UNIVERSITY OF BERGEN

Department of Natural Science and Mathematics

MASTERS THESIS

Application of Hilbert-Huang transform in Bearing Fault Detection

Author: Yapi Donatien Achou

Supervisor: Jarle Berntsen

February 20, 2020

Imagination is more important than knowledge. Knowledge is limited, but imagination encircles the world.

Albert Einstein

Abstract

Since the industrial revolution, machine of all sorts without doubt, have been, and still are the driving force of the world economy. Operating in harsh conditions, machines are prone to failure, which can incur financial losses, the degradation of the environment, and sometime human casualties.

One of the main component of nearly all machines are bearings. Due to their geometrical characteristics, the latter facilitate rotation movements while being continuously subjected to large loads . Consequently, bearing failures represent more than 40% of machine breakdown. To mitigate bearing degradation, they are monitor in order to detect incipient faults as early as required. This is achieved by mounting sensors on strategics area surrounding a bearing, in order to measure, collect and analyze its vibration movement (vibration signal).

The analysis processes for bearing fault detection in most industrial applications, rely on the Fourier transform. This is due to the fact that bearings faults occur at specific frequencies, called characteristic defect frequencies. A defect in a bearing will emit a periodic high frequency pulse as the bearing rotates. The frequency at which the pulse is emitted is the defect frequency. As a result, the core of bearing monitoring, is the early detection of onset of failure, expressed by the defect frequencies. The latter can be found in the frequency spectrum, derived by the Fourier transform.

Bearings are in general mounted on machines comprising several components, each emitting its own vibration signal. This render the detection of defect frequencies in a bearing frequency spectrum challenging, due to interfering signals, and induced noise. Therefore, the frequency spectrum obtained directly from a bearing vibration signal, contains little diagnostic information. To circumvent the latter issue, a series of signal filtering operations and mathematical transformations are applied to a bearing vibration signal. The goal is to isolate and expose the relevant signal, that contains the defects frequencies. This process can be achieved through various methods. However, the most prevalent scheme is the so called high frequency resonance technique (HFRT). The HFRT removes or dump irrelevant signals, and extract relevant one, be-

fore applying the Fourier transform. This result in a bearing frequency spectrum, that reveals the defect frequencies.

Although efficient and widely used, the high frequency resonance technique can generate a noisy spectrum, in particular when bearing defects are severe. This can render bearings faults detection challenging. To circumvent the above challenge, this thesis posit a new scheme, that relies on a set of filtering techniques, in order to generate a noiseless spectrum. To test the validity of the proposed new method, it is applied to a case study. The results derived from this test, proves that the new scheme is able to reveal bearing defect frequencies, in a relatively noiseless spectrum, as opposed to the HFRT.

Acknowledgment

I am profoundly gratefully towards my supervisor professor jarle Berntsen, for the supports he had given me. His ability to see the bigger picture, coupled with a great sense of detail, was instrumental and necessary in the process of writing this thesis. I am also thankful for his life advice pertaining to health and good living in general. I want to also thank my colleague Dr Ru Yan with whom I had fruitful discussions, that was beneficial to this thesis.

Y.D.A

Contents

List of Figures

List of abbreviations

ACF Autocovarience Function

COV Covariance

E Expectation value

BPFI Ball Pass Frequency Inner race defect

BPFO Ball Pass Frequency Outer race defect

BSF Ball Spin Frequency

DP D'Agostino Pearson

EMD Empirical Mode Decomposition

FFT Fast Fourier Transform

FT Fourier Transform

FTF Fundamental Train frequency

HHT Hilbert-Huang Transform

HFRT High Frequency Resonance Technique

IMF Intrinsic Mode Function

LOESS Locally Estimated Scatter Smoother

NASA National Aeronautics and Space Administration

PSD Power Spectral Density

STFT Short Time Fourier Transform

STL Seasonal Trend decomposition based on LOESS

Chapter 1

Introduction

1.1 Background

The associate press, a leading news organization, wrote: “Norwegian probe: Gearbox failure caused fatal 2016 crash”. It was a news report of an airbus helicopter crash, on a small island outside of Bergen, the second largest city in Norway, cutting short the life of 13 people. The official cause of the crash: “A fatigue fracture in the main rotor gearbox.”. The anticipated question that arises after such catastrophic event is: was it preventable?

The latter question has been the driving force, behind the adoption of what is commonly known as predictive maintenance. Since, machines are prone to failure, they must be monitored and maintained regularly, to avoid catastrophic breakdown. The set of methods and strategies used to monitor, detect or predict onset of failure, and plan maintenance, are grouped under the umbrella term predictive maintenance.

Formally, predictive maintenance for machines and industrial equipment can be defined as a maintenance philosophy or more generally a framework, with a set of standards and methods, used to predict and prevent machine failure. This maintenance philosophy, when correctly implemented, increases machine life time, reduces downtime and maintenance cost. The aim here, is to detect as early as possible incipient failure and take appropriate actions.

Most industrial machines or equipment, rely on one or more rotating component(s) such as a gear box or a shaft. These rotating components, are sometime coupled with bearings in order to facilitate their rotation. In rotating machines, more than 40% of failure can be attributed to bearing faults (Albrecht et al. (1986)), as shown in figure 1.1, which displays the distribution of faults in rotating machines. By facilitating rotation movements, bearings can be subjected to large load and mechanical forces, which can lead to slowly propagating defects.

Most commonly applied methods in predictive maintenance for bearing fault detection rely on the Fourier transform. Here is why: To monitor bearing for eventual failures, one or more sensors are used to record their vibrating movement periodically. Each record or sample is called a bearing vibration signal. Because of their geometry, bearings failures are associated to specifics frequencies call defect frequencies, derived from each bearing type geometrical properties. Furthermore, the core of bearings monitoring, rest on the identification of the failure frequencies and their secondary effects. This is accomplished by Fourier transforming a target vibration signal into its corresponding frequency spectrum, and searching for eventual failure frequencies.

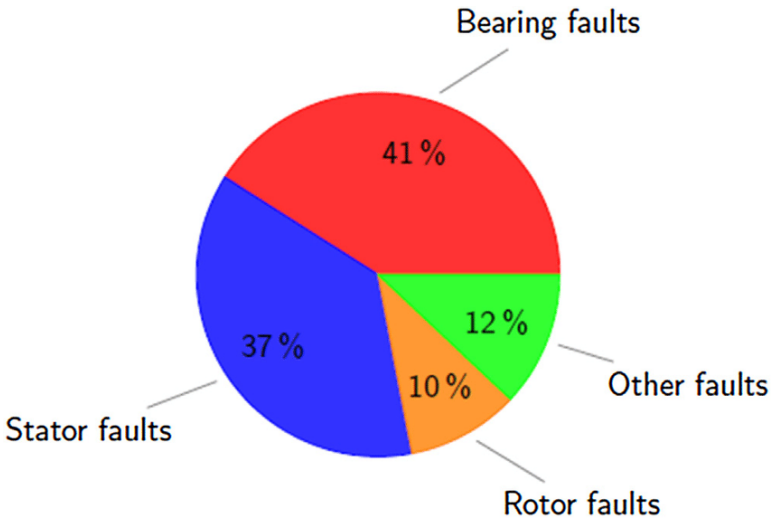

Figure 1.1: Failure distribution in a rotating machine.

Because a bearing vibration signal can be "contaminated" by other signals from different sources, most bearing fault detection methods apply a set of transformations to the target signal, before applying the Fourier transform. These "pre-processing" aim at isolating part of the signal that contained relevant diagnostic information. One of the most widely used method which accomplishes this with elegance, is the so call high frequency resonance technique (HFRT). Its apply a series of "filtering" operations in order to seep through, only part of the vibration signal, that contained potential traces of failure. The filtering operations, which are mathematical operations, are: high, low and band pass filtering. As their name indicate, a high pass filtering operation removes all low frequencies in a target signal, while a low pass filter accomplishes the apposite. A band pass filter removes all frequencies bellow and above predefined thresholds.

As will be shown later, a bearing with a defect, generates pulse like signals, buried deep inside the vibration signal, which can introduce non linearity as well as rendering the signal non stationary. Such pulses are temporal events, that the high frequency resonance technique (HFRT) attempts to isolate through filtering operations. However, the latter are not designed to resolve

such phenomenon. Therefore, the spectrum derived from the HFRT can be noisy, making failure frequency detection challenging at times. In addition, the Fourier transform operates in the frequency domain, therefore can not properly resolve temporal events. Furthermore, the basis functions of the Fourier transform are trigonometric extensions that are not compactly supported. Recall that a compactly supported function on a closed interval, has non zero value within the interval and zero else where. This property allows capturing pule like signals, since a pulse is approximated as a compactly supported function. To circumvent the aforementioned issues, this thesis presents an alternative method, based on Hilbert- Huang transform (HHT), for bearing fault detection. By design, the Hilbert-Huang transform can efficiently deal with signals that exhibit temporal features. Furthermore, in applying the HHT, the goal is to resolve all temporal events before generating the frequency spectrum of a target vibration signal. This strategy not only expose any potential anomalies in the signal, caused by a bearing failure, it also generates a "clean" frequency spectrum.

The Hilbert Huang transform was developed recently by Huang, (Huang et al. (1998)) to deal efficiently with non linear and non stationary processes. It decomposes data adaptively into its sub-components by using the so called empirical mode decomposition (EMD). Unlike Fourier transform where trigonometric functions are used to decompose signals, adaptive decomposition means that the basis functions are completely determined by the data itself, (Huang and Wu (2008)). This allow in theory, to access intrinsic and salient properties of data

In this brief introduction, the high frequency resonance technique (HFRT) was introduced as one of the prevalent methods used in bearing fault detection. Some of its weaknesses were briefly outlined, and the Hilbert Huang transform (HHT) was introduced as a possible fix. To further understand the HFRT and the HHT, section 1.2 presents a literature survey, covering both methods while outlining their respective strengths and weaknesses. To set apart the contribution of this thesis from what has been previously accomplished, section 1.3 describes in detailed the proposed method for bearing fault detection, based on Hilbert Huang transform.

In general, at the core of bearing fault detection are: the Fourier transform, signal filtering and the Hilbert transform, which was not mentioned so far. Therefore, Chapter 2 gives a brief introduction of the Fourier analysis, outlines the Hilbert transform and presents the basics of signal filtering. Moreover, the high frequency resonance technique, which is the predominant method in bearing fault detection, is applied to a case study, in order to detect bearing failure frequencies. To show the contribution of this thesis, Chapter 3 gives a thorough presentation of the proposed method. It covers the Hilbert-Huang transform (HHT), describes in detailed the application of the HHT to bearing fault detection, and uses a case study to demonstrate the efficiency of this method. Chapter 4 sums up what has been done, outline some of its weaknesses and gives a road map on how to improve up on them.

1.2 Literature review

As part of nearly all rotating machines, rolling bearing elements are one of the most frequent reasons for machine breakdown (Randal and Antoni (2010)). A rolling bearing element is made of mainly four parts, which are shown in Figure 1.2: The outer race indicated by 1, the balls trapped in the cage and labeled by 2 and 3 respectively. Finally, the inner race identified by 4 and 5. To facilitate machine rotation, a rotating shaft, which is a long cylindrical tube, is placed within the inner ring.

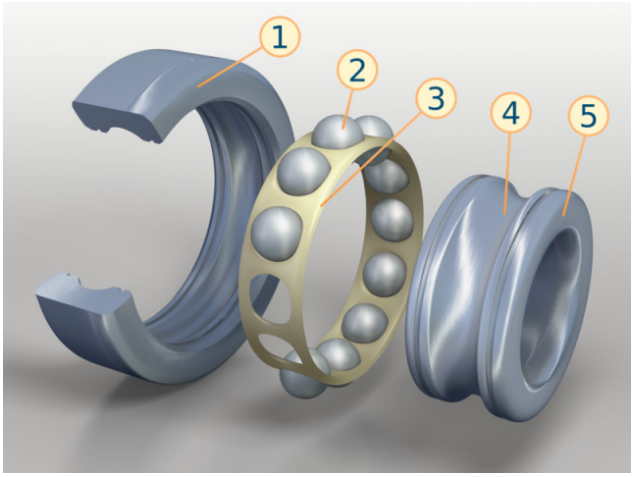

Figure 1.2: Exploded view of a rolling bearing. 1: outer race or ring, 2: balls or roller elements, 3: cage, 4 and 5: inner race or ring.

Premature and unexpected bearing breakdown, can halt production and incur high cost. In the worst case scenario, human lives can be impacted. To avert dramatic consequences of bearing failure, it is therefore crucial to detect incipient faults and take appropriate actions.

The geometry of a bearing is important in understanding its dynamic, and detecting early sign of failure. A key fact in bearing fault detection, is that a failure can be detected at a given frequency, called characteristic defect frequency or failure frequency. Here is why. A defect in one surface of a rolling element bearing also called balls, generates an impulse as it hits an other surface, (McFadden and Smith (1984a), McFadden and Smith (1984b)). As the bearing rotates, the impulses will occur periodically with a frequency (the defect frequency) which is uniquely determined by the location of the defect (McFadden and Smith (1984a)). Consequently, the failure frequency is derived from the bearing physical characteristics, and the rotational speed of the machine housing the bearing, (McFadden and Smith (1984a)). Therefore, finding the failure frequencies is the basis for bearing fault detection in most cases.

There are typically four failure frequencies: ball pass frequency outer race (BPFO), which is the

frequency at which a defect strikes the outer ring. Similarly, the ball pass frequency inner race (BPFI) is the frequency at which a defect hits the inner race. The fundamental train frequency cage (FTF) and the ball spin frequency (BSF) are frequencies at which a fault tricks the cage and the rollers (balls), respectively.

Finding the failure frequencies entails, transforming the bearing vibration signal into its corresponding frequency spectrum, where the defect frequencies reside. However, this apparent simple task is challenging. The vibration signal of a bearing is "infected" by the signals of other machines components or other vibration sources (Zhao et al. (2014), McFadden and Smith (1984a)). In addition, at the onset of failure, the fault frequencies are very weak (Zhao et al. (2014)). To efficiently detect failure frequencies, a de-noising of the signal is necessary and the weak early defect frequencies must be enhanced (Zhao et al. (2014)).

To address the aforementioned issues, several signal processing techniques have been proposed. The most prevalent are (Zhao et al. (2014)): the high frequency resonance technique (HFRT) (Darlow et al. (1974)), Spectral kurtosis (Antoni (2006), Antoni and Randall (2006), Antoni (2007)), wavelet analysis (Lin and Qu (2000), Qiu et al. (2006)), Hilbert Huang transform and the empirical mode decomposition (Yu et al. (2005), Lei et al. (2011)), cyclostationary approach (Antoni et al. (2004), Borghesani et al. (2013), Girondin et al. (2013)), minimum entropy deconvolution (Sawalhi et al. (2007), Jiang et al. (2013)), and stochastic resonance (Tan et al. (2009), He et al. (2012)). The most widely used method is however the high resonance frequency technique, because it is able to efficiently extract bearing diagnostic information, trough a sequence of filtering operations (Zhao et al. (2014)).

One of the first work on bearing fault detection was by (Balderston (1969)), who investigated bearings rings and roller elements (balls) natural frequencies, and observed that the signal induced by bearings defects are located in the high frequency zone of resonance exited by the internal impact of the faults, (Randal and Antoni (2010)). Let elaborate on this. As previously mention, a bearing defect will generate periodic impulses, which are forces that strike the bearing at the location of the defect, at a given frequency (failure frequency). When the failure frequency is equal, or nearly equal to the natural frequency of the bearing, this cause the latter to vibrate at a higher amplitude and frequency (McFadden and Smith (1984a)), then it would have at a different frequency. This phenomenon is called mechanical resonance. The latter will also occur in the machine housing the bearing, and in a potential sensor mounted on the bearing for collecting vibration data, (McFadden and Smith (1984a)).

It is therefore critical to detect, and isolate the resonance exited by the impulses generated by bearing defects, which in theory can be visible in the bearing vibration signal. However, bearing

diagnostic information in the form of failure frequencies are difficult to be directly observable in the raw signal (Randal and Antoni (2010)). This is due to the fact that the energy generated by the impulses induced by faults, are widely distributed over a wide range of frequencies (Randal and Antoni (2010)).

To circumvent the latter difficulty, the high frequency resonance technique (HFRT) was developed and allowed early detection of bearing failure (Broderick et al. (1972), Burchill (1973), Burchill et al. (1973), Darlow et al. (1974), Darlow and Badgley (1975b), Darlow and Badgley (1975a), Board (1975), Randal and Antoni (2010), Gupta and pradhan (2016), Khadersab and Shivakumar (2018)). In the high frequency resonance technique, a bearing vibration signal is band pass filtered, envelop-detected, low pass filtered and finally decomposed into its frequency spectrum by the Fourier transform.

The resonance induced by defects, are responses of the bearing, the machine housing the bearing, and possibly other surrounding machines (McFadden and Smith (1984a)). Therefore band pass filtering allows only the bearing signal to be recovered. The envelop detection procedure, extracts the high frequency resonance signal produced by the bearing defects. This signal is the superposition of two components: the high frequency resonance component and the low frequency bearing failure component. The latter is recovered by a low pass filter.

Recall that a band pass filtering process in signal processing, filters a signal by only letting through desire frequencies. On the other hand, the envelop detection process, takes a target signal and returns its envelope. The latter is the curve formed by joining all peaks in the signal. In practice, the envelope signal is derived by taking the Hilbert transform of the band pass filtered signal. The low pass filter procedure, sips out low frequency components of a signal.

The diagnostic power of the high frequency resonance technique rests on the key fact that it uses the envelope of the band passed raw vibration signal, before performing the Fourier transform (McFadden and Smith (1984a), Randal and Antoni (2010)). The envelope signal contains nearly all the information generated by bearing faults. It is worth noting that the rollers elements located in the bearing cage are subjected to random slip, (McFadden and Smith (1984a)), and the bearing failure frequencies variation is of the order of 1-2% (Randal and Antoni (2010)). This random slip changes the characteristics of the raw vibration signal, which makes it difficult to extract useful diagnostic information directly from the vibration raw signal, (Randal and Antoni (2010)). However, the sequence of operation applied to the raw signal to obtain the envelope addresses specifically this situation of slip (Randal and Antoni (2010)).

