(x) - \sum_{j=1}^{q} m_j(x) - f(x) + \sum_{j=1}^{n} m_j(x)| = |\sum_{j=q}^{n} m_j(x)| < \epsilon \tag{3.20}$$

So $f_k \to f^*$ for some $f^* \in L^1(E)$. This means the sifting process converges point-wise on the domain of f. However, the convergence is actually uniform on E, as shown next: First, define a sequence M_k as

$$M_k = \max(|m_k(c_1)|, |m_k(c_2)|, ..., |m_k(c_r)|) \tag{3.21}$$

Since each of the finitely many terms that determine M_k converge to zero, M_k converges to zero, too. Specifically, there exists $N \in \mathbb{N}$ such that for $k \geq N$, $M_k < \epsilon$. Note that:

$$\sup\{|m_k(x) - 0| \big| x \in E\} \leq M_k < \epsilon \tag{3.22}$$

This shows that $m_k \to 0$ uniformly as $k \to \infty$. Now to show this extends to f_k, we need to show that the series of M_ks converge:

$$\sum_{j=1}^{\infty} M_k \leq \sum_{i=1}^{r} (\sum_{j=1}^{\infty} m_j(c_i)) < 2rM \tag{3.23}$$

Hence the series of M_k is bounded and summable. Therefore, there exists N in $\mathbb{N}$ such that $\sum_{j=N}^{\infty} M_j < \epsilon$.

Therefore, for $p, q > N$,

$$\sup\{|f_p(x) - f_q(x)| \big| x \in E\} = \sup\{|\sum_{j=p}^{q} m_k(x)| \big| x \in E\} \leq M_p < \epsilon \tag{3.24}$$

And finally the f_k are uniformly Cauchy, and therefore, uniformly convergent. This mean the limit function, f^*, is continuous, since it inherits continuity from f.

□

While this result is nice, the sifting process cannot be repeated indefinitely. In practice a stopping criterion is used to determine when to stop the sifting process. So, does such a stopping criterion exists?

Theorem 3.7 (Stopping Criterion). *For any $\epsilon > 0$, there exists $N \in \mathbb{N}$, such that for each $k > N$ m_k is within ϵ of zero, and the number of of extreme values of f_k and the number of zeros of f_k differ by at most one.*

Remark: This means f_k is an IMF, except that the mean of the envelopes may be as small as needed, rather than exactly zero.

Proof. One must show that f_k is indeed an Intrinsic Mode Function. The mean of the envelopes is m_k, which has been shown to converge uniformly to zero for k sufficiently large.

Now, from Proposition 3.2, we know that the maxima of f_k are positive and the minima are negative. Therefore f_k meets the second definition of IMF.

□

This means that stopping the sifting process will give us an IMF. This requires selecting an adequate stopping criterion.

While these results are nice, they only apply to sifting with tight envelopes. Is this condition necessary for convergence? We conjecture that it is more strict than necessary. Furthermore, sifting with this type of envelopes does not allow the sifting process to create new minimums or maximums that were not present in the original function. This behavior may be desirable in some applications. While the use of tight envelopes may not be required, there exist envelopes for which no IMF is admitted.

Now we would like to extend this result to other envelope functions. First, we will extend this to a limit of a sequence of tight envelopes.

Proposition 3.8 (Convergence of Limits of Tight Envelopes)**.** *Let $U_{n,k}$ and $L_{n,k}$ be a sequence of tight envelopes converging in n to an envelope U_k uniformly. Then the EMD process converges as before using U_k and L_k as envelopes.*

Proof. Fix $\epsilon > 0$. For each n, $U_{n,k}$ and $L_{n,k}$ are tight. Therefore we know

$$\lim_{n\to\infty}\lim_{k\to\infty} m_{n,k} = \lim_{n\to\infty} 0 = 0 \tag{3.25}$$

Therefore all we need to do is justify the switch the order of the limits above. This can be done, as we know that the convergence in n and k is uniform, and the U_k exists for each k by the assumptions of the theorem. Note that since $U_{n,k}$ and $L_{n,k}$

converge uniformly, for n large,

$$m_{n,k} - m_k = \frac{U_{n,k} + L_{n,k}}{2} - \frac{U_k + L_k}{2} = \frac{U_{n,k} - U_k}{2} + \frac{L_{n,k} - L_k}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \quad (3.26)$$

Therefore, the limits may be exchanged and

$$\lim_{k\to\infty} \lim_{n\to\infty} m_{n,k} = \lim_{n\to\infty} \lim_{k\to\infty} m_{n,k} = 0 \quad (3.27)$$

This completes the proof. □

Now that the convergence of such envelopes has been established, one should construct such a sequence of tight envelopes.

Proposition 3.9 (Extension of Convergence Theorem)**.** *Let $f : E \to \mathbb{R}$ be a continuously differentiable function which has finitely many extreme values and whose first derivative is Riemann Integrable. Assume also that $|f(x)| \leq M \ \forall x \in E$. Let a_i be the i^{th} maximum of f and let b_i be the i^{th} minimum of f. Then the sequence of tight envelopes defined by*

$$U_n(x)|_{[a_i,b_i]} = f(a_i) + \frac{P_{n,i} - f(a_i)}{f(b_i) - f(a_i)} \int_{a_i}^{x} f'(t)dt \quad (3.28)$$

where $P_{n,i}$ is defined as

$$P_k = \begin{cases} (1 - \frac{1}{n})|f(b)| & f(b) < 0 \\ (1 - \frac{1}{n}) \min(f(a_i), f(a_{i+1})) + \frac{1}{n} f(b_i) & f(b) \geq 0 \end{cases} \quad (3.29)$$

converges uniformly on E as $n \to \infty$ to a limit $U(x)$, and hence the EMD process converges in the sense that $m_k \to 0$ as $k \to \infty$.

Proof. Fix $\epsilon > 0$. let $U(x)$ denote the limit of the $U_n(x)$. Observe for $x \in [a_i, b_i]$ since $f(t)$ is monotone on that interval.

$$\left| \int_{a_i}^{x} f'(t)dt \right| = |f(x) - f(a_i)| \leq |f(b_i) - f(a_i)| \quad (3.30)$$

Therefore

$$\left|\frac{1}{f(b_i)-f(a_i)}\int_{a_i}^{x} f'(t)dt\right| \leq 1 \tag{3.31}$$

If $f(b_i) < 0$

$$|U_n(x)-U(x)| = \left|\frac{(1-(1/n))|f(b)|-f(a_i)}{f(b_i)-f(a_i)}\int_{a_i}^{x} f'(t)dt - \frac{|f(b)|-f(a_i)}{f(b_i)-f(a_i)}\int_{a_i}^{x} f'(t)dt\right| \tag{3.32}$$

$$= \left|\frac{\frac{1}{n}|f(b_i)|}{f(b_i)-f(a_i)}\int_{a_i}^{x} f'(t)dt\right| \leq \left|\frac{1}{n}|f(b_i)|\right| \leq \frac{M}{n}$$