The envelope signal obtained from the High frequency resonance technique was also leveraged in other methods for bearing fault detection. For example, the spectral kurtosis method, uses

the frequency spectrum of the envelope signal from the short time Fourier transform (STFT), to find the frequency band of the pulses generated by bearing fault (Randal and Antoni (2010)). The spectral kurtosis uses fourth order statistical moment, to decompose the power of a signal with respect to frequencies, (Randal and Antoni (2010)). Note that the short time Fourier transform uses fixed window size where the Fourier transform is applied. Akin to the spectral kurtosis, is the power spectral density (PSD) method, which uses second order statistics, to obtain the energy contribution of frequencies. The bearing failure frequency will generally produce a larger energy distribution.

Although powerful, the high frequency resonance technique can produce a noisy spectrum for inner race or rolling element defect (McFadden and Smith (1984a)). The spectrum can become even more noisy when bearing defects are extensive (McFadden and Smith (1984a)). The dependency of the failure frequency on the machine rotation speed, can also be an issue as the rotating speed changes continuously or is unknown. Time varying rotation speed and load can cause the vibration signal to be non-stationary (Zhao et al. (2014)), which makes the Fourier transform inadequate in some cases (Huang et al. (1998), Huang and Wu (2008)).

In addition, the Fourier transform uses trigonometric basis functions, that are not locally compact, and can not handle efficiently nonlinear and non-stationary signal (Huang et al. (1998)). The reader might argue that the short time Fourier transform, can handle non-stationary signal. However, the issue in this case is the selection of the window size. An other important issue for all frequency based method so far mentioned, are the selection of a resonance frequency band for band pass filtering the raw vibration signal (Zhao et al. (2014)). The resonance frequency band, is the frequency band within which, the frequencies of resonance induced by the faults are located.

To extend bearing fault detection to cases where a Fourier transform based method or more generally frequency based methods are limited, numerous contributions have been made towards alternative methods such as Hilbert-Huang transform, wavelet transform and machine learning (Zhang et al. (2019), Xiaoan and Minping (2018), Rai and Upadhyay (2016), Konar and Chattopadhyay (2011), Rai and Mohanty (2006)). These methods can directly operate in the time domain (on the time signal), as opposed to Fourier transform based methods that require a frequency spectrum.

The Hilbert-Huang transform (HHT) as most signal analysis tool, is predicated on the key assumption that a signal has multiple components, and can be decomposed into single oscillatory modes called intrinsic mode functions (IMFs) (Fosso and Molinas (2019), Huang and Wu (2008), Huang et al. (1998)). It was developed to deal efficiently with nonlinear and non-stationary signals. As opposed to Fourier transfrom, the HHT method does not use a-priori basis functions.

It uses local properties of a signal such as the extrema of a signal, the mean, with a series of operations to obtain a hierarchy of single component signals, that range from high to low frequency time signals. By using local properties, the HHT is able to model local events such as pulses emitted by bearing faults (which will be shown in this thesis). The wavelet transform on the other hand, utilizes locally compact basis functions, and was primarily used in geophysics, to model high frequency short duration seismic pulses, (Bogges (2009)), similar to those emitted by bearing faults .

For bearing fault detection, (Fan et al. (2016)) decomposed the vibration signal obtained from a motor into intrinsic mode functions, and evaluated the Hilbert-Huang energy spectrum (also called marginal spectrum) of each IMF, to detect sign of fatigue, oxidation and mechanical structure deformation. In the same fashion (Peng et al. (2004), Soualhi et al. (2015), Osman and Wang (2014),Osman and Wang (2013a), Osman and Wang (2013b), Li et al. (2009b)) Applied the marginal spectrum to identify bearing characteristic defect frequencies. The energy spectrum also called power spectrum or energy density, is the energy contribution of each frequency, derived from the intrinsic mode functions. To compute the energy density, the Hilbert transform of the absolute value of the square of an IMF is first computed. Secondly, the integral of the latter is evaluated over the domain of variability of the signal. The Hilbert transform of a signal is the convolution of the signal with the function $\frac{1}{\pi t}$, where t is a dummy variable.

One of the issue of the Hilbert Huang transform, is selecting the appropriate intrinsic mode functions (Fosso and Molinas (2019)). The IMFs obtained from a target signal, are hierarchy of mono component signals, ranging from high to low frequency. However, due to a phenomenon called mode mixing, an IMF, rather then having a single oscillatory mode, can have more than one mode, resulting an IMF to lose its physical meaning (Fosso and Molinas (2019)). For a given process such as the vibration of a bearing in a motor, there exists multiple sub processes corresponding to the vibration of sub components of the motor and the bearing. Therefore mapping the correct IMFs to the corresponding sub processes is a daunting task.

To resolve this issue, (Osman and Wang (2013a), Osman and Wang (2013b)) used a linear combination of two similarities measure (Linear and non-linear similarity) to select the target IMF and applied the energy spectrum of the IMFs to identify bearing failure frequencies. The similarities measures, quantifies the “sameness” of two distributions. In the same fashion (Osman and Wang (2014)) applied a weighted D’Agostino Pearson (DP) normality test to select the more relevant IMFs (IMF representing defect component). The DP test uses both the skewness and the kurtosis to assess normality. The skewness is a statistical estimator that measures the symmetry of a probability distribution, while the kurtosis also a statistical estimator, measures the “tailedness” of a distribution. The Kurtosis and skewness of a normal distribution are 3 and 0 respectively. (Peng et al. (2004)), used the correlation coefficient as a criteria for IMF selection.

1.3 Contributions

The ubiquity of Fourier transform in science and technology, is the proof of its efficiency in solving complex scientific and industrial problems, such as bearing fault detection. In the latter, Fourier transform is the core of several methods that rely on frequencies spectrum to identify bearing failure frequencies. One of the predominant methods is the so call high frequency resonance technique (HFRT). As outlined earlier, the high frequency resonance technique, first uses successive filtering operations to remove "noisy signals" in a target vibration signal, then apply Fourier transform. This result in a frequency spectrum, which may potentially contain a bearing failure frequency.

in spite of its wide spread application, the Fourier transform is not without "limitations". The latter are the results of the mathematical assumptions imposed, in order to formulate a rigorous theoretical framework. One such assumption, is to say that a target signal is the superposition of trigonometrical extensions, each with constant amplitude and frequency. This makes it hard to apply the Fourier transform to amplitude and frequency modulated signals, generated for example by changing load. In addition, Fourier transform operates in the frequency domain, thus unable to access temporal phenomenon in a signal. Lastly, applying the Fourier transform to a signal can produce a noisy frequency spectrum when bearing faults become pronounced. This can introduce challenges in the process of identifying bearing failure frequencies.

To address the aforementioned issues, when it comes to bearing fault detection, this thesis introduces a mixed method. It comprises the Hilbert-Huang transform coupled with a robust seasonal trend decomposition method. There are three steps in this method: In step 1, the Hilbert-Huang transform (HHT) is used to decompose a signal into nearly mono component signals. This is akin to the Fourier transform, except that there is no predefined basis functions, and the resulting components reside in the time domain. In step 2, the seasonal parts of each component are extracted through the so called seasonal trend decomposition by LOESS method (STL). The seasonal parts are the oscillatory components. In the last step, the power spectrum density of each oscillatory part is computed through the periodogram which is an estimation of the power spectrum density. The latter is the energy distribution of the frequency spectrum. The power spectrum, essentially shows the energy contribution of each frequency component of a signal.

As opposed to the high frequency resonance technique, the method proposed in this thesis, generates a relatively noiseless frequency spectrum.

Chapter 2

Application of the High frequency resonance technique in bearing Fault detection

This chapter covers the application of the high frequency resonance technique (HFRT), to bearing fault detection. It starts with the background materials in section 2.1, where section 2.1.1 presents an overview of the Fourier analysis, while section 2.1.2 describes signal filtering in general, and in particular the Hilbert transform, which is a special kind of linear filter.

After discussing the background materials, section 2.2 instantiates a case study for the high frequency resonance technique, applied to detecting bearing failure frequencies. The data for this case study, was provided by the intelligence system division, of the national aeronautics and Space administration (NASA).

2.1 Background materials

2.1.1 Fourier analysis

From solving differential equations to analyzing sound waves, images and signals in general, Fourier analysis has a profound impact in science and engineering. It provides a convenient way to transforming signals into a series of frequencies called frequency spectrum. And in doing so, it reveals hidden aspects of data. The bulk of Fourier analysis is to decompose as well as reconstruct a signal, or more generally a function, into trigonometric extensions. In bearing fault detection, the main goal is to decompose a signal into its elementary frequency spectrum, which contains the bearing defect frequencies, if they exist.

In an application view point, a signal can be viewed as a series of observations generated by a

given process, and recorded at discrete or continuous time intervals. The underlying process might be the sum of given sub-processes. In this case, the frequency spectrum will reveal all the sub-processes characteristics. This is illustrated in Figure 2.1, where a signal (in dark) is decomposed into four sinusoidal components (in blue). Each sinusoidal component is uniquely defined by its amplitude and its frequency (in red). The representation of all amplitudes versus their corresponding frequencies is called the frequency spectrum.

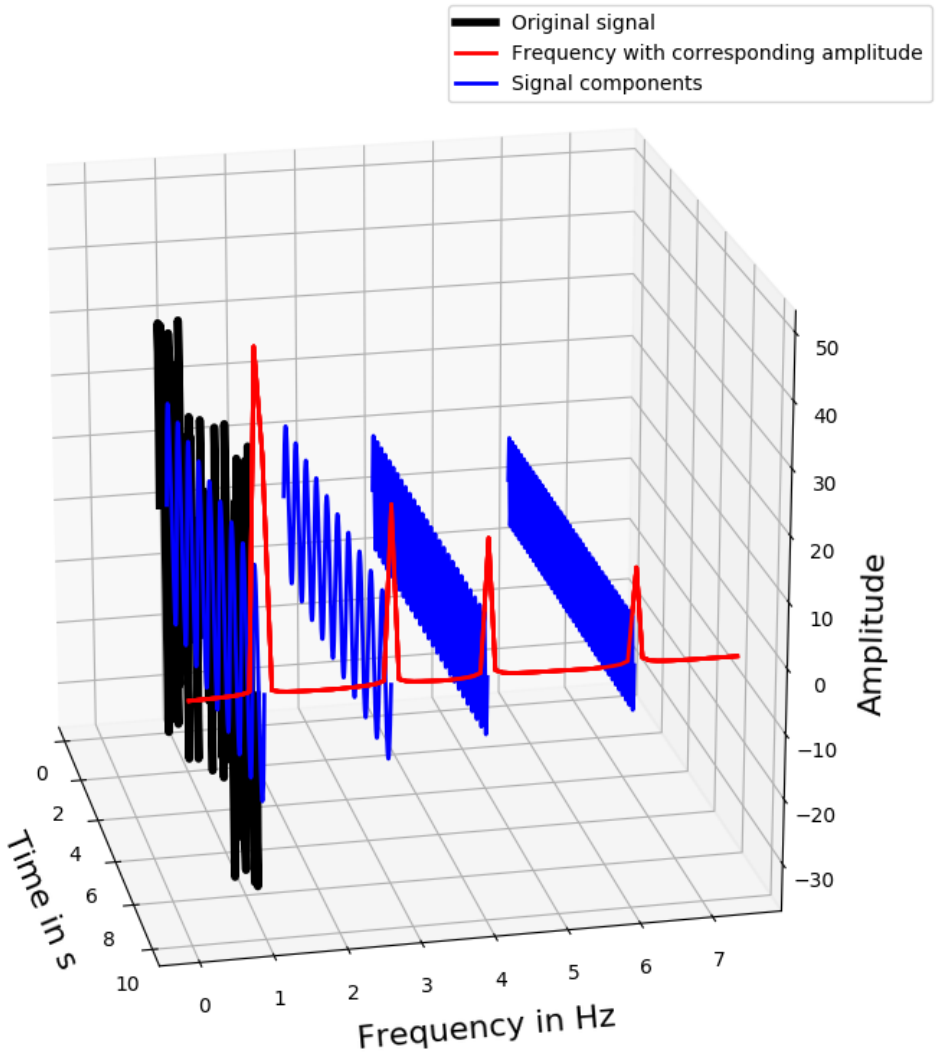

Figure 2.1:

In a theoretical view point, however, the results that underlie Fourier analysis can be sum up as followed: Given a function (or a signal), find a trigonometric series that converges to the target function. The latter statement falls under the umbrella of the general problem of function

approximation.

The basic ingredients required to approximate a function in this scenario are: a vector space, a basis, which is a subspace of the vector space, and a mathematical operation such as an inner product that maps two vectors to a real number. If a vector space has an inner product, we say that the vector space is an inner product space. The inner product operation, can be used for example to express the orthogonality property. The latter plays an import part in simplifying some algebraic operations.

Before continuing, let clarify some symbols. In this section, the letters f, V, V_0 are used for an arbitrary function, a vector space, and a subspace of a vector space, respectively. Basis functions will be denoted by $\{\varphi_0, \cdots, \varphi_n\}$, where n can either be a finite integer or infinite. Having made this clarification, let elaborate on the concept of function approximation.

The function approximation process in light of Fourier analysis goes like this: Given an arbitrary function f, that is to be approximated, pick an appropriate vector space V, such that $f \in V$. Furthermore, define a subspace V_0 of the vector space V, and construct an inner product on V_0, if it does not exist. Next, fine an appropriate basis of V_0. A basis of V_0 is a set of linearly independent vectors $\{\varphi_0, \cdots, \varphi_n\}$ in V_0, that span V_0. This means that any vector in V_0 can be written as a linear combination of the basis vectors.

Having all this in place, the best approximation of the function f is its orthogonal projection in the inner product space V_0. Figure 2.2 shows an illustration of a generic mechanism of function approximation by orthogonal projection, where f_0 is the orthogonal projection of f in the subspace V_0 of V.

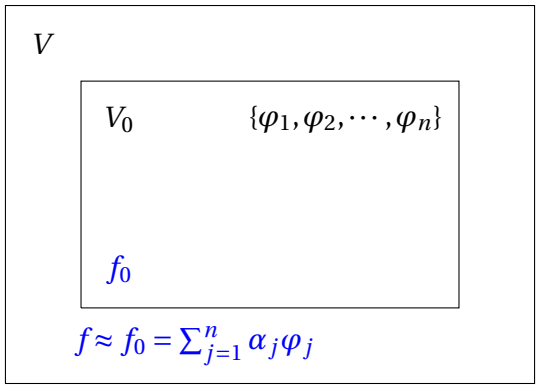

Figure 2.2: Illustration of a generic function approximation process of a function f into a subspace V_0 of a vector space V. φ_j are the basis functions and α_j are real numbers for $j = 1, \cdots, n$.

Having the generic function approximation defined in Figure 2.2 as a blue print, the Fourier space V_0 is a subspace of the space of all continuous functions of the interval $[0, T]$ and denoted by $C[0, T]$, which is V in Figure 2.2. The subspace V_0 of V is spanned by

$$\left\{1, \cos\left(\frac{2\pi t}{T}\right), \cdots, \cos\left(\frac{2\pi N t}{T}\right), \sin\left(\frac{2\pi t}{T}\right), \cdots, \sin\left(\frac{2\pi N t}{T}\right)\right\}.$$

which corresponds to

$$\{\varphi_1, \varphi_2, \cdots, \varphi_n\},$$

from figure 2.2.

Let sum up this introductory section of Fourier analysis. Fourier analysis is concerned with the general problem of periodic and non periodic functions approximation. The former and the latter are treated by Fourier series and Fourier transform, respectively. Given a periodic function, its Fourier series is given as a discrete superposition of trigonometric functions, and its Fourier transform is expressed as continuous superposition of exponential functions. To extend the discussion of Fourier analysis, section 2.1.1 presents the Fourier series, while section 2.1.1 elaborates on Fourier transform. In section 2.2, the application of Fourier analysis to bearing fault detection is presented.

Fourier series

Let f be an arbitrary function, defined on an interval of length $L = \pi$. Its Fourier series representation is given by the infinite series

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right)\right), \tag{2.1}$$

where the coefficients $a_0, a_1, \cdots, b_1, b_2, \cdots$, corresponding to the α_j in Figure 2.2, are given by

$$\begin{aligned} a_n &= \frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{m\pi t}{L}\right)\mathrm{d}t, \quad n = 0, 1, 2, \cdots \\ b_n &= \frac{1}{L}\int_{-L}^{L} f(t)\sin\left(\frac{n\pi t}{L}\right)\mathrm{d}t, \quad n = 1, 2, \cdots \end{aligned} \tag{2.2}$$

Approximating a function by an infinite series such as in equation (2.1), raises few interesting questions. The right hand side of (2.1) being an infinite series, must obviously converge for this expression to be valid. As such, under which condition does it converge? and what type of convergence is required? what are the characteristics of the function that it converges to?.

At first glance, one can observe that the trigonometric functions in the infinite series are periodic with period $2L$. This suggests that the class of functions admitting a Fourier series must be periodic, with period $2L$. The second observation is that the trigonometric extensions are continuous, therefore some sort of continuity property must be imparted on the function f. It turns out that, the convergence of Fourier series and the characteristics of the function f are intimately tied. This is formerly expressed in the following theorem

Theorem 1 *The function $f(t)$ will have a convergent Fourier series, with coefficients given by (2.2), provided $f(t)$ is periodic with period $2L$, and both $f(t)$ and $f'(t)$ (the derivative of $f(t)$) are at least piecewise continuous.*

A useful interpretation of what it means for a function $f(t)$ to have a Fourier representation, can be expressed in terms of the partial sum

$$S_N = \frac{a_0}{2} + \sum_{n=1}^{N} \left(a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right)\right). \tag{2.3}$$

The function $f(t)$ has a Fourier representation given by (2.1) means that

$$\lim_{N\to\infty} S_N(t) = f(t) \tag{2.4}$$

where N is an integer.

Theorem 1 ensures that if a function is periodic, and both the function and it first derivative are piecewise continuous, then it will have a convergence Fourier transform. Now, the question which remains to be addressed is, what type of convergence should be expected? The convergence type of Fourier series discussed in this section are: uniform convergence, pointwise convergence and mean-square convergence. The two first convergence type deal with continuous functions, while mean-square convergence treats discontinuous functions.

Uniform convergence is more stringent than pointwise convergence, while pointwise convergence is more strict than mean-square convergence. Specifically, uniform convergence implies pointwise convergence, and pointwise convergence implies mean-square convergence. Therefore, together, they form a hierarchy of convergence type.