This guarantees convergence of the sifting process in terms of the third definition of IMF. □

3.4 THE HILBERT TRANSFORM OF AN INTRINSIC MODE FUNCTION

This section discusses the Hilbert Transform; first it is proven the Hilbert Transform maps harmonic conjugate functions to one another. The section starts with the simplest harmonic conjugates,

Lemma 3.10. *The Hilbert Transform of* $\cos(ax)$ *is* $\sin(ax)$*, and the transform of* $\sin(ax)$ *is* $-\cos(ax)$*.*

Proof. In order to show this result, Hilbert Transform of cosine is calculated directly. These calculations draw heavily on the fact that sine is odd and cosine is even.

$$\begin{aligned}
&\lim_{\epsilon\to 0}\int_{\epsilon}^{\infty}\frac{\cos(t+ax)-\cos(t-ax)}{\pi t}dt \qquad (3.33)\\
&= \lim_{\epsilon\to 0}\int_{\epsilon}^{\infty}\frac{\cos(t)\cos(ax)-\sin(ax)\sin(t)-\cos(t)\cos(-ax)-\sin(t)\sin(ax)}{\pi t}dt\\
&= \frac{-2\sin(ax)}{\pi}\lim_{\epsilon\to 0}\int_{\epsilon}^{\infty}\frac{\sin(t)}{t}dt\\
&= \frac{-2\sin(ax)}{\pi}\mathrm{Si}(0) = \frac{-2}{\pi}\sin(ax)\frac{-\pi}{2}\\
&= \sin(ax)
\end{aligned}$$

In the above equations, Si denotes the sine integral. Using the same technique on the Hilbert Transform of cosine is tempting, However, using the same trick with the sum of angles for sine gives an expression which is not defined. Instead, the relation between sine and cosine gets the desired result from the above. Specifically, setting $ay = ax - \pi/2$, $\sin(x) = \cos(y)$. Using the above result,

$$H[\sin(ax)] = H[\cos(ay)] = \sin(ay) = -\cos(ax) \tag{3.34}$$

as needed. This completes the proof.

□

We will now move on to show the relation between the real and complex parts of a function that is analytic in the upper half of the complex plane which decays faster than $1/|z|$. This is known in physics and engineering as the Kramers–Kronig relations [38].

Theorem 3.11 (Kramers-Kronig Relation)**.** *Let $\phi : U \to \mathbb{C}$ be analytic on the upper half plane in $\mathbb{C}$. Then the The real and complex parts of ϕ are related by the Hilbert Transform. That is, if $\phi(z) = \phi_1(z) + i\phi_2(z)$, then $H[\phi_1] = \phi_2$ and $H[\phi_2] = -\phi_1$.*

Proof. Fix $\epsilon > 0$. Let t be a real number. First, note that

$$\oint \frac{\phi(z)}{z - t} = 0 \tag{3.35}$$

for each closed, rectifiable path in the upper half of the complex plane not containing t. Consider the path from a point $R > 0$ on the real line, tracing a semicircle counterclockwise through the upper half of the complex plane to $-R$, and then along the real axis. The path then jump over t and traces a semicircle of radius ϵ, and then returns to the real line to trace its way back to R.

Now break this path into four parts; the first, γ_1, is the section on the real line from $-R$ to $t - \epsilon$.

The second, γ_2, is the part on the real line from $t+\epsilon$ to R. The third, γ_3, is the small semicircle centered at t, and the last, γ_4, is the large semicircle meeting the real line at R and $-R$.

Consider γ_1 and γ_2 together. For $z = x + iy$, in this case, $y = 0$, and one can write

$$\int_{\gamma_1} \frac{\phi(z)}{z-t}dz + \int_{\gamma_2} \frac{\phi(z)}{z-t}dz = \int_{-R}^{t-\epsilon} \frac{\phi(x)}{x-t}dx + \int_{t+\epsilon}^{R} \frac{\phi(x)}{x-t}dx \tag{3.36}$$

Now taking the limit as $R \to \infty$,

$$\begin{aligned} &\lim_{\epsilon\to 0^+} \lim_{R\to\infty} \int_{-R}^{t-\epsilon} \frac{\phi(x)}{x-t}dx + \int_{t+\epsilon}^{R} \frac{\phi(x)}{x-t}dx \\ &= PV \int_{-R}^{R} \frac{\phi(x)}{x-t}dx \end{aligned} \tag{3.37}$$

Now it remains to evaluate the other pieces of the path under the correct limit. Next consider γ_3

$$\begin{aligned} \int_{\gamma_3} \frac{\phi(x)}{x-t}dx &= \lim_{\epsilon\to 0^+} \int_0^{\pi} \frac{\phi(t+\epsilon e^{i\omega})}{t+\epsilon e^{i\omega}-t} i\omega\epsilon e^{i\omega} d\omega \\ &= \lim_{\epsilon\to 0^+} \int_0^{\pi} \phi(t+\epsilon e^{i\omega}) i\omega d\omega \\ &= \int_0^{\pi} \lim_{\epsilon\to 0^+} \phi(t+\epsilon e^{i\omega}) i\omega d\omega \\ &= i\pi\phi(t) \end{aligned} \tag{3.38}$$

Now taking the limit as $\epsilon \to 0$ under the integral sign can be justified by the Dominated convergence theorem; as the domain is compact and the function continuous, it must have a maximum modulus to serve as the dominating function.

Which gives the needed result. Lastly for γ_4,

$$\begin{aligned} \int_{\gamma_4} \frac{\phi(x)}{x-t} &= \int_0^{\pi} \frac{\phi(Re^{i\omega})}{Re^{i\omega}-t} i\omega Re^{i\omega} d\omega \\ &= \int_0^{\pi} \frac{\phi(Re^{i\omega})}{1-\frac{t}{Re^{i\omega}}} i\omega d\omega \end{aligned} \tag{3.39}$$

Now recalling that ϕ is assumed to have faster than linear decay as $R \to \infty$, we see that this integral tends to zero as R becomes large.

This implies

$$
\begin{aligned}
0 = \oint \frac{\phi(z)}{z-t} &= PV \int_{-\infty}^{\infty} \frac{\phi(x)}{x-t} dx - i\pi\phi(t) \qquad (3.40) \\
\therefore \ \phi(t) &= \frac{i}{\pi} PV \int_{-\infty}^{\infty} \frac{\phi(x)}{x-t} dx \\
\therefore \ \phi_1(t) &= \frac{-1}{\pi} PV \int_{-\infty}^{\infty} \frac{\phi_2(x)}{x-t} dx \\
\phi_2(t) &= \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{\phi_1(x)}{x-t} dx
\end{aligned}
$$

Now splitting the real and imaginary components gives the desired result. All that remains is to observe that the final lines of the preceding calculations are exactly the Hilbert Transform, as needed. □

This relationship has a physical interpretation as well. Interpreting $\phi(t)$ as an impulse response function, the conditions on the function also make sense. First, that ϕ is analytic in the upper half plane is the same as requiring ϕ defines a causal system (i.e. forces only acting on particles after the force is applied). The decay condition corresponds, then, to the requirement that the effect of a given impulse must decrease over time (i.e. a force on a pendulum will cause the pendulum to swing, but eventually the system will come to rest).