In terms of function approximation, convergence of infinite series starts with examining the error incurred by approximating the function by the infinite series. This error is expressed as

$$E_N(t) = f(t) - S_N(t) = \sum_{n=N+1}^{\infty} \left(a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right)\right). \tag{2.5}$$

It is now apparent that the Fourier series converges pointwise to $f(t)$ if and only if

$$\lim_{N\to\infty} E_N(t) = 0 \quad \text{for all } t. \tag{2.6}$$

Pointwise convergence can also be formulated in terms of the series in equations (2.3, 2.4). Formally, the sequence S_N converges pointwise if for some fixed t and fixed ϵ (a small real number), it exists some $M \in \mathbb{Z}$, such that for $N \geq M$

$$\left|S_N(t) - f(t)\right| \leq \epsilon. \tag{2.7}$$

All the previous discussions can be summarized in the following theorem

Theorem 2 *if f is differentiable on $[-\pi,\pi]$ then the Fourier series converges pointwise and 2.1 holds at every point t where $f(t)$ is continuous.*

Now let turn our attention to uniform convergence. since the trigonometric functions in the infinite Fourier series are bounded above by 1, we can write

$$\left|E_N(t)\right| \leq \sum_{n=N+1}^{\infty} \Big(\left|a_n\right| + \left|b_n\right|\Big), \tag{2.8}$$

and provided that the left hand side converges, uniform convergence is warranted. This is summed up in the following theorem

Theorem 3 *The Fourier series for $f(t)$ converges uniformly, and therefore $f(t)$ is continuous if*

$$\sum_{n=N+1}^{\infty} \Big(\left|a_n\right| + \left|b_n\right|\Big), \tag{2.9}$$

converges

Since a discontinuous function can not have a uniform convergence Fourier series, what should be expected from discontinuous function in terms of convergence? this question can be answered in terms of mean-square convergence. The mean-square value of the approximation error $E_N(t)$, also interpreted physically as the average energy of the approximation error $E_N(t)$ (define in equation (2.5)) in the interval of length L, is given by

$$\frac{1}{2L}\int_{-L}^{L} \big(E_N(t)\big)^2 \,\mathrm{dt} = \frac{1}{2}\sum_{n=N+1}^{\infty} \Big(a_n^2 + b_n^2\Big). \tag{2.10}$$

If the right hand side of equation (2.10) converges then

$$\lim_{N\to\infty}\left(\frac{1}{2L}\int_{-L}^{L}\left(E_N(t)\right)^2\,\mathrm{dt}\right)=0. \tag{2.11}$$

This result can be formulated in the following theorem.

Theorem 4 *The Fourier series for $f(t)$ converges in the mean-square or almost everywhere if*

$$\sum_{n=N+1}^{\infty}\left(a_n^2+b_n^2\right), \tag{2.12}$$

converges

Theorem 4 ensures that the Fourier series will converge everywhere except at the points of discontinuity of $f(t)$. This is also refer to as almost everywhere convergence.

From the previous discussion, The coefficients a_n and b_n of the Fourier series, somehow drive convergence. Therefore, they determine the series its self. Being so important, let find an interpretation. By setting

$$\begin{aligned}
A_n &= \sqrt{a_n^2+b_n^2}\\
\phi_n &= \tan^{-1}\left(\frac{b_n}{a_n}\right)\\
\delta_n &= \frac{L}{n\pi}\tan^{-1}\left(\frac{b_n}{a_n}\right)\\
A_0 &= \frac{a_0}{2}
\end{aligned} \tag{2.13}$$

The Fourier series of $f(t)$ can be rewritten as

$$f(t)=A_0+\sum_{n=1}^{\infty}A_n\cos\left(\frac{n\pi}{L}(t-\delta_n)\right). \tag{2.14}$$

Now from equation (2.14), it is clear that the Fourier series decomposes the function $f(t)$ into frequency components, each with amplitude A_n, phase angle ϕ_n and delay δ_n. The amplitudes together with the frequency components form the frequency spectrum.

So far, only periodic functions have been examined. However, in real application, non periodic functions or signals are prevalent. For this reason, the Fourier transform is used to decompose such functions.

Fourier transform and the fast Fourier transform

On an application point of view, the Fourier series is limited. Let explain why. Most signal found in real application are non periodic waveform that vibrate at non integer frequencies. However, the Fourier series only deals with periodic functions. Its decomposes signals into trigonometric extensions in $[-\pi,\pi]$, that vibrates at integer frequencies. To circumvent the short coming of the Fourier series, the Fourier transform is introduced. Its decomposes non periodic functions into sinusoidal extensions that vibrates at frequencies that are real number on infinite time interval [ref]. This is of great practical importance, at least for bearing fault detection, since failure frequencies are often real numbers.

The Fourier transform $\hat{f}(s)$ of a given function f, is a complex valued function defined by the integral

$$\hat{f}(s) = \int_{-\infty}^{\infty} e^{-2\pi i s} f(t) \mathrm{d}t, \tag{2.15}$$

assuming that f is defined for all real numbers t and $s \in \mathbb{R}$. The Fourier transform defined by (2.15) gives a continuous spectrum of frequencies as opposed to the Fourier series of a periodic function which generates a discrete spectrum of frequencies. The function f can be recover from the continuous spectrum of frequencies by taking the inverse Fourier transform to obtain

$$f(t) = \int_{-\infty}^{\infty} e^{2\pi i s} \hat{f}(s) \mathrm{d}s, \tag{2.16}$$

In technical semantics, we say that the Fourier transform $\hat{f}(s)$ is defined on the frequency domain, while $f(t)$ is defined on the time domain. This define a time and frequency duality. For any real number s, the square magnitude $|\hat{f}(s)|^2$ is called the power spectrum or the spectral power density. Its gives the energy of a signal in terms of frequency. The frequency domain and the time domain are related by the so called Parcevals identity given by

$$\underbrace{\int_{-\infty}^{\infty} |f(t)|^2 \mathrm{d}t}_{\text{energy of the signal f}} = \underbrace{\int_{-\infty}^{\infty} |\hat{f}(s)|^2 \mathrm{d}s}_{\text{Energy spectrum}} \tag{2.17}$$

Equation (2.17) says that the total energy in the time domain is equal to the total energy in the frequency domain. The left hand side of Equation (2.17) defines the total energy of the signal f while the right hand side gives what is called the energy spectrum, which is the total energy in the frequency domain. The latter quantifies the energy contribution of all frequencies.

The time and frequency duality defined above, allows time domain information to be recover in the frequency domain. To illustrate this fact, let assume that a vibration sensor records the vibrating movement of a bearing over a period of time. This measurement can be view as a

function of time and reside in the time domain. By applying the Fourier transform, we can decompose the signal into frequencies which reside in the frequency domain. Assume further that the goal is to detect any failure incur by the bearing. For a bearing, the failure frequency can be computed based on the bearing geometrical characteristics. Once we known the failure frequency, it remains to search for it in the frequency domain. In the subsequent section we will discus how Fourier transform can be use to detect specific bearing failure frequencies. For the time being, let discuss the conditions under which the integrals defined in equations (2.15, 2.16) exist.

Theorem 5 *If f is a continuously differentiable function with*

$$\int_{-\infty}^{\infty} |f(t)| \mathrm{dt} < \infty, \tag{2.18}$$

then

$$f(t) = \int_{-\infty}^{\infty} e^{2\pi i s} \hat{f}(s) \mathrm{ds},$$

where $\hat{f}(s)$ the Fourier transform is given by

$$\hat{f}(s) = \int_{-\infty}^{\infty} e^{-2\pi i s} f(t) \mathrm{dt}.$$

Thus if f is continuously differentiable and the integral defined by (2.18) is absolutely convergent, the Fourier transform of f exists. In cases where the latter condition is not satisfied, the Fourier transform still exists if the integral defined by (2.18) is conditionally convergent.

So far, in our discussion of Fourier transform, only continuous function (signal) have been mention. However, in many applications such as signal analysis, we are primarily dealing with discrete signals. Recall that a discrete signal has values at discrete times. For such signals, we need to apply the discrete Fourier transform.

Definition 1 *Let $x = \{x_j\}_{j=0}^{n}$ be a sequence of real numbers. The discrete Fourier transform of x denoted by $\hat{x}$ is the sequence $\hat{x} = \{\hat{x}_k\}$, where*

$$\hat{x}_k = \sum_{j=0}^{n-1} x_j \overline{w}^{jk} \, with \quad w = \exp\left(\frac{2\pi i}{n}\right),$$

where i is the imaginary complex number, n is an integer and $\overline{w}$ denote the complex conjugate of w.

Practically, the computation of the Discrete Fourier transform reduces to the following matrix

computation

$$\underbrace{\begin{pmatrix} \hat{x}_0 \\ \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_{n-1} \end{pmatrix}}_{\hat{x}} = \underbrace{\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w & w^2 & \cdots & w^{n-1} \\ 1 & w^2 & w^4 & \cdots & w^{2(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & w^{n-1} & w^{2(n-1)} & \cdots & w^{(n-1)^2} \end{pmatrix}}_{F_n} \underbrace{\begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{pmatrix}}_{x}, \tag{2.19}$$

where F_n is a symmetric matrix. The matrix operation defined in equation (2.19) requires n^2 multiplications. For large n, this can incur a significant overhead. Fortunately, since the matrix F_n is symmetric, the number of multiplications can be reduced to $5n\log_2(n)$, by applying the so called fast Fourier transform algorithm.

$$\hat{x} = \sum_{j=0}^{\frac{n}{2}-1} \overline{W}^{jk} + \overline{w}^k \left(\sum_{j=0}^{\frac{n}{2}-1} x_{2j+1} \overline{W}^{jk} \right) \tag{2.20}$$

The Fourier transform not only computes the frequency spectrum of a signal. It is also a tool in the process of deriving the analytical representation of a real value signal. An analytical representation of a real valued signal is its complex representation that makes it possible to represents a signal in terms of frequency modulation (frequency variation). The analytical representation of a signal is derived in terms of the Hilbert transform. In terms of the Fourier transform, the Hilbert transform of a signal is found by computing its Fourier transform, remove the negative spectrum and finally take the inverse Fourier transform. The Hilbert transform is also used in addition with the Fourier transform for bearing fault detection. In section 2.1.2 the Hilbert transform is presented followed by section 2.2 which apply the Fourier transform and the Hilbert transform to bearing fault detection.

2.1.2 Digital filters and Hilbert transform

Filter

A digital filter can be defined as a dynamic system in general, and in particular a discrete time system, that removes undesirable part of a signal such as noise, or extracts desirable components that reside within certain frequency ranges (band). The input signal to the system is some time refer to as the excitation, while the output is called the response. To be more specific, the discrete time system is a mathematical function that maps the input signal to the output. Furthermore, a discrete time system operates on discrete time signals.

A digital filter is characterized by its frequency response, which when known, completely defines

the output signal frequency and phase variation. The frequency response is also defined as the transfer function of the system evaluated on the unit circle. The transfer function is the function that gives the relationship between the output signal produced by the filter, and the input signal.

As illustration, let ξ be the mapping that represents a discrete time linear system. From the eigenfunction property,

$$\xi(z^k) = T(z)z^k \tag{2.21}$$

where z is a complex number, k an integer, and $T(\cdot)$ the transfer function of the discrete linear system. In addition, if ξ is stable, meaning that it can be evaluated on the unit circle, ($z = e^{jk\Omega}$, $j = \sqrt{-1}$), then (2.21) becomes

$$\xi\left(e^{jk\Omega}\right) = T\left(e^{j\Omega}\right)e^{jk\Omega} \tag{2.22}$$

where Ω is the phase of the complex number z lying on the unit circle. It is worth noting that a stable system produces a bounded signal from an input bounded signal, which is a desire property in real life applications. Moreover, the response of a single frequency component of the sinusoidal $2\cos(\Omega k)$, is given by

$$\begin{aligned} 2\xi\left(\cos(\Omega k)\right) &= 2\left(\xi\left(e^{jk\Omega}\right) + \xi\left(e^{-jk\Omega}\right)\right) \\ &= 2\left(T\left(e^{j\Omega}\right)e^{jk\Omega} + T\left(e^{-j\Omega}\right)e^{-jk\Omega}\right) \\ &= 2|T\left(e^{j\Omega}\right)|e^{j\Omega k} \\ &= \left|T\left(e^{j\Omega}\right)\right|\cos\left(k\Omega + \angle T\left(e^{j\Omega}\right)\right) \end{aligned} \tag{2.23}$$

where $\angle T\left(e^{j\Omega}\right)$ is the angle of the frequency response and $\left|T\left(e^{j\Omega}\right)\right|$ is the amplitude response. The observation that can be made from equation (2.23) is that, the response of a single phase sinusoidal function, is given by the amplitude of the frequency response, and the sum of the phase $k\Omega + \angle T\left(e^{j\Omega}\right)$. In retrospect, equation (2.22) says that the frequency response of a filter, completely determined the excitation of an input signal.

In this thesis, the Butterworth filter is used in the implementation of the high frequency resonance technique. This filter tends to approximate an ideal one, which "should not only completely reject the unwanted frequencies (stopband) but should also have uniform sensitivity for the wanted frequencies (passband)" (Butterworth (1930)). Uniform sensitivity for the wanted frequencies means that, the wanted frequency band should be "flat". As a low pass filter, the frequency response of the Butterworth filter in terms of frequency ω is given by

$$T(\omega) = \frac{1}{\sqrt{1+\left(\frac{\omega}{\omega_c}\right)^{2n}}} \tag{2.24}$$

where n is the order of the filter and ω_c is called cutoff frequency, and segregates the passband and the stopband. This low pass filter is the building block of other filters type(bandpass, high-pass and so on). Figure 2.3 shows the graph of a normalized Butterworth amplitude response with respect to the angular frequency ω, with cutoff frequency $\omega_c = 1$. As a low pass filter, the passband region includes frequency range from 0 to 1 radian per second. In this region, as the filter order n increases, the passband region converges to a flat curve, and the amplitude response converges to a step function.

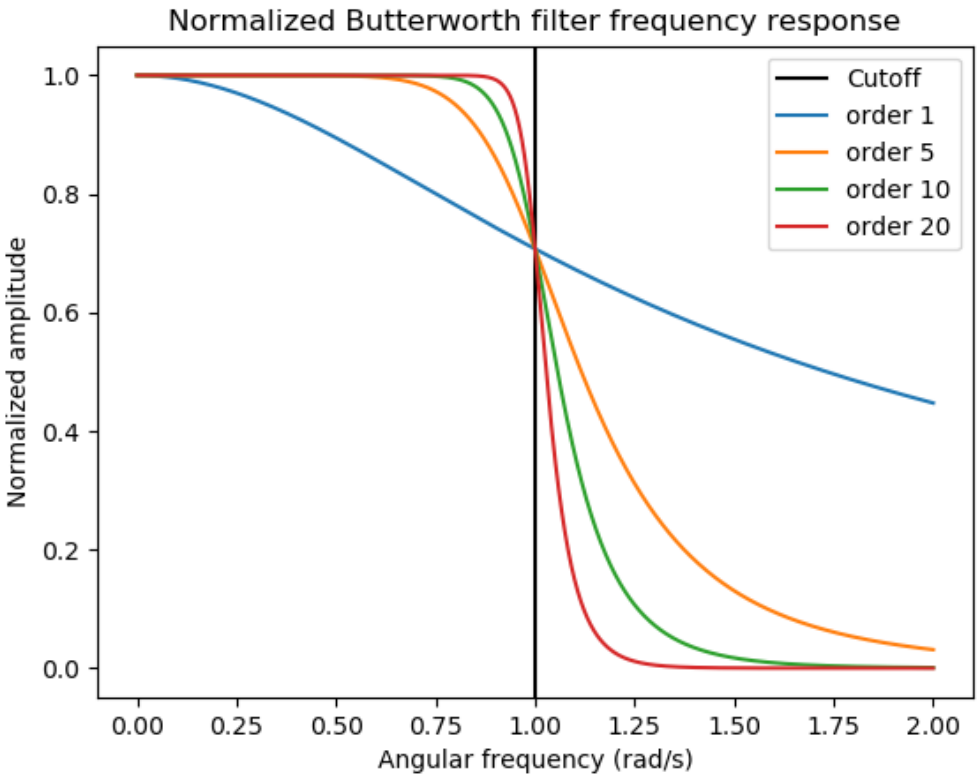

Figure 2.3:

Hilbert transform

The Hilbert transform is an integral transform, first introduced by David Hilbert to solve integral equations in mathematical physics (Gabor (1982)). Its physical interpretation is equivalent to a special kind of linear filter, which shift a signal spectral component phase by $\frac{\pi}{2}$, while keeping

its amplitude unchanged (Feldman (2010)). This is the physical meaning behind expressing the Hilbert transform of a signal as its convolution with the function $\frac{1}{\pi t}$ (Feldman (2010)). Moreover, the generalization of Euler formula

$$e^{i\theta} = \cos(\theta) + i\sin(\theta) \tag{2.25}$$

to a complex function as

$$Y(t) = u(t) + i\,v(t), \tag{2.26}$$

due to Gabor, was made possible by the Hilbert transform (Feldman (2010)). In equation (2.26), the real value function $v(t)$ is the Hilbert transform of the real value function $u(t)$. Furthermore, if $Y(t)$ is a signal which depends on time t, then $Y(t)$ is an analytic signal, which is the complex representation of the signal $u(t)$ in the upper half complex plane.

Equation (2.26), has a profound implication in signal processing, notably in problems pertaining to non-stationary signal analysis. In the latter, spectral properties such as amplitude and frequency are modulated in time. Therefore, it is safe to posit that, an appropriate representation of such signals, should incorporate instantaneous amplitude and frequency. The former and the latter, are akin to amplitude and frequency variation in time. This is precisely achieved through the Hilbert transform, in part.