CHAPTER 4

AIR LEAK IDENTIFICATION IN SPACE MODULES

4.1 INTRODUCTION

In this chapter some applications of the Fourier Transform, Hilbert-Huang Transform and Artificial Neural Networks to air leak identification in pressurized space modules are discussed. First, data was collected using an experimental setup described in the following section. Data analysis and pre-processing is discussed following the data collection. The accuracy of FFT, HHT, and ANN based approaches are then compared.

4.2 EXPERIMENTAL SETUP

The first goal was to simulate an air leak in a space module. This required placing the ultrasonic recording equipment on the high pressure side of the leak aperture and recording the acoustics.

The data was collected from simulated leaks in a steel autoclave, as shown in 4.1a. The leak apertures were created using a half-inch end cap on an exhaust port of the autoclave. Aluminum, plastic and steel exhaust ports were used. Leaks in rubber were tested by securing a piece of rubber to the exhaust port using a steel end cap with a large hole. Small (0.032 inch), medium (0.1 inch), and large (0.25 inch) holes were drilled in the end caps at the University of Maine Advanced

Leak Size	Material	Pressure Gradient (psi)
None	Aluminum	+15
0.032	Steel	+12
0.10	Rubber	+9
0.25	Plastic	

Table 4.1: Test Conditions.

(a) An Autoclave

(b) Recording Equipment

Figure 4.1: Data collection equipment including a steel autoclave (a), oscilloscope, DC power supply and laptop computer (b).

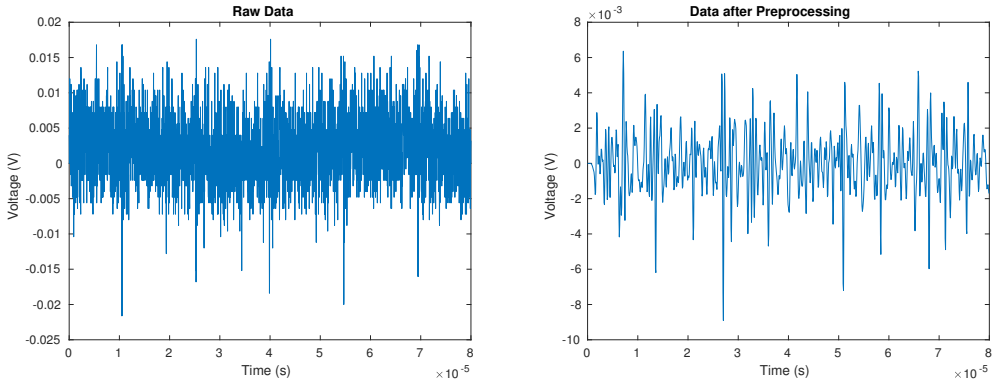
Manufacturing Center. The autoclave was pressurized by connecting it to a 80 PSI pneumatic line. Once the pressure reached 15, 12, or 9 PSI the high pressure line was disconnected and the audio signal was recorded. A table of the test conditions can be found in Table 4.1.

The acoustics were recorded using a Knowles High-frequency microphone, wired directly to an Agilent oscilloscope. The data was then transfered via Universal Serial Bus (USB) to a laptop. The data was then analyzed using the MATLAB programming suite. A survey of the recording equipment can be found in 4.1b

4.3 DATA ANALYSIS

After the data was collected, it was analyzed in MATLAB. The data from the oscilloscope comes in 4000 sample time series. A sample of the base data can be seen in Figure 4.2a.

Once the data was collected, it was downsampled by a factor of four, to create a 1000 sample time series. Due to the high sampling rate on the oscilloscope, it was necessary to reduce the effective sampling rate to have a good resolution on an



(a) Unfiltered Data (b) Filtered Data

Figure 4.2: A sample time series of ultrasonic data before and after filtering.

FFT. The downsampling was done using a moving average scheme to reduce some shot noise present in the data (occasional, very high or very low samples).

After this, the data was filtered with a frequency dependent FIR filter designed in MATLAB. The filtering was done in two stages, one high pass and one low pass filter. The filters were designed to isolate the frequency band from 30 to 55 kHz, as defined as critical [11]; emphasis was given to getting a good frequency response rather than computational efficiency.

Once the data was prepared, the sound spectrum was analyzed to create a test statistic for differentiating among the different conditions. The determination between the material is done with a classic algorithm from computer science, the "king of the hill".

For a batch of data to test, a summary statistic is calculated the same way as for the experimental data set. The test mean from our new data is compared to the old data sets' means. This is one at a time, and the one that is closest is the 'king of the hill,' or our best guess for the leak material. A block diagram of the process can be found in Figure 4.5.

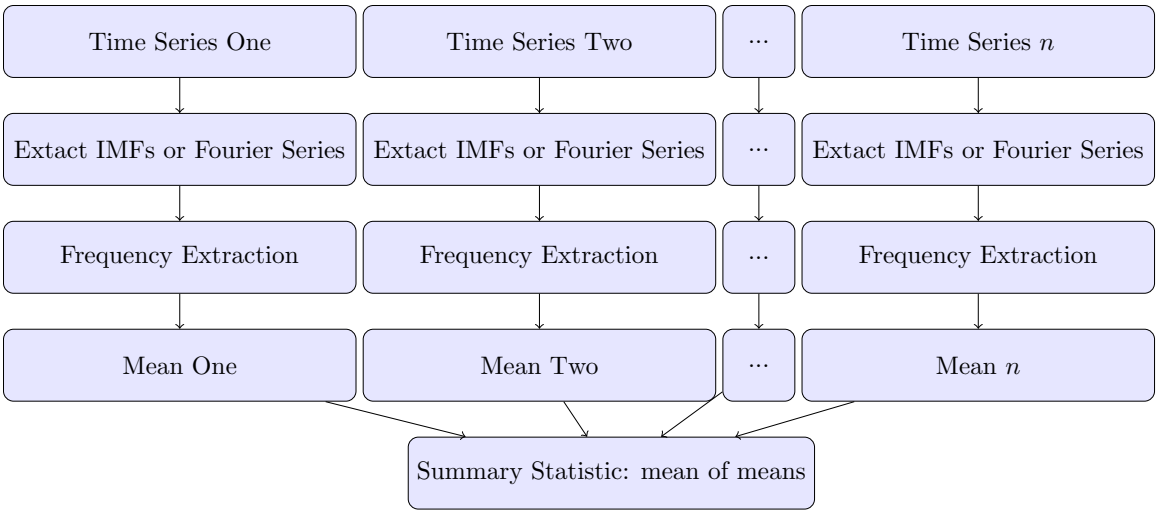


Figure 4.3: Generation of the summary statistics for the times series using varied transforms.