Before expanding the mathematical formulation of the Hilbert transform, a definition of an analytic signal, and the clarification as to why it is important for non-stationary problems, are required. Formally, an analytic signal is a complex signal whose imaginary part is the Hilbert transform of its real part. A real valued signal $s(t)$, can be extended to a well defined complex signal $Y(t)$, given by

$$Y(t) = s(t) + i\,H\{s(t)\} \tag{2.27}$$

where $H\{s(t)\}$ is the Hilbert transform of $s(t)$. If (2.27) holds, then $Y(t)$ is said to be an analytic signal. In addition, if $s(t)$ is a mono component signal, its instantaneous amplitude (envelop) $a(t)$ and instantaneous frequency $\omega(t)$ as a function of the time variable t, are well defined and given by

$$a(t) = \sqrt{s(t)^2 + Y(t)^2} \tag{2.28}$$

$$\omega(t) = \frac{d\Psi(t)}{dt} \tag{2.29}$$

where

$$\Psi(t) = \tan^{-1}\left(\frac{H\{s(t)\}}{s(t)}\right) \tag{2.30}$$

is the instantaneous phase. With these formulations, $s(t)$ can be extended to the analytic signal

$$Y(t) = \mathbf{R}_e \left[|a(t)| e^{i\Psi(t)} \right], \tag{2.31}$$

where $\mathbf{R}_e$ is the real part of the enclosed complex function.

An alternative way to describe a signal and its Hilbert transform, is to say that they are in quadrature. Although they differ in form, a signal and its Hilbert transform contain the same spectral component(s) (Gabor (1944)). In layman's terms, that is to say for example that, "a human ear could not distinguish between a sound wave and its Hilbert transform" (Gabor (1944)).

In contrast to (2.31), the Fourier analysis represents a signal in frequency domain, where the concept of instantaneous frequency and amplitude can not be defined de facto. The amplitude and frequency modulation expressed in (2.31) gives a satisfactory representation of non-stationary signals, where obviously the frequency and the amplitude varies continuously with time.

Mathematically, the Hilbert transform $H\{s(t)\}$ of a signal $s(t)$, is define as its convolution with the function $\frac{1}{\pi t}$ expressed as

$$H\{s(t)\} = \frac{1}{\pi t} * s(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{s(\eta)}{t-\eta} \mathrm{d}\eta. \tag{2.32}$$

Because of the singularity at $t = \eta$ in (2.32), the indefinite integral might not converge. As a circumvention, it is evaluated by applying the Cauchy principal value method, as indicated by the letter P in front the integral. Furthermore (2.32), can be written as

$$\begin{aligned} H\{s(t)\} &= \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{s(\eta)}{t-\eta} \mathrm{d}\eta \\ &= \frac{1}{\pi} \lim_{\epsilon \to 0^+} \left(\int_{-\infty}^{-\epsilon} \frac{s(\eta)}{t-\eta} \mathrm{d}\eta + \int_{\epsilon}^{\infty} \frac{s(\eta)}{t-\eta} \mathrm{d}\eta \right) \end{aligned} \tag{2.33}$$

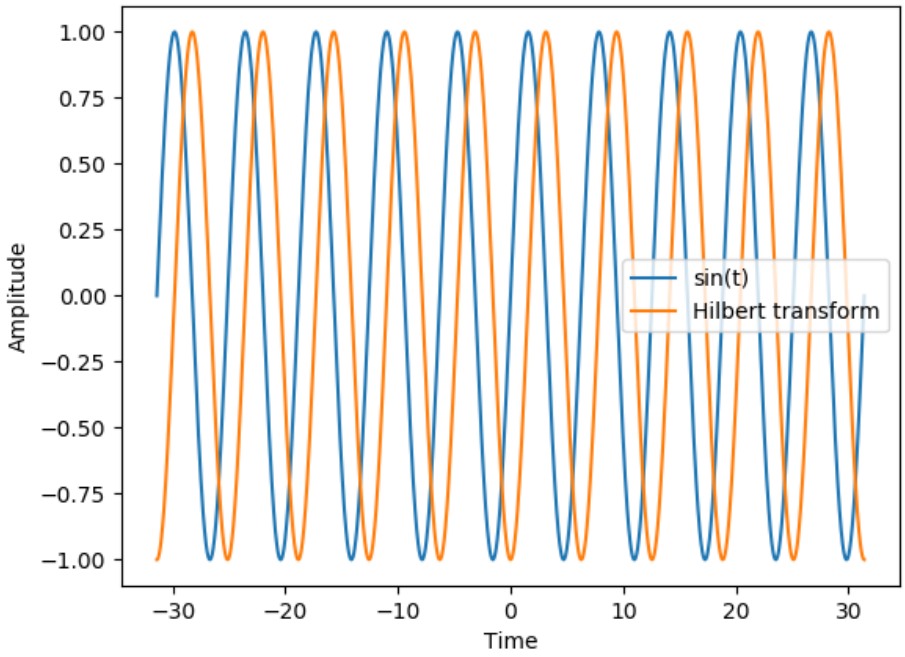

Figure 2.4: The Hilbert transform of $\sin(ct)$ given by $\sin(ct - \frac{\pi}{2})$. $c = 1$

Figure 2.4 shows the Hilbert transform of $\sin(ct)$ which is $\sin\left(ct - \frac{\pi}{2}\right)$. In the same fashion, the Hilbert transform of $\cos(ct)$ is $\cos\left(ct - \frac{\pi}{2}\right)$, where $c = 1$. Therefore, in the frequency domain, the Hilbert transform imposes a phase shift of $\frac{\pi}{2}$ to frequency components of a signal.

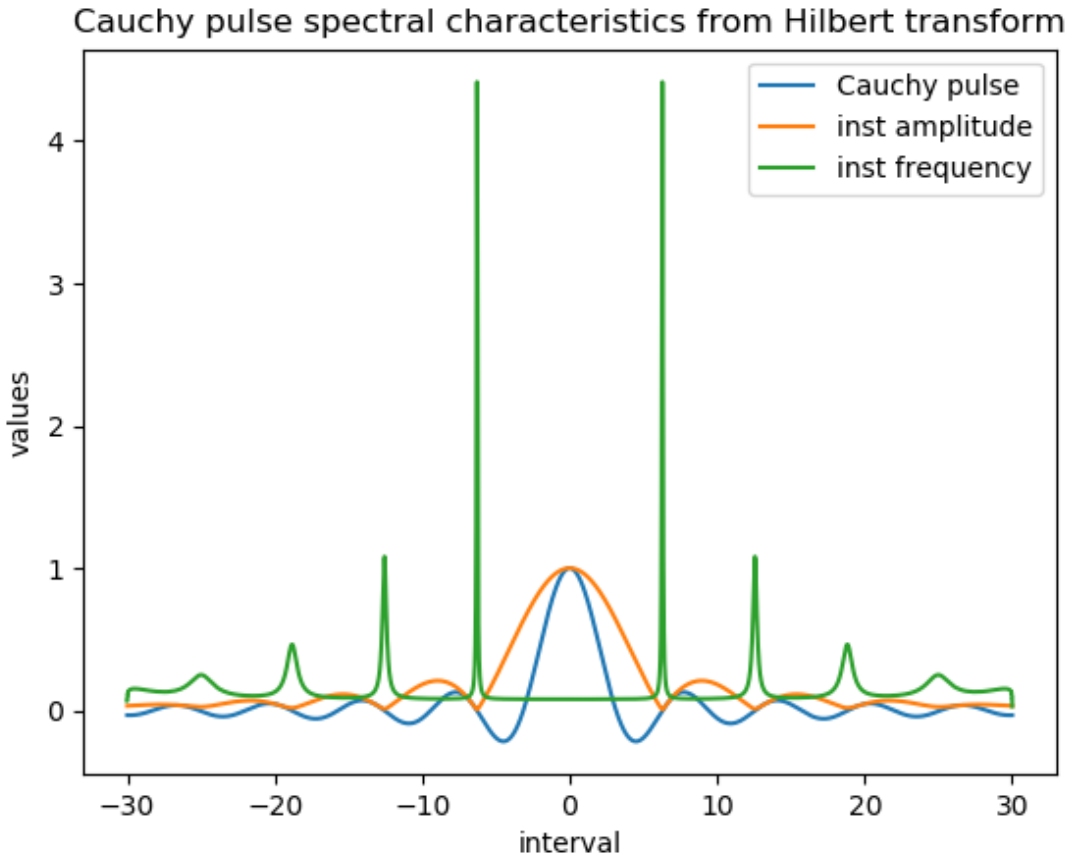

Figure 2.5: Cauchy pulse with its instantaneous (inst) amplitude and frequency. $c = 1$

Figure 2.5 shows the Cauchy pulse, which is given by ($c = 1$)

$$\frac{c}{c^2 + t^2}, \tag{2.34}$$

as well as its instantaneous (inst) amplitude and frequency on the interval $[-30, 30]$. Considering the Cauchy pulse on the time interval $[0, 30]$, its amplitude and frequency decay in time. This is well captured by the instantaneous amplitude and frequency in Figure 2.5.

Le further illustrate the importance of the Hilbert transform for frequency and amplitude modulated signals. This example is taken from (scipy community (2019)), the official page of the Python implementation of the Hilbert transform. It depicts the strength of the Hilbert transform. Consider the signal given in Figure 2.6.

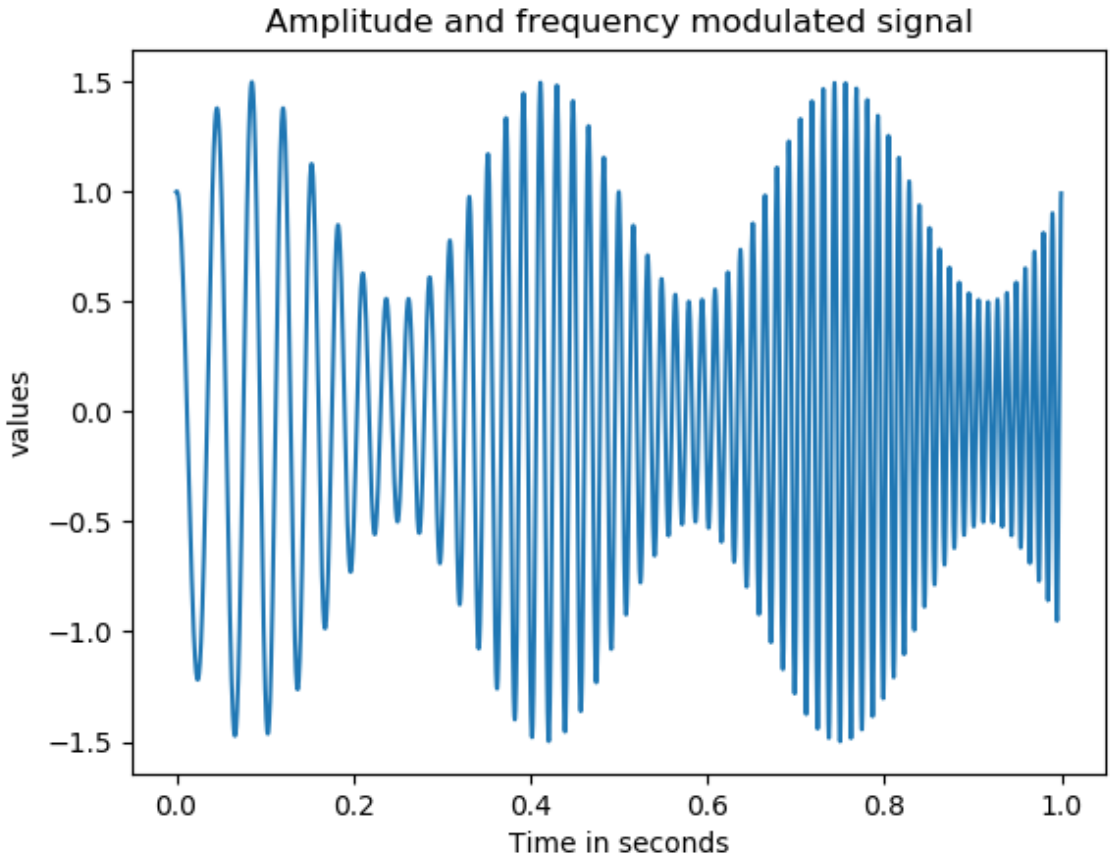

Figure 2.6: A signal with amplitude and frequency variation in time

As one can can observe, its frequency changes from 20 Hz to 100 Hz from 0 to 1 second, while its amplitude varies in time. By computing its Hilbert transform, the signal is extended to the upper complex plane, as an analytic signal. Therefore using (2.28-2.30), its instantaneous frequency and amplitude can be computed.

Figure 2.7 shows the corresponding instantaneous amplitude and frequency. It is clear from this that the frequency changes (nearly) from 20 Hz in the neighborhood of $t = 0$, to 100 Hz in the neighborhood of $t = 1$. This demonstrates the efficiency of the Hilbert transform in dealing with signals with modulated spectral characteristics.

Although the Hilbert transform provides a tool to analyze signals with modulated spectral characteristics, the instantaneous frequency obtained does not always provide a physical meaning (Huang et al. (1998), Huang and Wu (2008), unless the input signal is mono component (single frequency characteristics).

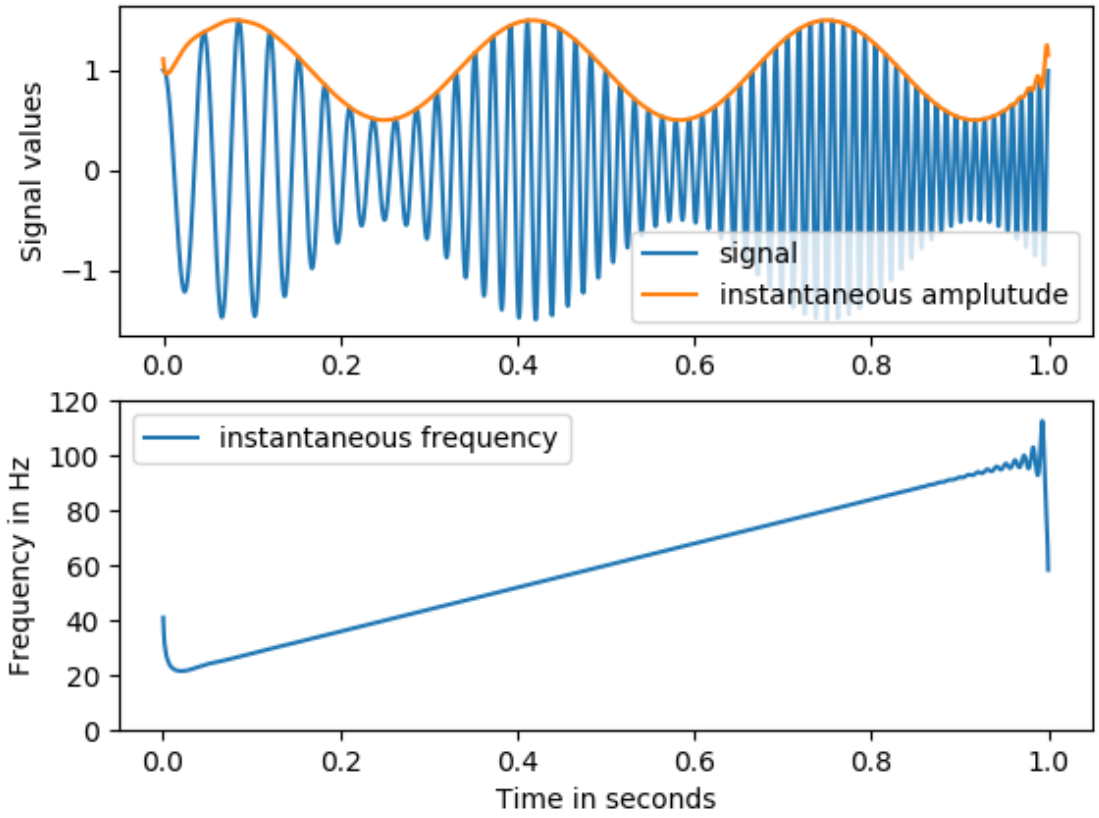

Figure 2.7: A signal with amplitude and frequency variation in time

To address the issue faced by the Hilbert transform for non mono component signals, the Hilbert-Huang transform was developed. Through the empirical mode decomposition (EMD), which is a series of transformation, a target signal is decomposed into (nearly) mono component signals called intrinsic mode functions, for which instantaneous amplitude and frequency are well defined in a physical sense. Therefore, the Hilbert-Huang transform adds value to the Hilbert transform, and both make it possible to correctly tackle problems with frequency and amplitude modulation. A concise exposition of the Hilbert-Huang transform is given in chapter 3, where it is applied in a proposed new method, for bearing fault detection.

The Hilbert transform acts as a linear filter in the frequency domain, where it imposes a phase shift to every Fourier component. In the time domain, it extends a real valued function to a complex function which is an analytic signal. The imaginary part of the latter is the Hilbert transform of the original signal. By using the analytic signal, the instantaneous frequency and the instantaneous amplitude of the original signal is derived. To have a meaning instantaneous amplitude and frequency, the original signal must be a mono component signal. The Hilbert-Huang trans-

form, makes it possible to decompose any signal to a set of (nearly) mono components called intrinsic mode function. Thus, the Hilbert transform, coupled with the Hilbert-Huang transform provide a complete set of tools, in analyzing frequency and amplitude modulated signals.

2.2 Application of the high frequency resonance technique for bearing fault detection

In the literature review section, the high frequency resonance techniques (HFRT) was presented as one of the most widely used method for bearing fault detection. In this section, the HFRT is applied to a case study, which consists of detecting defects in the outer and inner ring of a roller bearing.

Section 2.2.1 describes the experimental set up that generated the bearings vibration signals for this case study. In section 2.2.2 the theoretical equations for bearing defects frequencies are formulated. Furthermore, the characteristics of each fault are described. Finally, the results from applying the high frequency resonance technique are presented in section 2.2.3.