4.3.1 FOURIER ANALYSIS

A Fourier Transform of the recorded data yields a distribution of frequency components shown in Figure 4.4

Next To identify the material where the leak occured, one must create a test statistic. A natural candidate is the mean frequency component with the highest energy. In terms of the basis of $L^2(E)$, we consider the coefficient of the $e_n(t)$ which has the highest magnitude. The magnitude $|a_n|$ is also considered.

For each Fourier Series, select the frequency component with the highest magnitude. Then average the highest energy component from each of the Fourier series from that material. In this way, one is able to generate a summary statistic from the Fourier Transformed data. Another approach used was calculating the mean (or, really, first central moment) of the Fourier spectrum. Both offered similar results in practice. This summary statistic $\bar{M}_k$, where k is the index of the data set (the kth 1000 sample time series).

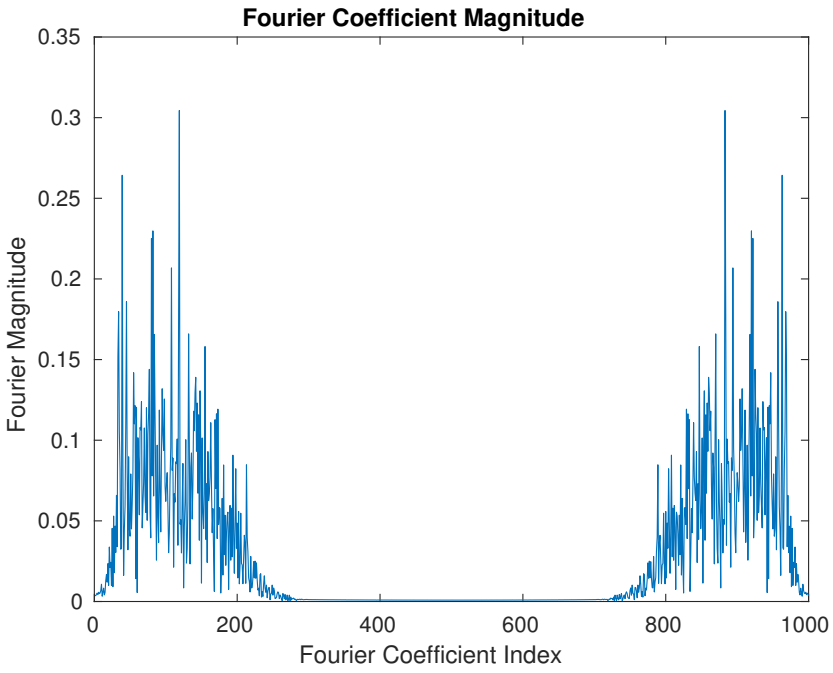


Figure 4.4: A sample Fourier Series of the ultrasonic acoustic data.

4.3.2 HILBERT-HUANG SPECTRAL ANALYSIS

A similar process for the Hilbert-Huang Transform, except that an HHT is used on place of an FFT. This gives us many different instantaneous frequencies, one for each IMF.The average instantaneous frequency for each IMF is calculated. This is denoted $\bar{\theta_k}$.

The data is again aggregated as in 4.3.1, except with $\bar{\theta_k}$ instead of $\bar{M_k}$. This allows for a similar process comparing data.

Comparing the $\bar{\theta_k}$ has a significant advantage over other methods; as we compare a mean of means, test statistic will converge to normal for large sample sizes. For this reason, the model is asymptotically accurate, so long as the test set and initial data set have arbitrarily large sample sizes. While this is not practical, a very large initial data set increases accuracy. Using such a data set, this becomes a simple mean estimation problem to solve, rather than a more difficult type identification problem.

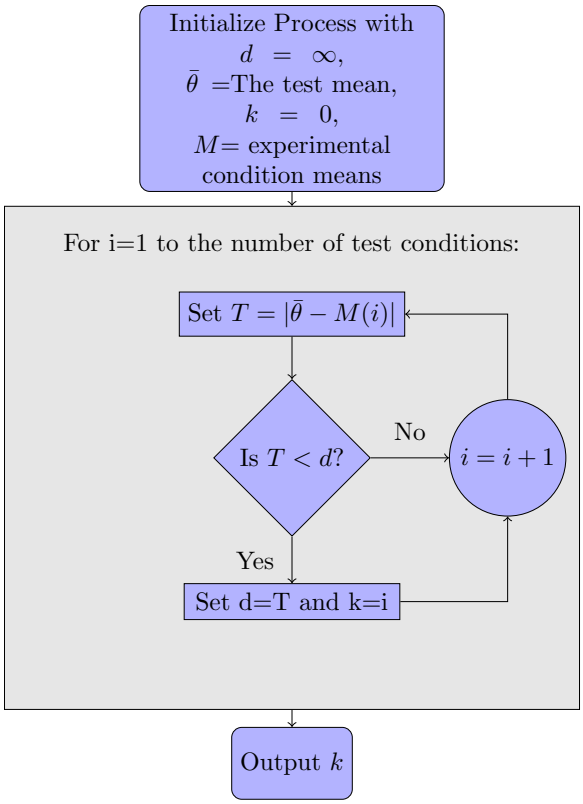


Figure 4.5: A block diagram of the material estimation process.

4.3.3 NEURAL NETWORK

Another approach uses the entire data set for the analysis through the use of a convolutional neural network. This method used the same data that was used for the other analysis to train a neural network using a gradient decent training scheme. The network was developed using MATLAB's nprtool. The same preprocessing scheme was used as in previous analysis.

Neural networks were trained using both the preprocessed time series and its Fourier Transform. The training data sets included about 2500 data vectors of one thousand samples. The network was trained with 20 layers; this means consists of a sigmoid transfer function on the hidden layers and softmax transfer functions on the output layers. The training was done with a Scaled Conjugate Gradient scheme.

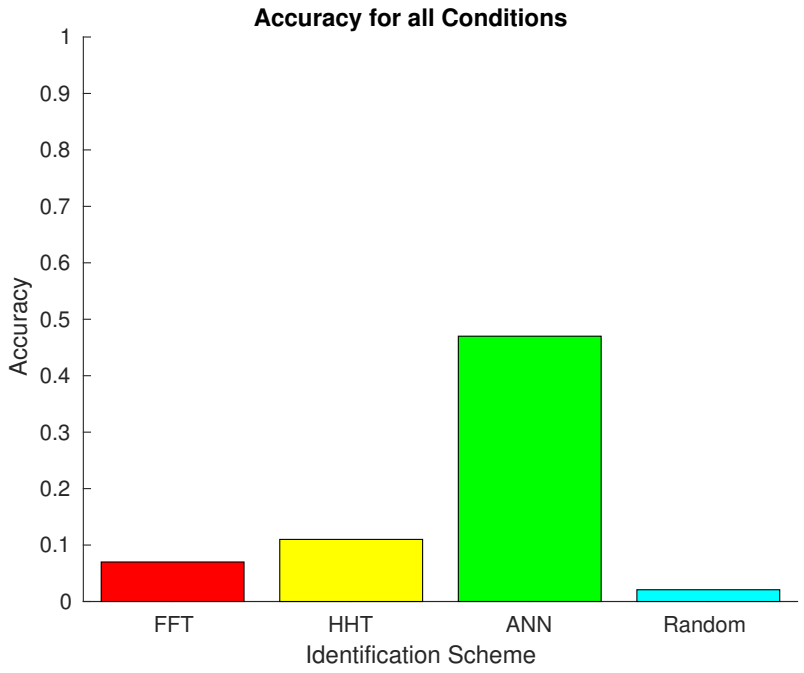


Figure 4.6: A comparison of the accuracy of the FFT, HHT and ANN methods for determining among all test conditions.