2.2.1 Description of the case study

The data for this case study was generated by the Intelligence Maintenance System (IMS), at the University of Cincinnati. It is freely available online and was measured by (Lee et al. (2007)). It has been the subject of numerous publications such as (Hai et al. (2006), Mejia et al. (2010), Fangtao et al. (2011), Mejia et al. (2011), Mortada and Yacout (2011), Rego et al. (2011), Yacout (2012), Sergey and Zigmund (2012) , OF et al. (2012), Jianbo (2012a), Jianbo (2012b), Mejia et al. (2012))

Three separate experiments involving four bearings were performed on a motor. In each experiment, a one second vibration signal snapshot was recorded every five or ten minutes. Each vibration signal (sample), consists of 20 480 data points with a sampling rate of 20 000 Hz. The sampling rate also called sampling frequency, is the average number of sample points obtained per second during the sampling process. The sampling process is the reduction of an analog signal to a digital signal. The analog signal is the continuous time signal, while the digital signal is the discrete time signal of interest. The experimental set up is shown in Figure 2.8 where

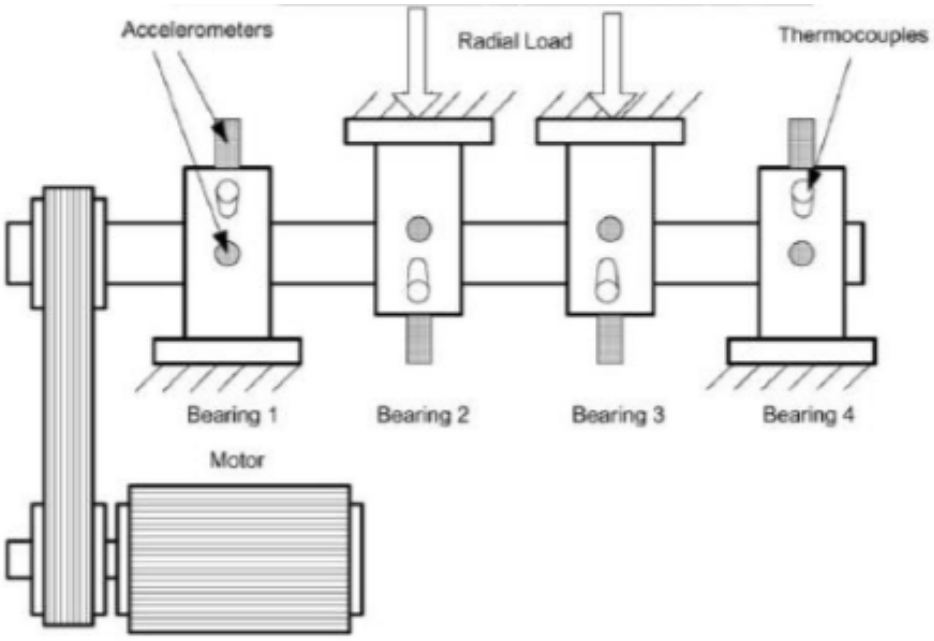

Figure 2.8: Experimental set up.

four bearings support a rotating shaft, driven by an AC motor through rub belts. The motor rotates at 2000 rotations per minute, while the bearings are subjected to a constant radial force of 6000 pounds. In the first experiment two accelerometers were mounted on the bearings. Each measuring vibration for the axial and radial position respectively. In the second and third experiment, four accelerometers were placed one each bearing, measuring axial vibration. Recall that an accelerometer is a sensor that measures vibration signals, in particular the acceleration experienced by the bearings in this case. Axial signals are measured along the axis of the motor shaft, while radial signals are measured along the perpendicular direction of the motor shaft.

At the end of the first experiment, an inner race defect occurred in bearing number 3 and a roller element defect occurred in bearing number 4. At the end of the second experiment, an outer race defect occurred in bearing number 1 and 3, respectively.

2.2.2 Bearing defects

Bearing failure frequencies depend on their physical characteristics as well as the rotating speed of the motor housing them. Figure 2.9 shows a graphical representation of a bearing, with the outer ring, the cages that encircle the balls (rolling elements), and the inner ring, where the motor shaft is attached. A bearing failure is instantiated in four different ways, each related to its parts : a fault in the inner, and outer ring, and defects in the cages and roller elements. As part of this section, only fault occurring in the inner ring and outer ring are considered.

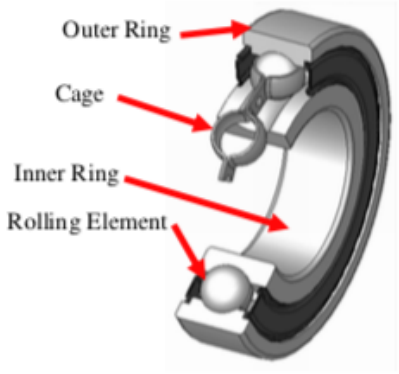

Figure 2.9: Geometrical view of a bearing

An inner ring defect occurs at a frequency called ball pass frequency inner race (BPFI), while an outer race defect occurs at a frequency named ball pass outer race frequency (BPFO). They are expressed in Hz in terms of the bearing geometry characteristics and the shaft rotation speed as

$$BPFI = \frac{nb}{2} S \left(1 + \frac{BD}{PD} \cos(\beta)\right) \tag{2.35}$$

$$BPFO = \frac{nb}{2} S \left(1 - \frac{BD}{PD} \cos(\beta)\right) \tag{2.36}$$

where

- S is the rotating speed of the motor
- nb is the number of cylindrical balls (roller elements)
- BP is the the ball diameter
- PD is the pitch diameter
- β is the contact angle

The pitch diameter is the perpendicular distance from the center of one ball to the center of the ball located at the end. The bearing type used in this case study are manufactured by the company Rexnord, and are of type Rexnord ZA 2115, with BPFO and BPFI given by 236.4 Hz, 296.8 Hz, respectively.

2.2.3 The high frequency resonance technique application to bearing failure detection

The high frequency resonance technique is made of series of transformations aiming at, in one hand signal filtering, and on the other hand detecting incipient bearing failure. The latter relies on the former, as it reveals hidden failure characteristics. the hight frequency resonance technique is made of four basic steps: The input vibration signal is bandpassed, Hilbert transformed, low passed and finally fast Fourier transformed. Figure 2.10 shows a schematic description of the entire process.

The band pass filter process allows signals within a prescribed frequency band to seep through, while attenuating non essential components. This passband comprises typically bearing failure frequencies. The band of allowed frequencies is bounded bellow and above by two parameters: The low cut-off frequency and the high cutoff frequency, respectively. In this application a low cutoff frequency of 2000 Hz and a high cutoff frequency of 9990 Hz were applied. The Hilbert transform extract high amplitude or envelop of the bandpassed signal. Since the Bearing frequencies are within the lower end of the frequency spectrum, the Butterworth lowpass filter is applied to capture them. The Fast Fourier transform is then applied to transform the resulting time signal to its corresponding frequency spectrum, where potential bearing failure frequencies can be observed

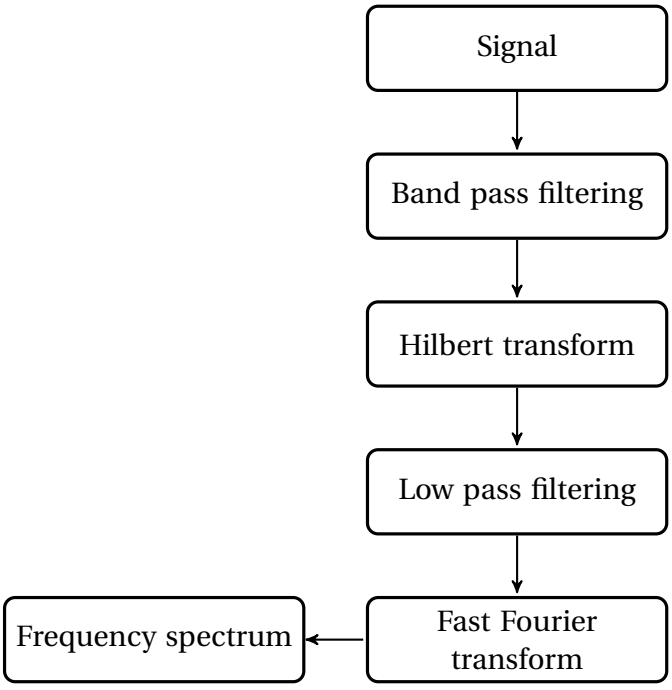

Figure 2.10: Description of the high frequency resonance technique.

Ball pass outer race(BPFO) defect frequency detection

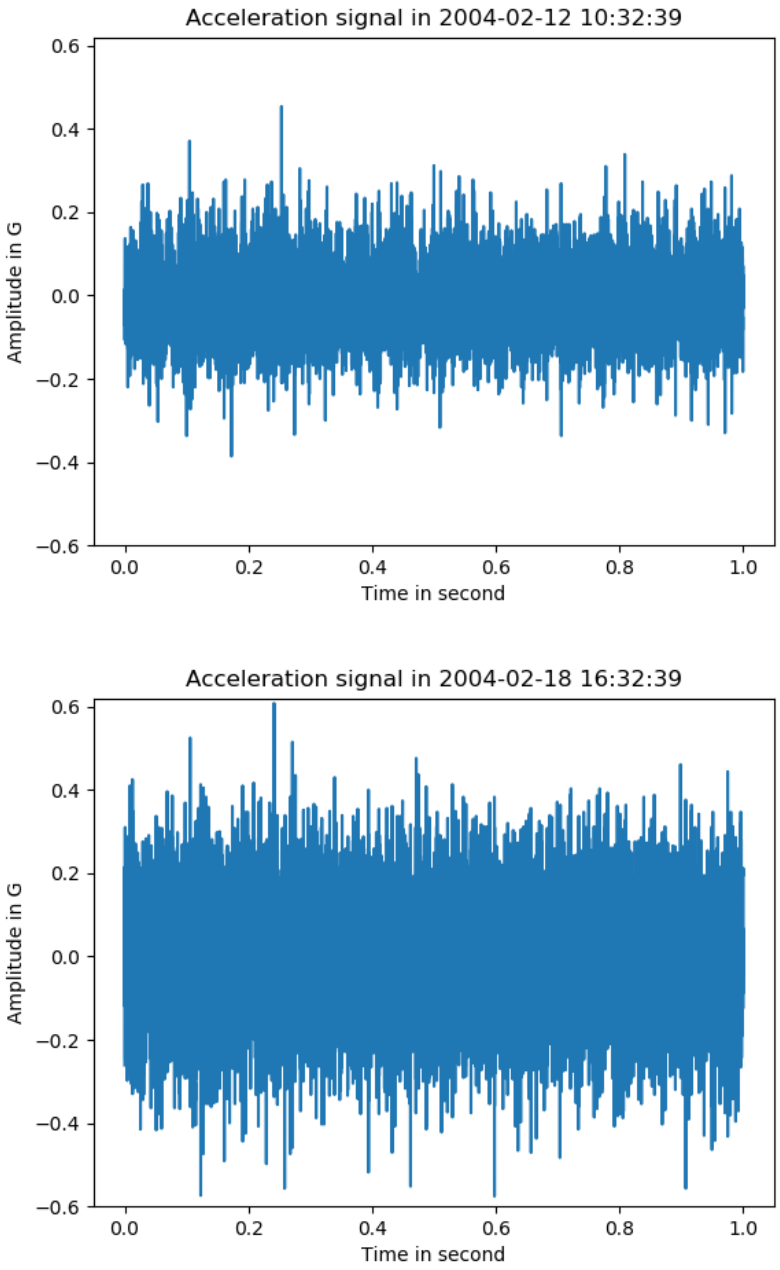

Figure 2.11: Vibration time signals of bearing number 1 recorded at the beginning (top) and six days after (bottom), in experiment 2.

In experiment 2, four bearings were run to failure, resulting in outer race defect in bearing number 1. Figure 2.11 shows the vibration time signals of bearing number 1, at the beginning (top) and six days after (bottom), from experiment number 2. A significant increase in the variance

can be observed. In fact, the variance has increased 28 time after 6 days. This can be seen in Figure 2.12, which shows a density plot at the beginning and six days after.

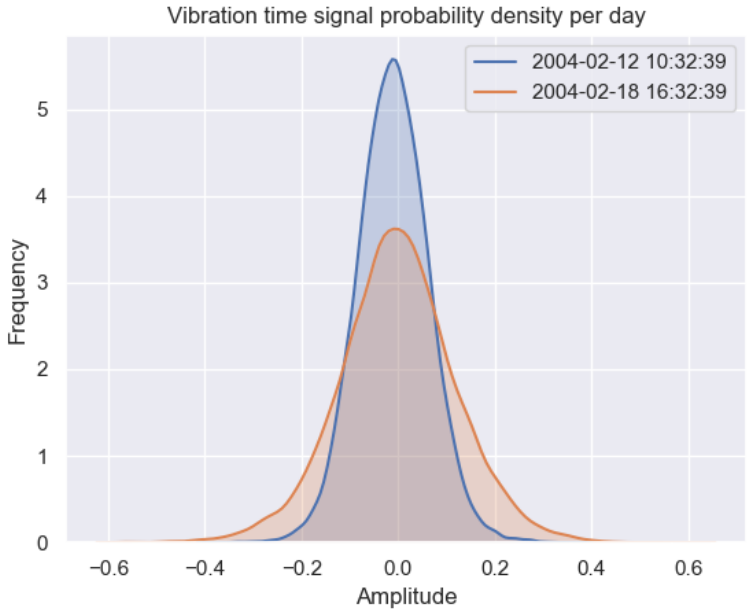

Figure 2.12: Density plot of the time signal at the beginning and six days after

After six days, the tail of the distribution increases considerably, which indicates an increase in the variance, due to bearing defect propagation. Figure 2.13 shows the corresponding frequency spectrum at the beginning (top) and six days after (bottom). At the beginning of the experiment, the frequency spectrum is relatively "clean". No high frequency peaks can be observed. This indicates that the bearing is relatively "healthy". However, after six days, the ball pass frequency as well as harmonics can be observed in the frequency spectrum.

Harmonics are evenly spaced frequency peaks that are integer multiple of the ball pass outer race defect frequency (Mobius (2014)). They are indication of transient, clipping or random impact in the vibration signal, and are characteristics of the presence of bearing defects (Mobius (2014)). Transient indicates the variation of the signal properties with time, while clipping refers to the distortion of a signal.

Figure 2.14 shows the evolution of defect frequency in terms of amplitude as a function of time, for all bearings. Clearly, bearing 1 exhibits high amplitude in terms of defect frequency compare to other bearings. Its amplitude increases and drops, before increasing in an oscillatory fashion. This is due to random slip or clipping caused by defect propagation in the bearing. This illustrates the fact that the vibration signal generated by bearing are non periodic and stochastic (Randall and Antoni (2011)).

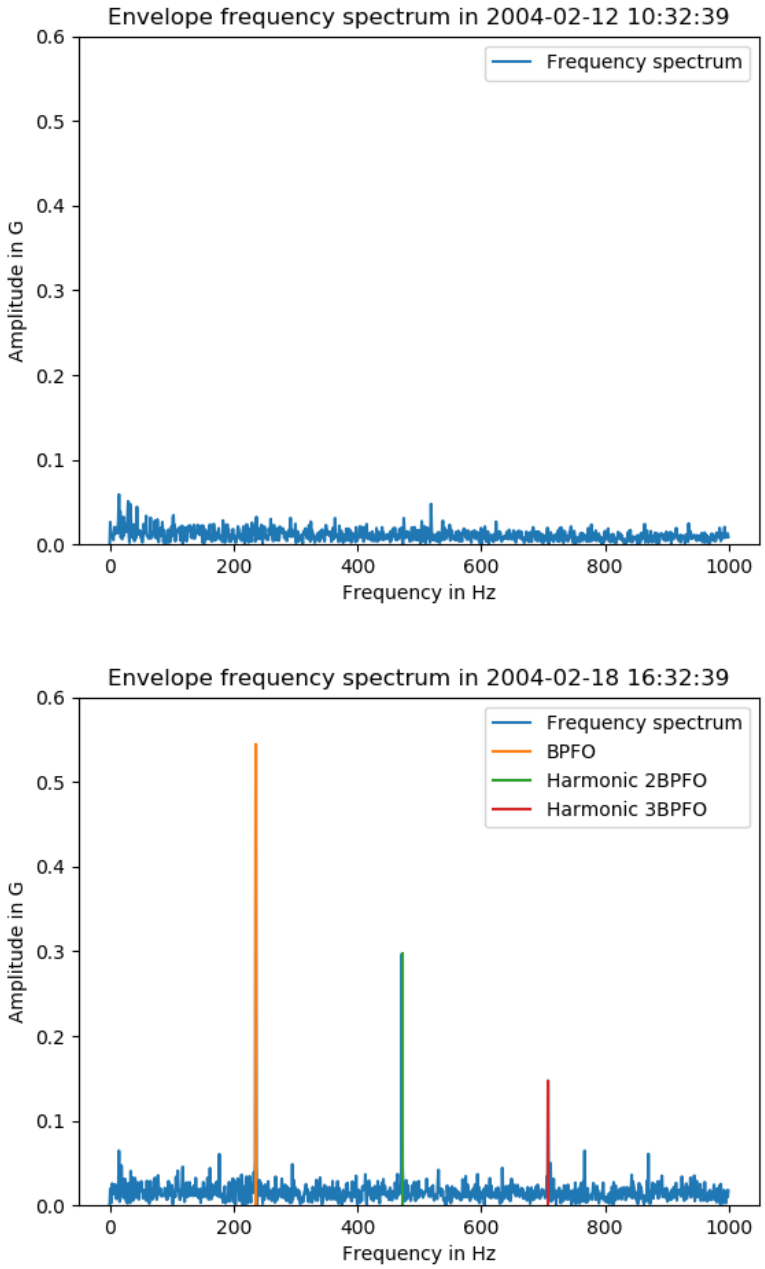

Figure 2.13: Frequency spectrum with the presence of outer race defect frequency and harmonics.

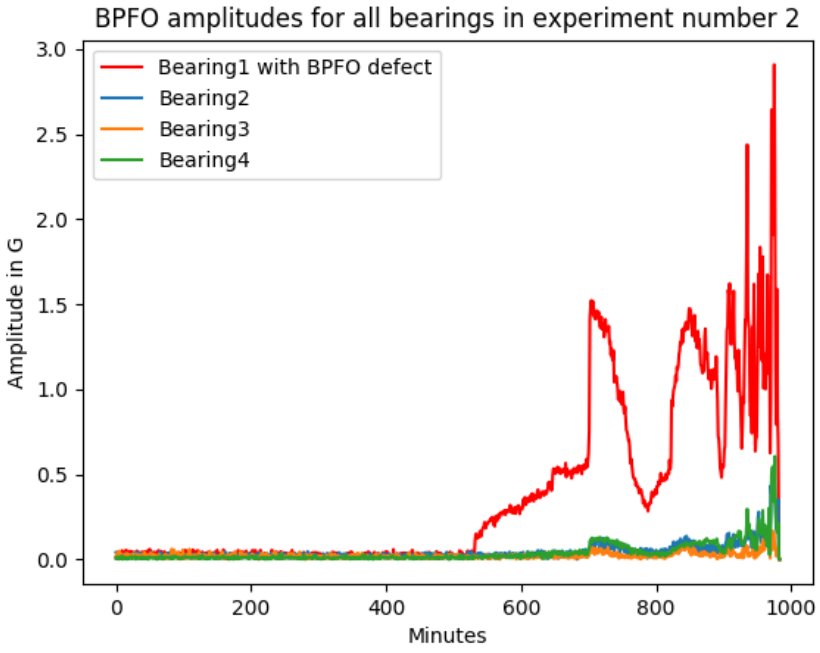

Figure 2.14: BPFO amplitude evolution for all bearings over time

Ball pass inner race (BPFI) defect frequency detection

Recall that in experiment 1, four bearings were run to failure, and at the end of the experiment, inner race defect occurred in bearing number 3. Figure 2.15 shows the vibration signals of bearing number 3, at the beginning (top) and at the end (bottom) of the experiment.

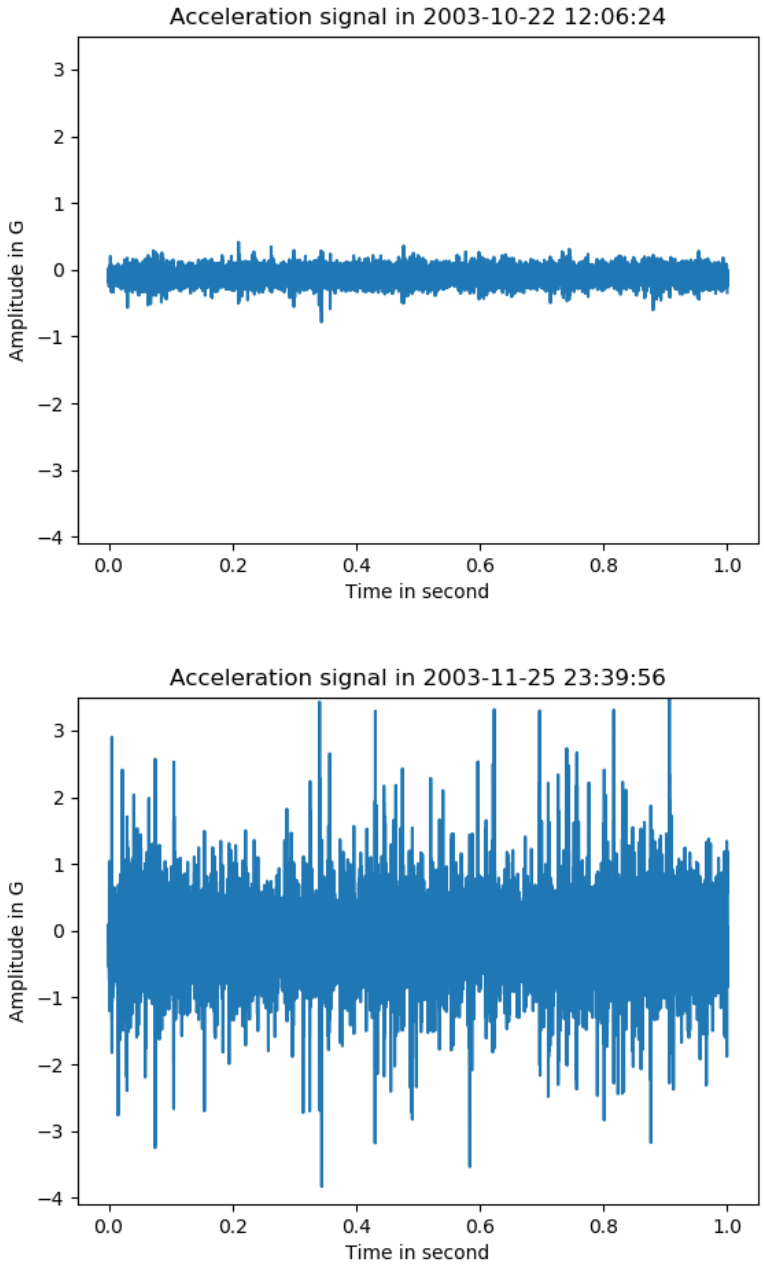

Figure 2.15: Vibration time signals recorded at the beginning (top) and at the end (bottom), of experiment 1.

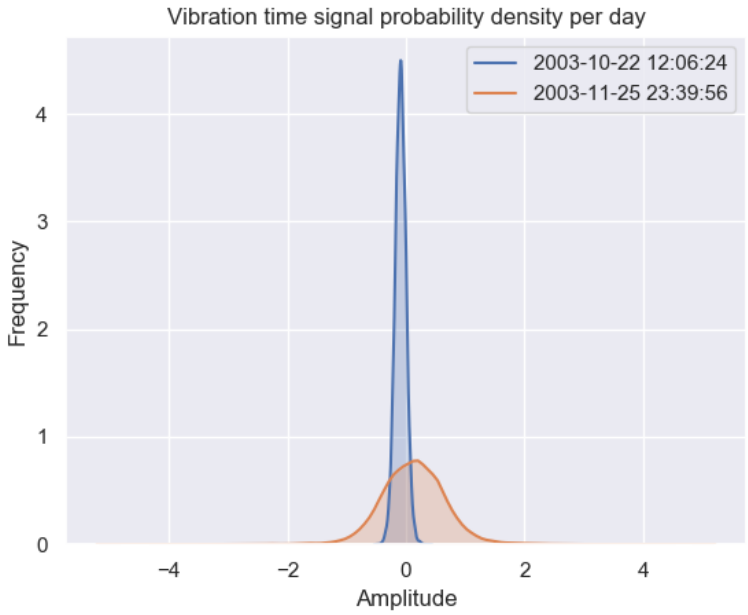

Figure 2.16: Density plot of the time signal at the beginning and at the end of the experiment 1, for bearing number 3.

It is clear that the variance of the time signal had increased significantly. This can be seen from the density plot in Figure 2.16, where the distribution of the signal at the end of the experiment exhibits a pronounced wide tail. In fact, the variance had increased by 41 times its initial value. The frequencies spectrum at the beginning (top) and at the end (bottom) of the experiment can be seen in Figure 2.17. The frequency spectrum at the beginning of the experiment does not contain any high peaks frequencies. It is completely "clean". This means that the bearing is relatively "healthy".

However, at the end of the experiment, the frequency spectrum exhibits a ball pass inner race frequency and form sidebands together with surrounding peaks. Sidebands are series of (almost) evenly spaced frequency peaks, centered around a center peak called carrier (Mobius (2014)). They are result of amplitude modulation between the shaft frequency and the ball pass inner race defect frequency (Mobius (2014)). This is due to the fact that the inner race fault will rotate in and out of the loaded region, causing modulation of the impact amplitudes by the rotation speed of the shaft (Courrech and Gaudet).

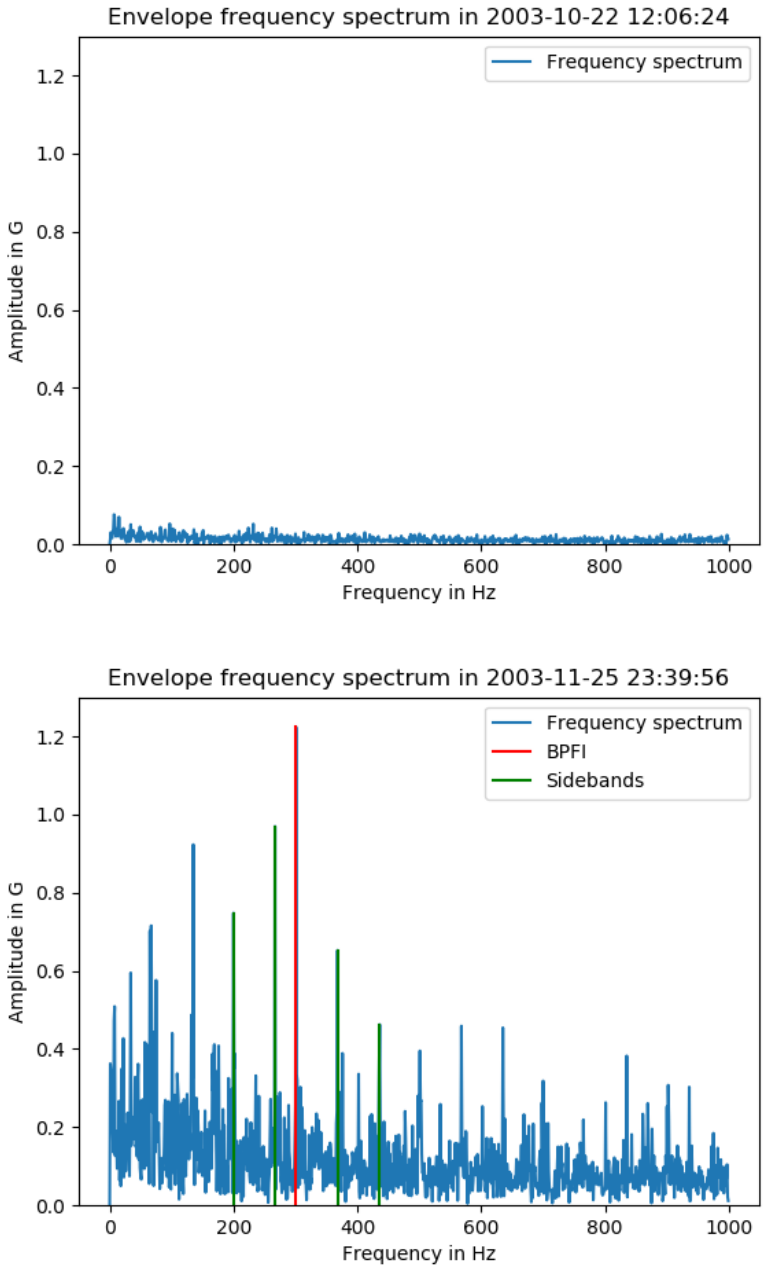

Figure 2.17: Frequency spectrum with the presence of inner race defect frequency and sidebands.

The high frequency resonance technique applied to inner race defect detection, can produce a confusion spectrum, with pattern of spectral line changing (McFadden and Smith (1984a)). This make identifying inner race defect difficult in situation where the defect is pronounced. In those cases the defect frequency values can vary up to 2% (Randal and Antoni (2010)). Although the rotating speed of the machine is set as a constant value, in practice however, there is always small variations, that will also generate a variation in the theoretical value of the frequency de-

fect. As can be seen from Figure 2.17, the frequency spectrum is some how noisy. This is one of the characteristics of the inner race defect when the defect is pronounced ((McFadden and Smith (1984a))). A noisy spectrum can render difficult the detection of failure frequencies. In the next chapter a new method is presented in order to tackle the issue of noisy spectrum.

2.3 Summary

Fourier analysis applied to bearing fault detection, is concerned with transforming the vibration time signal of a bearing, to its corresponding frequency spectrum. A defect generates an impulse, which induces resonance in the bearing, and the machine housing the bearing. The generated resonance represents the response of the former and the latter.

The diagnostic information for bearing fault detection resides in the high frequency resonance induced by defects. The latter are frequency components, expressed mathematically in terms of the bearings physical characteristics, and the rotational speed of the machine shaft. The most prevalent technique used to extract bearing diagnostic information is the high frequency resonance technique.

The latter uses a series of filtering operations, aiming at isolating the bearing signal containing the failure frequencies. Furthermore, the frequency spectrum is extracted from the time signal through Fourier transform. If defects are initiated, the frequency spectrum will contain the defects frequencies. Although efficient, the high frequency resonance technique is only valid within the boundary of the assumptions upon which it is predicated.

Due to slip of the roller elements in the bearing, the failure frequencies values can vary up to 2% from their theoretical values, or even more. In addition, bearing signal are relatively non periodic and stochastic in nature. Consequently, non-stationarity and non linearity can be observed in bearing signal. When inner race defects are pronounced, the frequency spectrum can contain a considerable amount of noise, rendering difficult failure detection. The need for alternative methods for bearing fault detection is therefore justified. In the next chapter, a new method which addresses some of the aforementioned issues is proposed.

Chapter 3

Hilbert-Huang transform for bearing fault detection

Most signal decomposition methods such as Fourier transform, impose a-priory basis functions on the signal to be analyzed. In Fourier analysis, the basis functions are trigonometric extensions. Although this implies a rigorous mathematical treatment, the resulting signal decomposition is limited by the mathematical assumptions (Huang et al. (1998), Huang and Wu (2008)). Limiting not in the sens of its mathematical truthfulness, rater in its ability to capture all intended salient physical properties of the target signal. Two such assumptions are linearity and stationary. As most phenomena in nature are nonlinear and non stationary, this mathematical approach, although rigorous, lacks an important property: Adaptivity (Huang et al. (1998), Huang and Wu (2008)). The latter refers to capturing the intrinsic properties of a signal, without imposing a-priory basis functions, (Huang et al. (1998), Huang and Wu (2008)).

The Hilbert-Huang transform (HHT) was precisely developed to deal with nonlinear and non stationary processes, in an adaptive fashion (Huang et al. (1998)). It combines Hilbert spectral analysis with the so call empirical mode decomposition (EMD), to adaptively decompose a signal into its fundamental components called intrinsic mode functions (IMF), (Huang et al. (1998)). The richness of the HHT spans from the analysis of differential equations, to the study of geophysical phenomena, as well as bearings faults detection (Huang and Wu (2008),Li et al. (2009a), Yan and Gao (2006) ,Soualhi et al. (2015), Sallo and Grif (2019)), and by no mean limited to them.

The HHT has been successful applied to the analysis of solutions of nonlinear differential equations, such as the Duffing and the Lorentz equations. The intrinsic frequency, the forcing function and the low intensity subharmonics of the numerical solutions for the nonlinear Duffing equation has been extracted through the EMD process, (Huang et al. (1998)).

The decomposition of the solutions of Lorentz equation, revealed “transient components with different frequencies and damping characteristics”, which agreed with previous studies, (Huang et al. (1998)). The HHT application to seismic waves propagation, identified high and low frequency seismic waves (Vasudevan and Cook (2000)). In particular, its decomposition of the seismic waves induced by the 1999 Taiwan earthquake, revealed that “the Fourier transform underestimated low frequency energy” (Huang et al. (2001)). The Hilbert-Huang transform emerged as a general signal decomposition tool, and in theory can be applied to any signals.

In this chapter, a new method based on Hilbert-Huang transform (HHT) in part, is postulated and tested, with signals generated by diverse bearings mounted on a motor. The novel technique, couples the HHT with a robust seasonal trend decomposition method called STL, to detect bearing faults. STL stands for seasonal trend decomposition based on LOESS. In short, the STL decomposes a target signal into a trend and oscillatory components also called seasonality. The trend is a monotone curve, while the seasonal components are periodic oscillations with constant period.

The bulk of the proposed scheme, consists of decomposing a target signal into (nearly) mono components signals also called intrinsic mode functions (IMFs). Furthermore, The seasonal trend of each IMF is extracted trough the seasonal trend decomposition based on LOESS. Moreover, the power spectral density of the resulting seasonal components are computed. The idea is that the resulting power spectral density, which is the energy distribution per frequency contribution, will encompass bearing failure if any.

the remaining parts of this chapter are organized as followed: Section 3.1 presents the empirical mode decomposition (EMD), which is the back bone of the Hilbert-Huang transform for decomposing a signal adaptively. Section 3.2, through a concrete case study, demonstrates the efficiency of the novel approach. It begins with a gentile reminder of bearing faults characteristics, followed by a description of the experimental setup, that generated the data for this case study. Afterwards, the seasonal trend decomposition based on LOESS (STL) is presented, followed by a schematic description of the new scheme. Finally, the results of the case study are explained. In closing, section 3.3 gives a short summary of this chapter.

3.1 The Hilbert-Huang transform

The Hilbert-Huang transform is the amalgam of the empirical mode decomposition (EMD) and the Hilbert transform. The former decomposes a target signal into components having the same spectral characteristics, and identified as intrinsic mode functions (IMFs), while the latter enables the extension of a real valued signal to its complex counterpart. Together, the empirical

mode decomposition and the Hilbert transform, form the basis for accurately (in the physical and mathematical sens), analyzing amplitude and frequency modulated signal (Huang et al. (1998)).

As will be shown subsequently, the empirical mode decomposition (EMD), can be (almost) interpreted as a filter bank. The EMD can be regarded as a process that maps a target signal frequency range, to each intrinsic mode function (IMF). That is, higher frequency regions will correspond to IMFs with high frequency oscillation, while lower frequency regions will correspond to lower frequency IMFs. The empirical mode decomposition is iterative in nature.

For a target signal $s(t)$, the goal is to obtained its n fundamental parts (IMFs), denoted here by $s_j, j = 1, \cdots, n$, such that

$$s(t) = \sum_{j=1}^{n} s_j(t) + r(t), \tag{3.1}$$

where $r(t)$ is the residual, which is either a constant or a monotone function. The derivation of the intrinsic mode functions s_j, relies on the following key assumptions.

Definition 2 *An intrinsic mode function (IMF) must satisfy the following conditions:*

1. *The number of extrema and the number of zero crossing must either equal or differ by one*
2. *At any data point, the mean value of the envelope defined using the local maxima and the envelope defined by using the local minima is zero.*

The empirical mode decomposition as an iterative algorithm, generates intrinsic mode functions, each satisfying definition 2 and obtained as follow:

1. Compute the upper and the lower envelope curve of the target signal $s(t)$
2. In the first iteration ($i = 1$), compute the mean $m_i(t)$ between the upper and the lower envelope cure of $s(t)$
3. Compute the first (pseudo) IMF as

$$h_{i1}(t) = s(t) - m_i(t). \tag{3.2}$$

   If h_{i1} satisfies definition 2, then it is an IMF and it is denoted by

$$c_i(t) = h_{i1}(t). \tag{3.3}$$

   otherwise

4. Set $h_{i1}(t)$ as the input signal and repeat step 1,2 and 3, k time, until $h_{ik}(t)$ satisfies definition 2.

5. After finding the first IMF, set the first IMF as input signal and repeat step 1,2,3 and 4 to obtain the remaining IMFs. If an IMF is monotone, set it as the residual and you are done.