4.4 RESULTS

The last thing to do is to compare the various methods suggested to determine their efficacy. To do this, we compare the relative accuracy of each proposed method to determine the best one. To do this,a test sample from each condition is selected and algorithm is used to determine the test condition.

A graph of the accuracies is shown in Figures 4.6 - 4.9. The best method for selecting the conditions overall was the HHT-based method. This method requires the most computational resources, making it difficult to use in practice. This methodology is useful for determining the leak condition, however

A comparison of the methods suggests that the cognitive neural network is actually the most effective at identifying the material where the leak occurs. However, once the number of output classes for the ANN grow too large, the

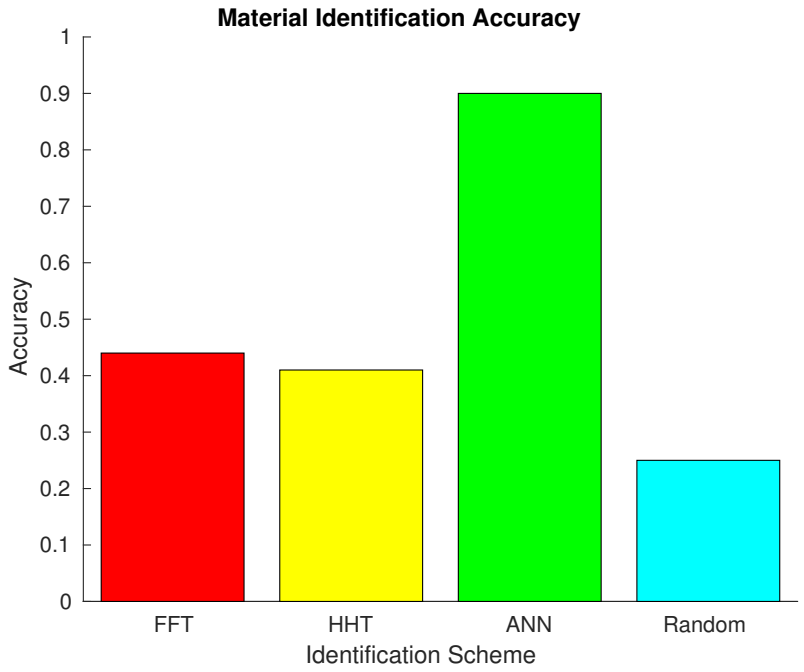


Figure 4.7: A comparison of the accuracy of the FFT, HHT and ANN methods for determining among materials.

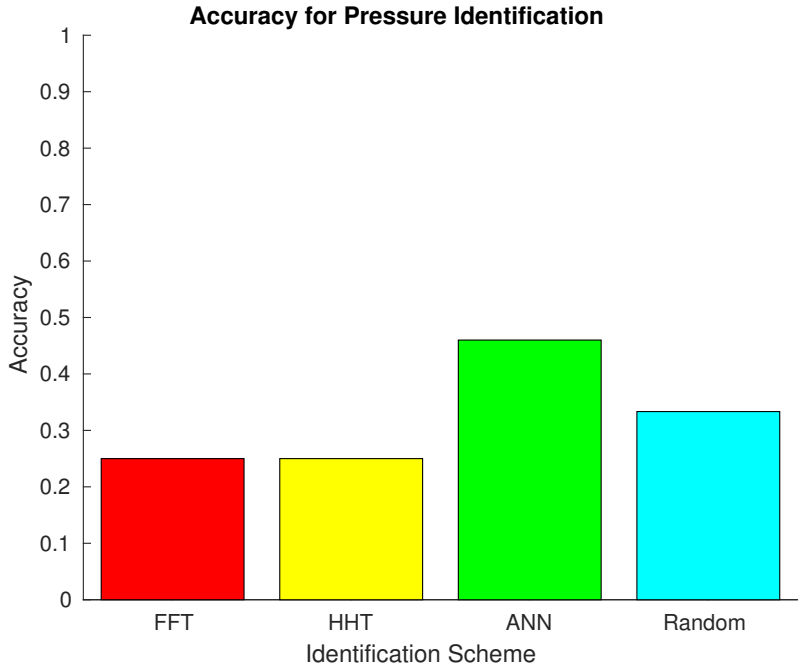


Figure 4.8: A comparison of the accuracy of the FFT, HHT and ANN methods for determining among pressure levels.

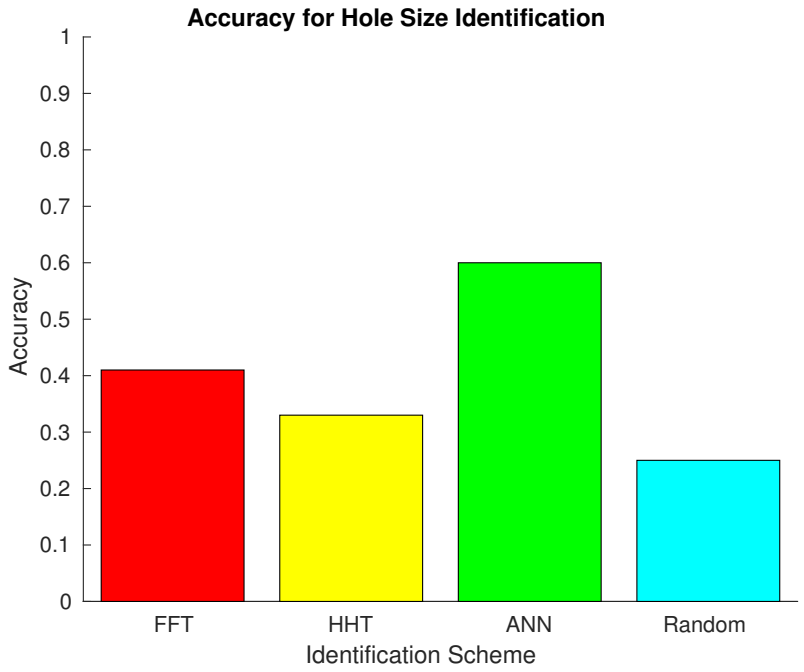


Figure 4.9: A comparison of the accuracy of the FFT, HHT and ANN methods for determining among hole sizes.

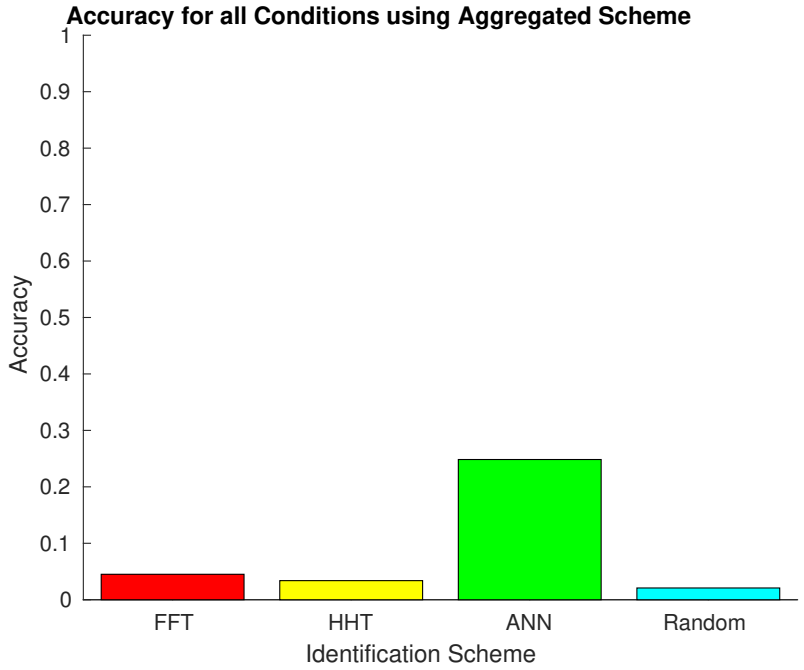


Figure 4.10: A comparison of the accuracy of the FFT, HHT and ANN methods for determining among all test conditions using aggregated single-factor identification.

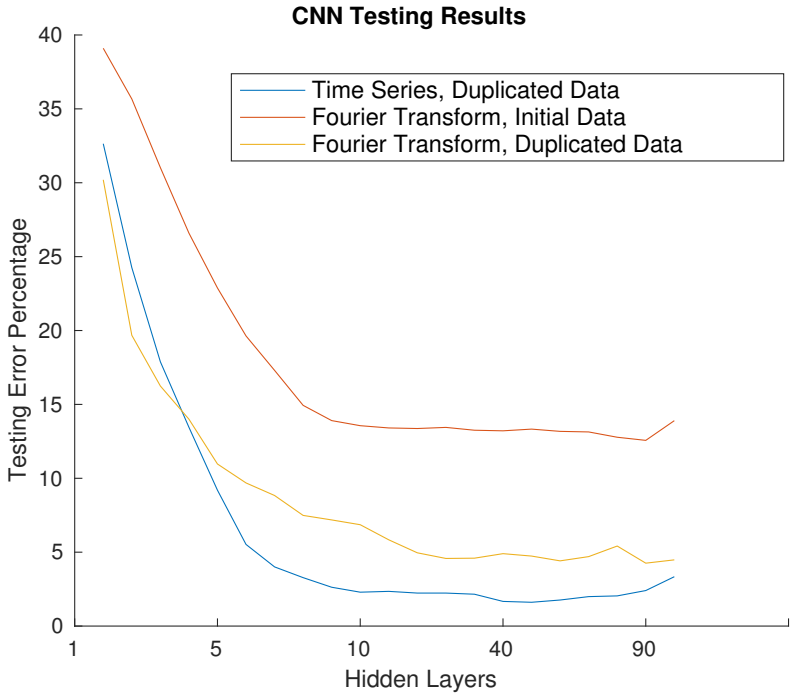


Figure 4.11: The performance of an ANN in material identification for different training data sets and varied numbers of hidden layers.

accuracy is greatly diminished; with only four output classes the ANN method offers 90% accuracy, while the full test conditions offers only 45% accuracy.

The performance of an ANN can also vary based on its architecture. Specifically, changing the number of hiddern layers can alter performance. In order to determine the best ANN for material identification, various numbers of hidden layers were used. This can be seen in Figure 4.11. Little improvement is seen after ten layers, and increasing past twenty offers virtually no improvement. Due to the increasing computational complexity, it is not recommended to use more than 25 layers for material identification.

The methods were much more effective in identifying the materials than the pressure level or the hole size. It is possible that the pressure level does not have much effect on the spectrum of the data. This means that none of the methods are particularly suited to identifying the pressure used to simulate a leak. The accuracy

when identifying the hole size alone was between the accuracy of determining the pressure and the overall accuracy.

It is perplexing that the overall accuracy can be higher than the accuracy when determining only one condition. This may be explained by high within groups variance that was not present without the clustering. This is to say that grouping the data in an unnatural way increases the variance rather than decreases it.

These results are not unexpected. The cognitive neural network leverages the entire dataset in a way that the other methods do not; instead the FFT and HHT methods use a summary statistic rather than conditioning on the entire dataset. The accuracy of these methods is only around 75%. For this reason, it is preferable to use the ANN approach, which offers 90% accuracy.

The FFT is undoubtedly the most suspect method; the FFT offers poor performance for turbulence. Since the characteristics of the signal are not constant over time, the HHT is a preferable method. However, the ANN is the most efficient method. The accuracy of the neural network approach is the greatest; only one time series beats the highest accuracy of the HHT approach. Using multiple time series for the test would allow even better estimation; that is using a best-of-three or best-of-five approach with the neural network. This would also allow the basic time series to be used, allowing a neural network to be placed on a smaller processor. In real time, this would allow for the identification of the leak in a fraction of a second.

We can model the accuracy of such as system using Bernoulli trials. Assume that the probability of correctly identifying a leak material from a give time series is p. Our accuracy for a best of $2n+1$ method using a neural network is estimated as the probability of getting at least half of the identifications correct. This can be calculated as:

$$P(x > n) = \sum_{k=n+1}^{2n+1} \binom{2n+1}{k} p^k (1-p)^{2n-k+1} \tag{4.1}$$

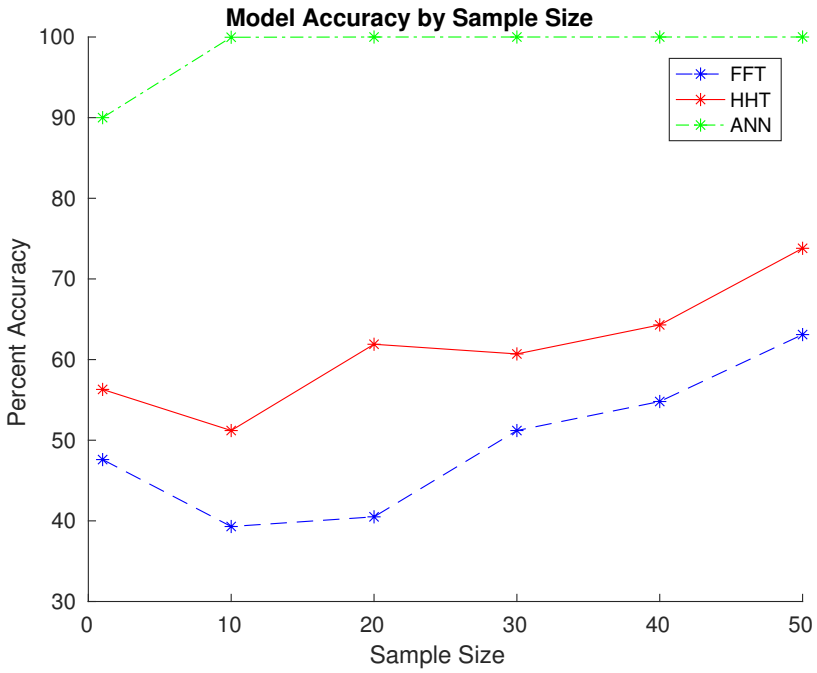


Figure 4.12: The accuracy of different methods using varied sample sizes.