In most function (signal) decomposition methods, a set of predefined n basis functions denote here by $\varphi_j(t)$, $j = 1, \cdots, n$, coupled with coefficients a_j, are used to decomposing a function, say $f(t)$ as

$$f(t) = \sum_{j=1}^{n} a_j \varphi_j(t). \tag{3.4}$$

However, the Hilbert-Huang transform decomposes $f(t)$ as

$$f(t) = \sum_{j=1}^{n} c_j(t) + r(t). \tag{3.5}$$

In equation (3.4) the basis $\varphi_j(t)$ are chosen before hand (a-priory), and the coefficients are computed through an integral or summation operation. A kind of "bias" is imposed on the the function $f(t)$. A change of basis function will affect the resulting decomposition. On the other hand, in equation (3.5), the basis functions are directly derived from the properties of the function $f(t)$. This illustrates the concept of adaptivity, central to the Hilbert-Huang transform, which can be regarded as basis function agnostic.

Consider the signal

$$s(t) = \sin(5\pi t) + \sin(10\pi t) + \sin(100\pi t), \tag{3.6}$$

given in Figure 3.1. It is made of one high frequency component (with angular frequency 100π) and two relatively low frequency components, with angular frequency 5π and 10π, respectively. By applying the empirical mode decomposition, the three components are extracted and can be seen in Figure 3.2. The first intrinsic mode function corresponds to the highest frequency component in the original signal (Figure 3.1), while the second and third intrinsic mode functions correspond to the lower frequency components. This justifies the assertion made earlier, comparing the empirical mode decomposition to a filter bank.

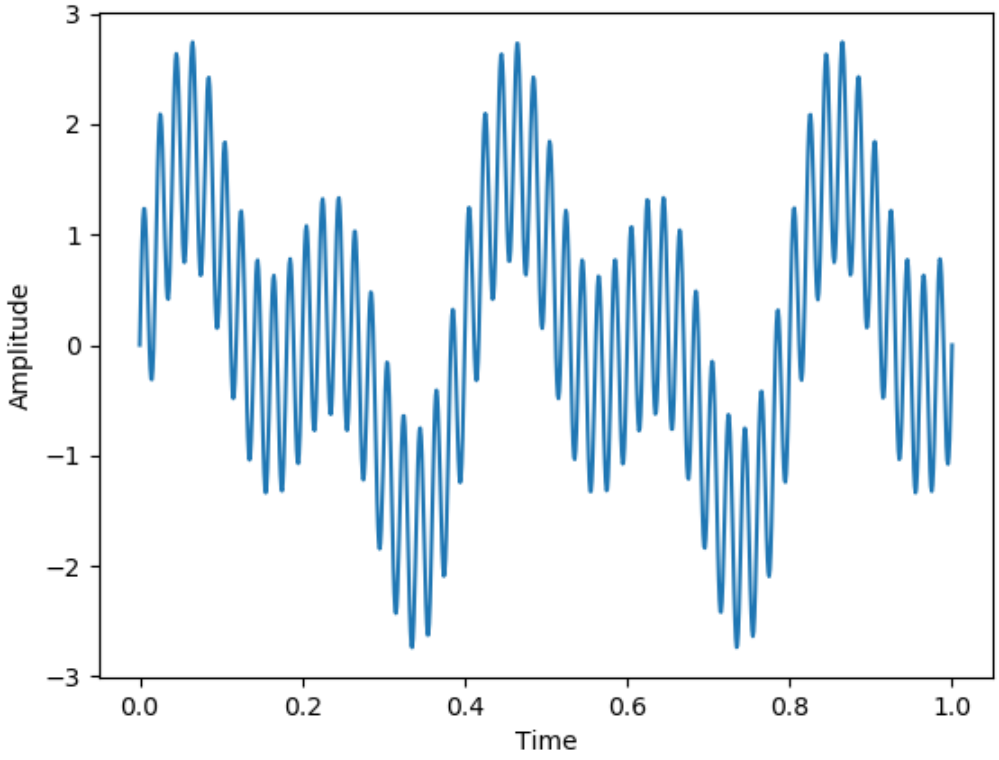

Figure 3.1: An input signal $s(t) = \sin(5\pi t) + \sin(10\pi t) + \sin(100\pi t)$, with three frequency components

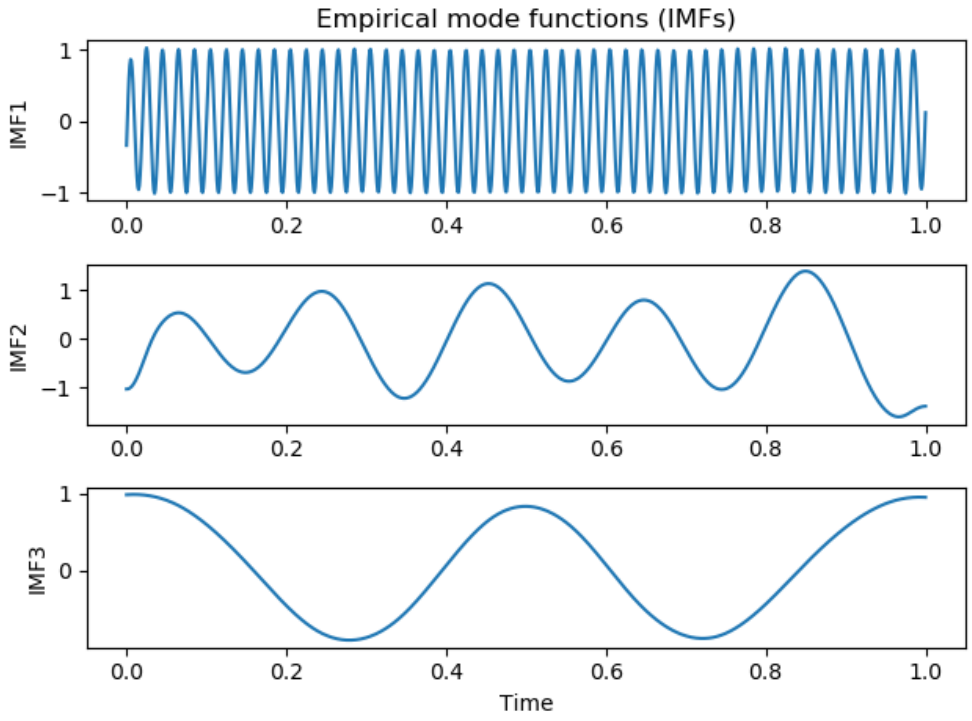

Figure 3.2:

The intrinsic mode functions obtained through the EMD process, constitutes an adaptive basis that satisfies the mathematical properties of convergence, completeness, orthogonality and uniqueness, (Huang et al. (1998)). Furthermore, if $a_j(t)$ and $\omega_j(t)$ are amplitude and frequency modulation corresponding to IMF j, then the original signal $s(t)$ can also be recovered as

$$s(t) = \Re\left(\sum_{j=1}^{n} a_j(t)\exp\left(i\int \omega_j(t)\mathrm{dt}\right)\right), \tag{3.7}$$

where the symbol $\Re(\cdot)$ represents the real part of the expression its encompasses, $i = \sqrt{-1}$, and n is the total number of IMFs obtained from decomposing a signal $s(t)$. Recall that the amplitudes $a_j(t)$ and the frequencies $\omega_j(t)$ can be computed through the Hilbert transform. An analog representation of equation (3.7) in terms of Fourier expansion would be

$$s(t) = \Re\left(\sum_{j=1}^{n} a_j\exp\left(i\omega_j\right)\right), \tag{3.8}$$

where this time, the amplitude a_j and the frequency ω_j are constant. The Hilbert-Huang transform (HHT) offers two different approaches to recover a decomposed signal. The first one is described by equation (3.5) and includes the IMFs, while the second approach is given by equation

(3.7) and includes the instantaneous amplitude and the instantaneous frequency. This gives the HHT a leverage over Fourier transform in nonlinear an non stationary data analysis.

The intrinsic mode functions resulting from the empirical mode decomposition, are not exactly mono components. That is, their spectral characteristics are not uniform in terms of their frequency content. This can be seen in Figure 3.4, which displays the intrinsic mode functions of an input vibration signal, given in Figure 3.3. The IMF number six (6), although representing a high frequency region of the original signal, displays bands of relatively low frequencies area. This mixture of high and low frequency regions, is called mode mixing.

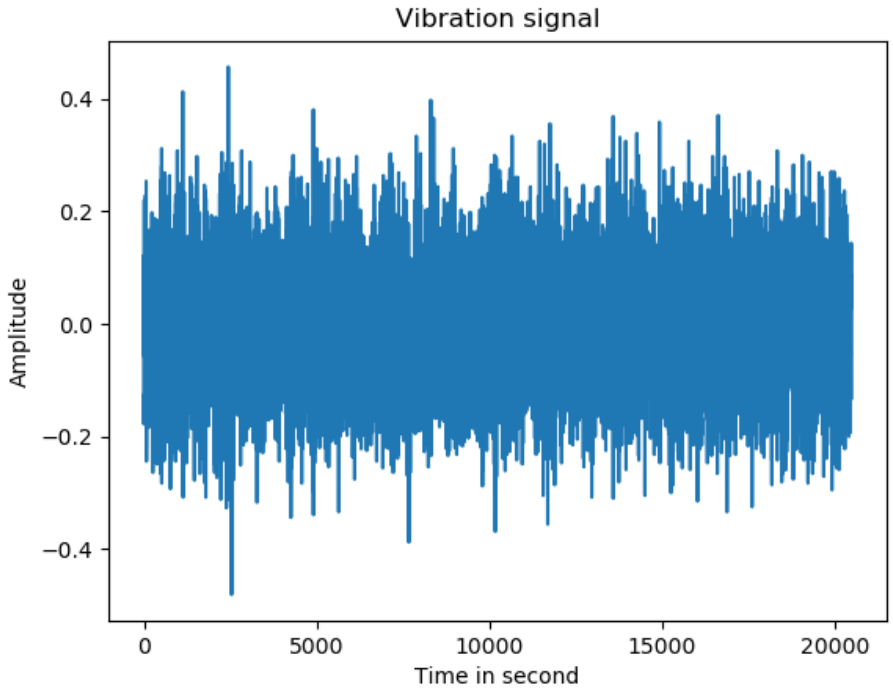

Figure 3.3: An input vibration signal

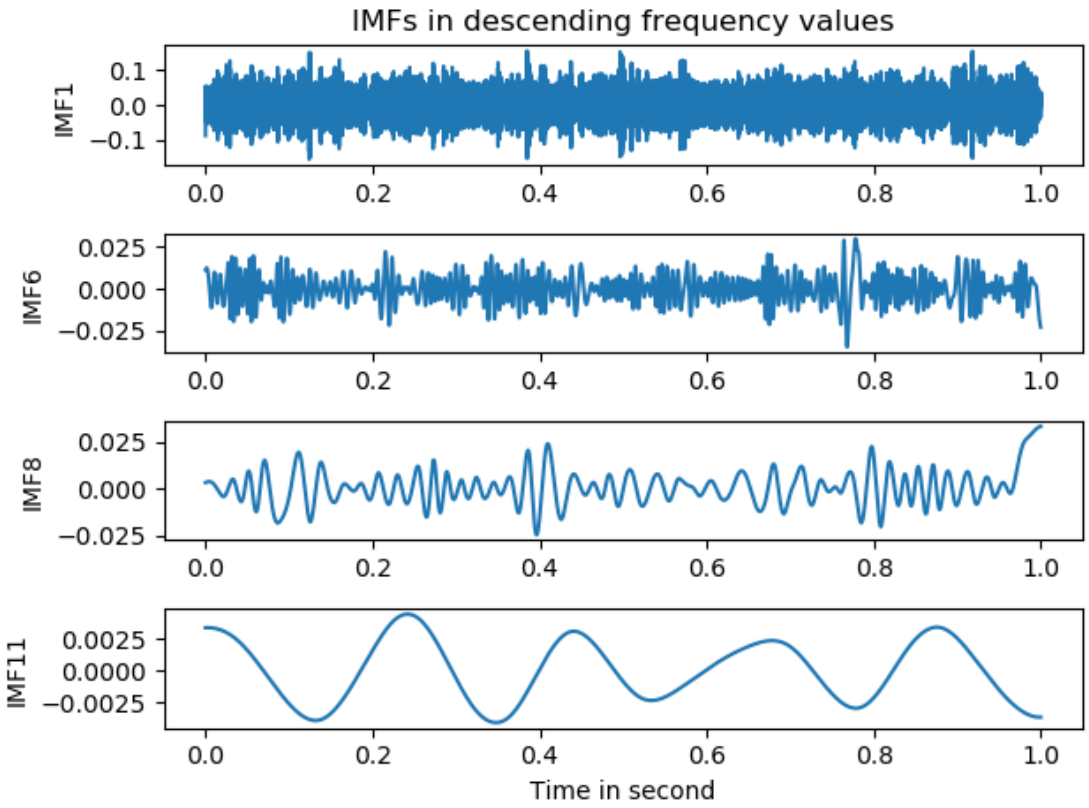

Figure 3.4: Selected intrinsic mode function in descending frequency values.

3.2 Application to bearing fault detection: a case study

In this section, the proposed new method is applied to a case study, in order to demonstrate its ability in detecting bearing failure. As a reminder, Figure 3.5 shows a geometry of a bearing, with different parts.

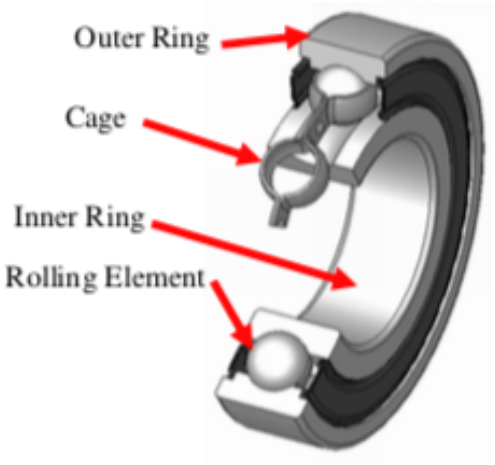

Figure 3.5: Geometrical representation of a bearing

A fault occurring in the outer ring is called a ball pass frequency outer ring (BPFO) defect, while a fault in the inner ring occurs at a frequency called ball pass inner race frequency (BPFI). Both are given in Hz in terms of the bearing geometrical characteristics an the rotating speed of the shaft as

$$BPFO = \frac{nb}{2} R \left(1 - \frac{BD}{PD} \cos(\beta)\right)$$
$$BPFI = \frac{nb}{2} R \left(1 + \frac{BD}{PD} \cos(\beta)\right), \tag{3.9}$$

where R is the rotating speed of the motor on which the bearing is attached. nb is the number of rolling elements (balls), BD is a rolling element diameter. The pitch diameter PD is half the height of the inner ring, and the contact angle β is the angle formed when a rolling element touches the cage.

The data and the experimental setup used for this case study were described in chapter 2. In this setup, four bearings are mounted on a motor rotating at 2000 rotations per minute (33.3 Hz). In the first experiment the motor runs until bearing number 3 is severely damaged with inner race defect, while in the second experiment outer race defect occurs in bearing number 1. For each experiment, successive vibration signals where obtained, with a sample rate of 20 KHz and corresponding Nyquist frequency of 10 KHz (half the sampling rate). The Nyquist frequency define a lower bound limit in order to avoid aliasing, which is a loss of information due to under sampling. That is, if the sample contains frequency components that are higher then the Nyquist frequency, those frequency components wont be able to be seen.

3.2.1 Seasonal Trend decomposition based on Loess (STL)

The STL sequentially applies the locally estimated scatter plot smoother (LOESS) in order to obtain cyclical and trend components of a signal (Cleveland (1979), Cleveland and Develin (1988)). Here a cyclical component is a periodically occurring pattern in a signal. The locally estimated scatter is a non parametric curve fitting procedure. It can be used for "Data exploration, diagnostic checking of parametric models and provides a non parametric regression surface". If $s(t)$ is a signal, the goal is to obtain the decomposition

$$s(t) = T_r(t) + C_y(t) + R_{res}(t), \tag{3.10}$$

where $T_r(t)$ an $C_y(t)$ are the trend and cyclical components, and $R_{res}(t)$ is the residual obtained from subtracting the trend and the cyclical component from the signal. The key ingredient in the STL, is the locally estimated scatter plot smoother procedure. The latter being a non parametric regression method, does not rely on any a-priory assumption on the shape of the curve that needs to be fitted. It provides a flexible approach to curve fitting, by capturing local, as well as global characteristics of a signal. A detailed account of the STL can be found in (Cleveland et al. (1990)).

In this thesis, the STL is applied to address the mode mixing issue discussed earlier. This approach advantage is two folded: In one hand it solves the mode mixing issue by extracting the dominant frequency mode, and on the other hand its generates a clear frequency spectrum, as will be shown later.

3.2.2 Results and interpretation

In this section we apply the empirical mode decomposition followed by the seasonal trend decomposition based on LOESS to extract features containing bearings diagnostic information. Figure 3.6 shows the flow diagram describing the new method for bearing fault detection.

An input bearing vibration signal is subjected to the empirical mode decomposition in order to generate intrinsic mode functions (IMFs). To mitigate mode mixing, the seasonal component of each IMF is extracted through the seasonal trend decomposition based on LOESS. In order to expose potential bearing failure frequencies, the power spectral density of the resulting signal is approximated by the periodogram method.

The power spectral density is technically the power or energy distribution of the autocovariance function (ACF) frequency spectrum. (Stoica and Moses (2004)). The autocovariance function of

a signal $s(t)$ is defined by

$$ACF(s(t)) = E\left[s^*(t)s(t-k)\right] = COV\left(s(t), s^*(t-k)\right) \quad k = 0, 1, \cdots \tag{3.11}$$

where $E(\cdot)$ is the expectation or mean, COV$(\cdot)$ is the covariance and $s^*(t)$ is the complex conjugate of $s(t)$. The covariance of the signal $s(t)$ and its complex conjugate measure their joint variability. Thus the autocovariance function is a sequence of covariance between the signal and its complex conjugate. The power spectrum density (PSD) of a signal s(t) can be defined by (Stoica and Moses (2004))

$$PSD\left(s(t)\right) = E\left[\frac{1}{N}\left|\sum_{t=1}^{N} s(t)e^{-i\omega t}\right|^2\right] \tag{3.12}$$

where N is an integer, $i = \sqrt{-1}$, $E(\cdot)$ is the expected value, and ω is the angular frequency. The estimate $\widehat{PSD}$ of the power spectral density given by

$$\widehat{PSD}\left(s(t)\right) = \frac{1}{N}\left|\sum_{t=1}^{N} s(t)e^{-i\omega t}\right|^2 \tag{3.13}$$

is called the periodogram.