For our material identification ANN, this offers excellent results. Plotting this again the HHT and FFT methods in Figure 4.12 shows the excellent accuracy of this system. The accuracy can be see to rapidly approach one. By 11 trials, the accuracy is estimated to be 99.97%. Naturally this is a much slower process. However, due to the high sampling rate of the system used, this corresponds to only 0.028 seconds of recording time.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

5.1 SUMMARY OF CONTRIBUTIONS

The thesis began by discussing issues with leak detection in pressurized space modules. Air leaks were simulated using a steel pressure chamber, and recorded the audio signals. We compared a number of methods for analyzing the leak acoustics and determining the leak source.

Three main methods were used in analyzing the leak spectra. First,a Fourier Transform was used to select frequency band with the highest energy, and characterize the leaks using this. This method however was not very effective.

From here, a better frequency analysis method for turbulent flow, the Hilbert-Huang Transform, was used. This method was more effective, with about 80% accuracy. However, this methodology is poorly understood, and this lead to many mathematical questions that are only partially answered.

A methodology that required a measure of center for characterizing the leaks may offer poor performance in determining the leak source. Instead, it is preferable to leverage the entire data set some how. We do this using a pattern-recognition neural network. This method obtained a 90% accuracy for identifying the leak material. Preliminary estimates suggest the accuracy of this method could increase to 98% using more data. This work has been the basis of the following published works:

1. Kenneth R. Bundy, Lonnie Labonte, Casey Clark, and Ali Abedi, "Collection and Analysis of Leak Spectral Signatures for Application to the ISS," IEEE Conference on Wireless Sensors in Extreme Environments, 2015, Orlando, FL [3]

2. Kenneth R. Bundy, Chitra Manjanai Pandian, Ali Abedi, and Vincent Caccesse, "Analysis of Leak Spectral Signatures in Pressurized Space Modules," IEEE Conference on Wireless Sensors in Extreme Environments 2016, Aachen, Germany [4]

3. Kenneth R. Bundy, Ali Abedi, "Air Leak Material Identification in Pressurized Space Vehicles using a Convolutional Neural Network," IEEE Conference on Wireless Sensors in Extreme Environments 2017, Toronto, Canada[5]

In our mathematical analysis of the Hilbert-Huang Transform, the convergence of the Hilbert-Huang Transform's Empirical Mode Decomposition has been demonstrated for some envelope functions. This convergence is uniform, but with zero as the limit function. Using a stopping criterion, is sufficient to express the function in terms of finitely many IMFs.

5.2 FUTURE WORK

We have proved some new results for providing a theoretical framework for the HHT, but there is much more to be done in this area. For example, the use of tight envelopes is required in the proofs given. Most likely this requirement can be removed, but this remains a conjecture.

In conclusion, both leak source identification and some theoretical aspects of the Hilbert-Huang Transform have been addressed. However, there is still much to be done on both fronts. Generalizing the proof of convergence of the transform, for instance, is a needed piece. Also the best envelope function has not been determined, which remains the largest problem in practice.

In the leak identification problem, a larger data data set for the neural network approach is needed. A new approach, or better data set, may offer good results. Furthermore, it would be useful to perform experiments with multiple leaks and different hole shapes. It is not known what effect the geometry of the leak aperture

has on the leak sound spectrum. This, and multiple hole conditions, should be tested.

BIBLIOGRAPHY

[1] N. C. C. W. C. Wilson and E. I. Madaras, "Leak detection and location technology assessment for aerospace applications," *National Aeronautics and Space Administration*, 2008.

[2] C. Clark, "Wireless leak detection using airborne ultrasonics with technology demonstration on the international space station," 2016, Master's Thesis.

[3] K. R. Bundy, L. Labonte, C. Clark, and A. Abedi, "Collection and analysis of leak spectral signatures for application to the iss," in *2015 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*, Dec. 2015, pp. 1–3. DOI: 10.1109/WiSEE.2015.7393106.

[4] K. R. Bundy, C. M. Pandian, A. Abedi, and V. Caccese, "Analysis of leak spectral signatures in pressurized space modules," in *2016 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*, Sep. 2016, pp. 54–58. DOI: 10.1109/WiSEE.2016.7877303.

[5] K. R. Bundy and A. Abedi, "Air leak material identification in pressurized space vehicles using a convolutional neural network," in *2017 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*, Oct. 2017, pp. 150–152. DOI: 10.1109/WiSEE.2017.8124909.

[6] S. Writers, "The crew that never came home: The misfortunes of soyuz 11," *Space Safety Magazine*, 2014.

[7] L. B. J. S. C. Robert W. Fricke, "Sts-37 space shuttle mission report may 1991 -nasa-cr-193062," *Extravehicular Activity Evaluation, NASA*, 1991.

[8] J. Oberg, "Crew finds 'culprit' in space station leak," *MSNBC*, 2004.

[9] C. Bergin, "Iss concern over s1 radiator - may require replacement via shuttle mission," *NASASpaceflight.com*, 2009.

[10] X. Qiyue, L. Pingping, and C. Maolin, "Research on different features between direct and reflected ultrasonic signals from air leakage sources," in *2010 First International Conference on Pervasive Computing, Signal Processing and Applications*, Sep. 2010, pp. 508–512. DOI: 10.1109/PCSPA.2010.128.

[11] D. E. C. e. a. R.S. Reusser, "Spacecraft leak location using structure-borne noise," *University of Iowa Digital Repository*, 2009.

[12] J. Zhao, F. Sun, and H. Wang, "Pipeline leak fault feature extraction based on wavelet packet analysis and application," in *2011 International Conference on Electrical and Control Engineering*, Sep. 2011, pp. 1148–1151. DOI: 10.1109/ICECENG.2011.6057772.

[13] Y. Jia, B. Gao, C. Jiang, and S. Chen, "Leak diagnosis of gas transport pipelines based on hilbert-huang transform," in *Proceedings of 2012 International Conference on Measurement, Information and Control*, vol. 2, May 2012, pp. 614–617. DOI: 10.1109/MIC.2012.6273368.

[14] J. Steckel and H. Peremans, "Ultrasound-based air leak detection using a random microphone array and sparse representations," in *SENSORS, 2014 IEEE*, Nov. 2014, pp. 1026–1029. DOI: 10.1109/ICSENS.2014.6985178.

[15] C. Ai, X. Sun, H. Zhao, R. Ma, and X. Dong, "Pipeline damage and leak sound recognition based on hmm," in *2008 7th World Congress on Intelligent Control and Automation*, Jun. 2008, pp. 1940–1944. DOI: 10.1109/WCICA.2008.4593221.

[16] W. Sun, R. Yan, L. Sun, D. Meng, Z. Li, H. Guo, and W. Li, "Study of the space station on-orbit leak detection based on the differential pressure gas sensor," in *2013 IEEE International Conference on Green Computing and*

Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, Aug. 2013, pp. 1718–1721. DOI: 10.1109/GreenCom-iThings-CPSCom.2013.316.

[17] H. Roufarshbaf, J. Castro, and A. Abedi, "Stochastic modeling of leak detection and localization using ultrasonic sensor array," in *IEEE International Conference on Wireless for Space and Extreme Environments*, Nov. 2013, pp. 1–1. DOI: 10.1109/WiSEE.2013.6737573.

[18] C. Ai, X. Sun, H. Zhao, R. Ma, and X. Dong, "Pipeline damage and leak sound recognition based on hmm," in *2008 7th World Congress on Intelligent Control and Automation*, Jun. 2008, pp. 1940–1944. DOI: 10.1109/WCICA.2008.4593221.

[19] K. D. Champaigne and J. Sumners, "Wireless impact and leak detection and location systems for the iss and shuttle wing leading edge," in *2005 IEEE Aerospace Conference*, Mar. 2005, pp. 1–8. DOI: 10.1109/AERO.2005.1559538.

[20] W. B. Andreas Johansson and A. Medvedev, "Model-based gas leakage detection and isolation in a pressurized system via laguerre spectrum analysis," in *Proceedings of the 1998 IEEE International Conference on Control Application*, Trieste, Italy, 1998.

[21] C. Clark, L. Labonte, J. Castro, A. Abedi, and V. Caccese, "Wireless leak detection using airborne ultrasonics and a fast-bayesian tree search algorithm with technology demonstration on the iss," in *2015 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*, Dec. 2015, pp. 1–5. DOI: 10.1109/WiSEE.2015.7392983.

[22] A. P. A. Mohinder S. Grewal, *Kalman Filtering: Theory and Practice Using MATLAB*. Wiley, 2015.

[23] H. Royden and P. Fitzpatrick, *Real Analysis*. Pearson, 2010.

[24] I. Y. A. Oppenheim A. Willsky, *Signals and Systems, First Edition*. Prentice Hall, Englewood Cliffs, New Jersey, 1983.

[25] E. Titchmarsh, *Introduction to the Theory of Fourier Integrals*. Oxford University Press, 1948, pp. 121–143.

[26] N. E. Huang, *Introduction to the Hilbert-Huang Transform and its related mathematical problems*. Greenbelt, MD: Goddard Institute for Data Analysis.

[27] B. N. E. H. et al., "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society*, 1996.

[28] R. Kress, *Linear Integral Equations*. Springer-Verlag, 1989, p. 91.

[29] J. Conway, *Functions of One Complex Variable*. Springer, 2003.

[30] E. Bedrosian, "A product theorem for hilbert transforms," *RAND Corporation*, 1962.

[31] D. Graupe, *Principles of Artificial Neural Networks*. World Scientific Publishing Co. Pte. Ltd., 2014.

[32] P. Lorentz, *Artificial Neural Systems: Principles and Practice*. Bentham Science Publishers, 2015.

[33] I. Kipyatkova and A. Karpov, "Recurrent neural network-based language modeling for an automatic russian speech recognition system," in *2015 Artificial Intelligence and Natural Language and Information Extraction, Social Media and Web Search FRUCT Conference (AINL-ISMW FRUCT)*, Nov. 2015, pp. 33–38. DOI: 10.1109/AINL-ISMW-FRUCT.2015.7382966.

[34] R. C. Sharpley and V. Vatchev, "Analysis of the intrinsic mode functions," *Constructive Approximation*, 2006.

[35] P. T. Thomas Y. Hou Zuoqiang Shi, "Sparse time frequency representations and dynamical systems," *Communications in Mathematical Sciences*, pp. 673–694, 2013, ISSN: 1539-6746.

[36] G. P. M.C. Peel and T. McMahon, "Empirical mode decomposition: Improvement and application," 2014.

[37] K. L. Zhengguang Xu Benxiong Huang, "An alternative envelope approach for empirical mode decomposition," *Elsevier*, 2009.

[38] R. G. Brown, "Kramers-kronig relations," *Duke University*, 2007.

BIOGRAPHY OF THE AUTHOR

Kenneth R. Bundy was born in Minot, Maine. He studied first at Minot Consolidated School, then Poland Regional High School, where he graduated in 2008.

He began his undergraduate studies in the Honors College at the University of Maine the same year. In 2012 he was working on his Honors' Thesis when his advisor, Dr. Ali Ozluk, passed away unexpectedly. Ken returned home to Minot to complete a new Honors Thesis with the help of Dr. George Markowski and Dr. Ali Abedi.

While working on his first thesis, he began working at the Poland Spring Inn on third shift. He was promoted to manager the following year, and continued to work there until he return to the greater Orono area to complete his Bachelor's degree. After defending his Honors Thesis calculating the channel capacity of interplanetary data transmission using the Voyager and New Horizons spacecraft specifications, he completed his Bachelor of Arts in Mathematics in 2014 with a concentration in statistics.

After returning to Orono, he worked at Nexxlinx, Inc. providing technical support for the OpenTable Internet-based reservation service. He began his master's studies at the University of Maine in 2014, studying Electrical Engineering. He later returned to the mathematics department to complete his degree.

At the time of printing, he lives in Maine with his wife Ayesha. He is a member of the American Mathematical Society (AMS) and Institute of Electrical and Electronics Engineers (IEEE), and has published three papers in IEEE conferences. These papers are as follows:

1. Kenneth R. Bundy, Lonnie Labonte, Casey Clark, and Ali Abedi, "Collection and Analysis of Leak Spectral Signatures for Application to the ISS," IEEE

Conference on Wireless Sensors in Extreme Environments, 2015, Orlando, FL [3]

2. Kenneth R. Bundy, Chitra Manjanai Pandian, Ali Abedi, and Vincent Caccesse, "Analysis of Leak Spectral Signatures in Pressurized Space Modules," IEEE Conference on Wireless Sensors in Extreme Environments 2016, Aachen, Germany [4]

3. Kenneth R. Bundy, Ali Abedi, "Air Leak Material Identification in Pressurized Space Vehicles using a Convolutional Neural Network," IEEE Conference on Wireless Sensors in Extreme Environments 2017, Toronto, Canada[5]

Kenneth Richard Bundy is a candidate for the Master of Arts degree in Mathematics from The University of Maine in August 2018.