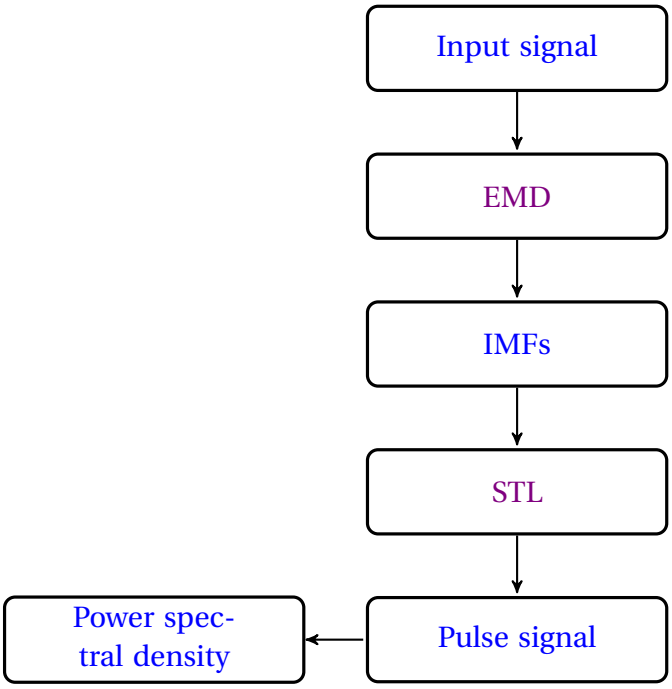

Figure 3.6: Schematic description of the new scheme for bearing fault detection.

Figure 3.7 shows an input vibration signal from a bearing with a defect, its fifth intrinsic mode function, and the resulting seasonal trend extracted through the STL.

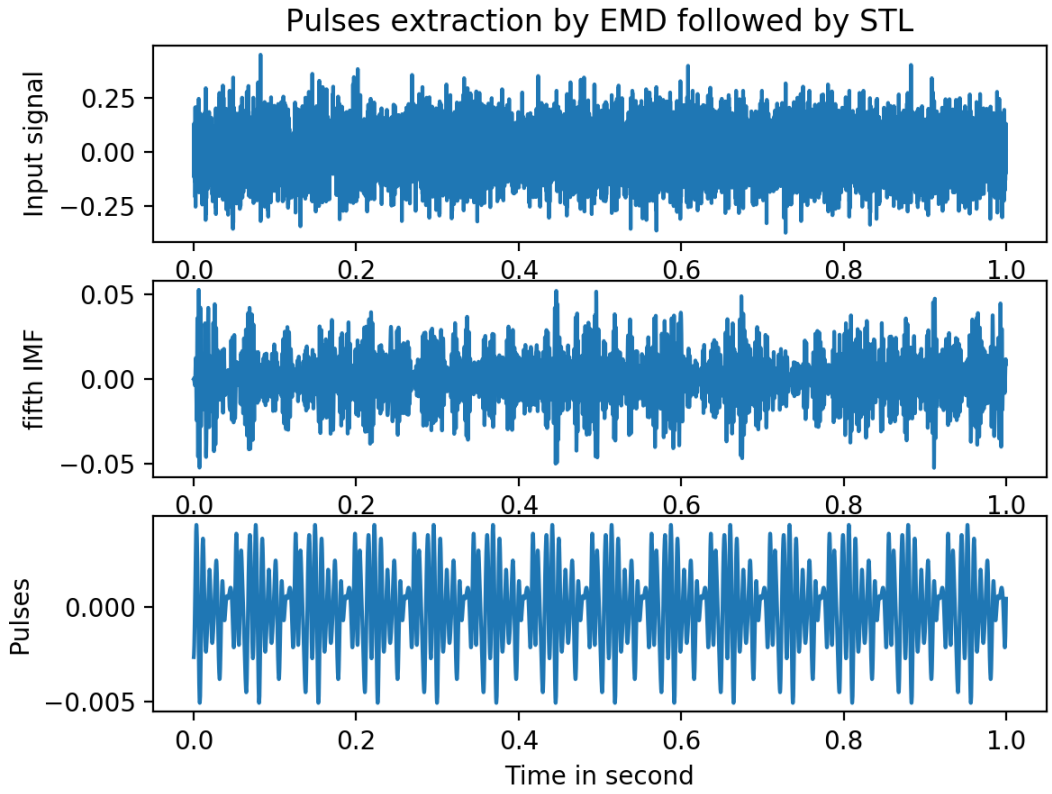

Figure 3.7: A pulse signal extracted by applying EMD followed by STL. The top graph represents the vibration time signal. The middle graph is the fifth intrinsic mode function. The bottom graph is the signal resulting from applying the STL on the IMF. The pulses represent the periodic high frequency signal emitted by bearings defects.

The seasonal component of the intrinsic mode function has a pulse like shape (Figure 3.7, bottom graph). As the bearing rotates, the roller elements (balls) strike the defect area, resulting in a sharp amplitude increase of the vibration signal. This process is repeated periodically through the machine operation. As will be shown shortly, this pulse signal contains all diagnostic information such as bearing failure frequencies, and their characteristics, the rotating speed of the machine, and other unidentified sources.

In the subsequent subsections, the power spectrum density of the seasonal component, from selected intrinsic mode function, is used to detect two types of bearing failure frequencies: Ball pass frequency outer race defect and ball pass inner race frequency defect. Recall that the power spectrum density is approximated here by the periodogram method.

Ball pass outer race frequency defect detection

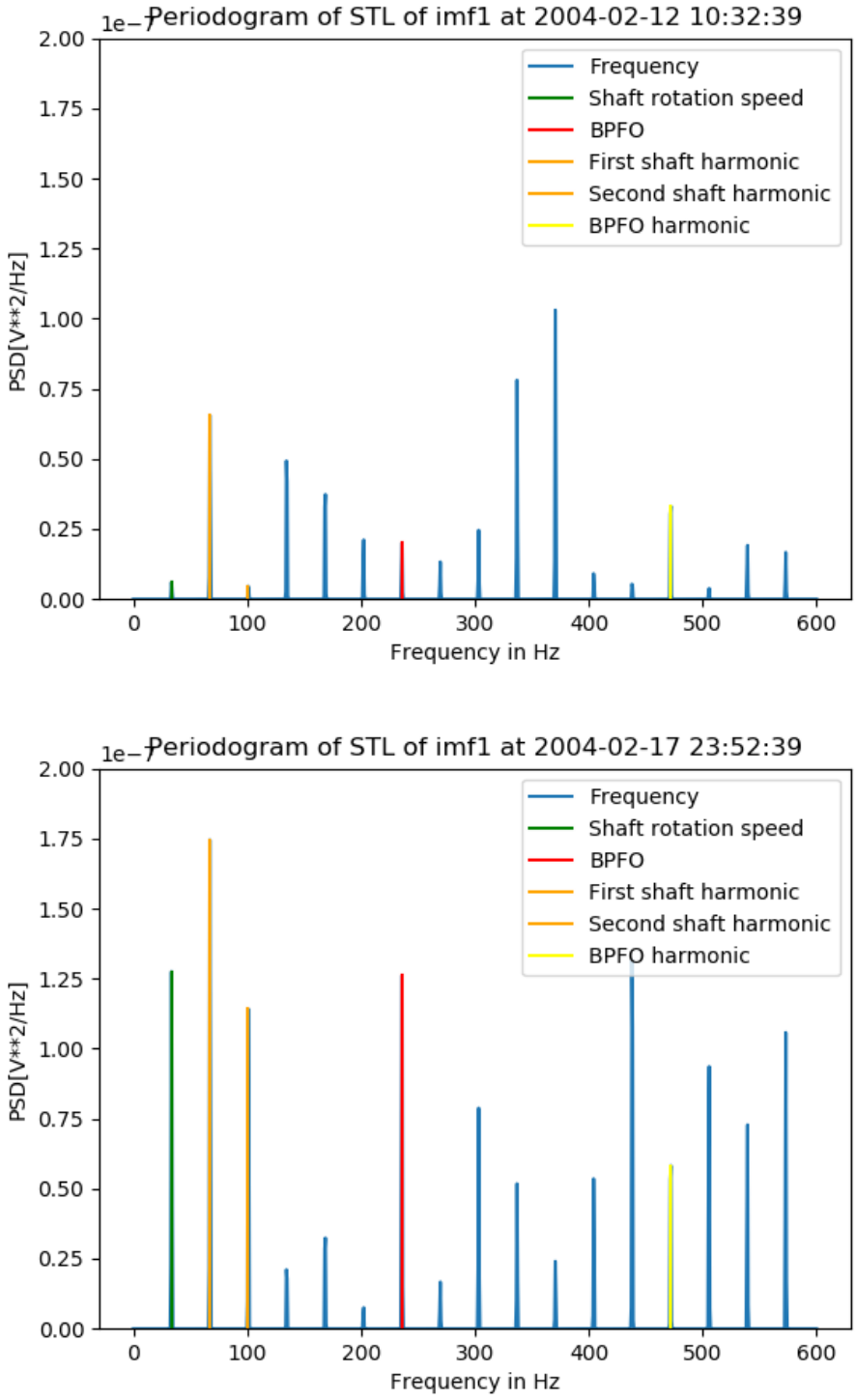

Figure 3.8: Periodogram of the pulse signal obtained from the first IMF, at the beginning (top) and the end (bottom) of experiment number 2

Figure 3.8 shows the periodogram of the seasonal component from the first IMF obtained at the beginning (top) and at the end of experiment number 2. The periodogram shows conspicuous frequency peaks from 0 to 600 Hz approximately. At the beginning of the experiment, the ball pass outer race frequency defect (236.4 Hz) is visible with a relatively low amplitude. At the end of the experiment however, there is a relatively high increase in the outer race frequency defect amplitude, which exhibits the severity of the damage incurred by the bearing.

In addition, the rotating speed (33.3Hz) of the machine shaft is visible in both cases. Harmonics of the defect frequency are also visible. The magnitude of the harmonics have also increased, as the bearing defect is accentuated. Recall that harmonics are characteristics of bearing outer race defect, and are integer multiple of the defect frequency (236.4 Hz) and machine speed (33.3 Hz). In bearing vibration analysis, harmonics are the tail tail sign of a ball pass outer race defect. The rest of the peaks in the periodogram represent unexplained phenomenon. Possibly signals from other parts of the machine.

The difference between the theoretical value of the ball pass outer frequency defect (236.4 Hz) and the one from the periodogram (236.03 Hz) is about 0.16%. While the difference between the theoretical shaft frequency (33.33 Hz) and the on the periodogram (33.63 Hz) is about 0.9 %. This shows that the proposed method is relatively efficient in detecting important diagnostic information from the vibration signal.

As pointed out earlier, the empirical mode decomposition, maps each intrinsic mode function to region of high and low frequencies in the original signal. As such we would expect low frequency processes such as the rotation speed of the machine, in the periodogram of lower frequency intrinsic mode functions. Figure 3.9 shows the periodogram of a lower frequency intrinsic mode function (IMF8). This IMF displays only a peaks corresponding to the rotation speed of the machine (33.3 Hz).

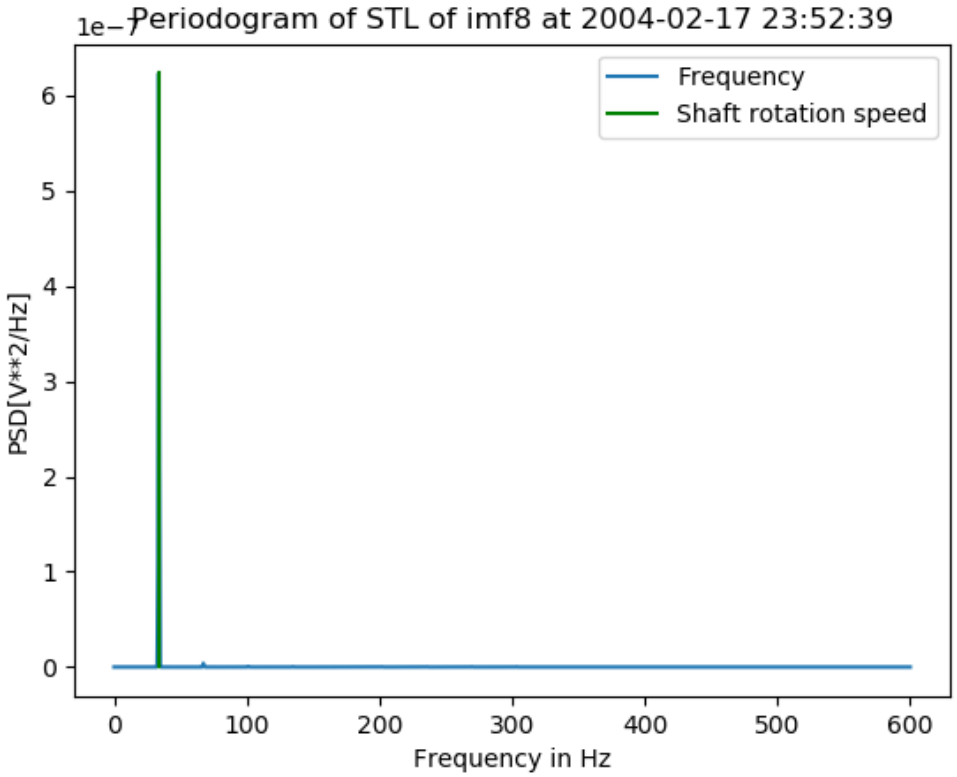

Figure 3.9: Periodogram of a lower frequency intrinsic mode function, displaying a peak corresponding to the rotation speed of the machine housing the bearing

Ball pass inner race frequency defect detection

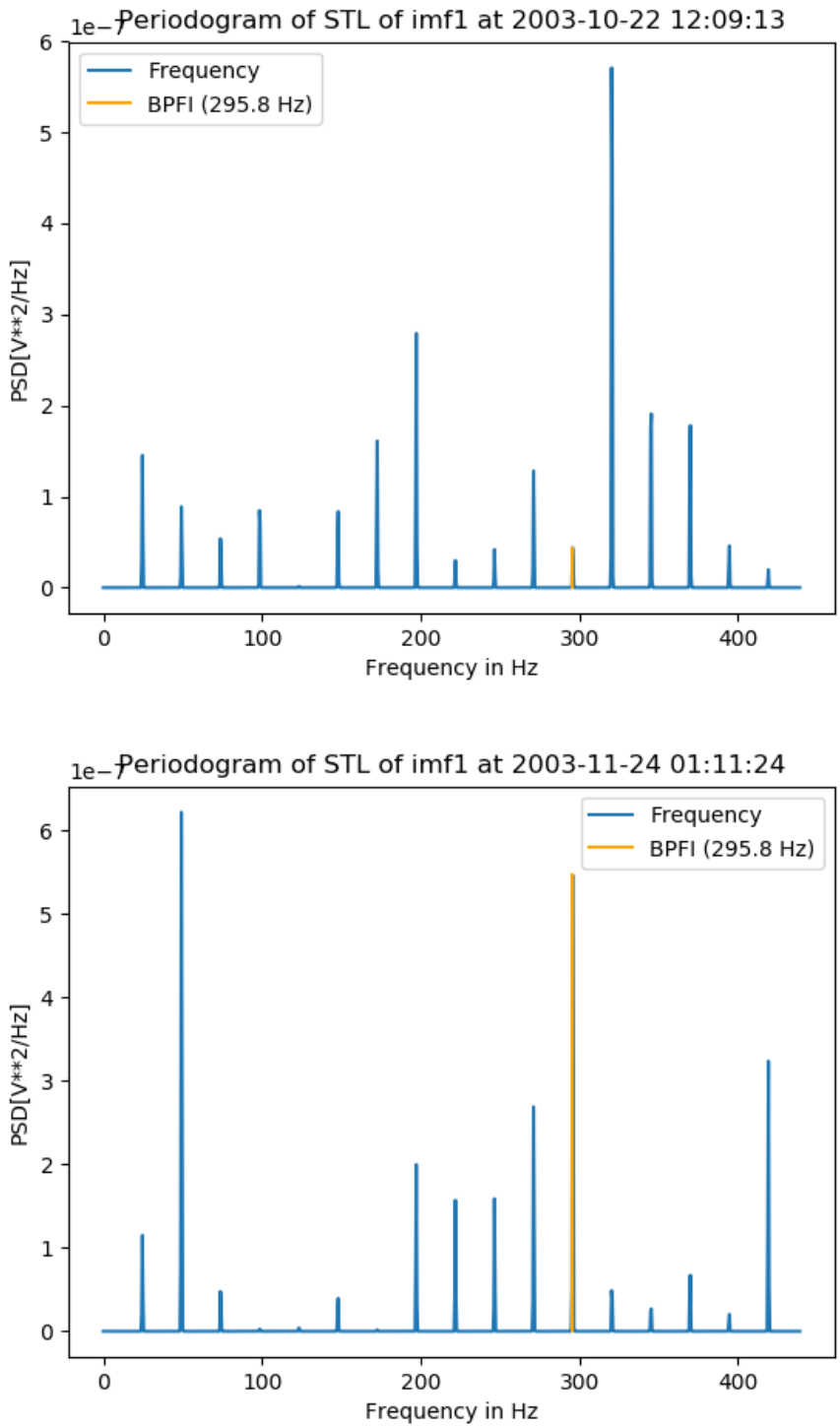

Figure 3.10: Periodogram of the pulse signal obtained from IMF number 1, at the beginning (top) and the end (bottom) of experiment number 1.

Figure 3.10 shows the periodogram obtained from the seasonal component of the first intrinsic mode functions at the start (top) and at the end (bottom) of experiment number 1. The inner race defect frequency (296.8 Hz) is visible in both cases. The amplitude of the inner race defect frequency had increased considerably towards the end of the experiment. This again indicates that the defect incurred by the bearing had increased in severity. The difference between the theoretical value of the BPFI (296.8 Hz) and the detected value (295.8) is about 0.33 %. The periodogram also exhibits other frequency peaks for which we can not account for, due to lack of additional information.

3.3 Summary

In this chapter, a new method for bearing fault detection was presented. This new technique uses the empirical mode decomposition (EMD), the seasonal trend decomposition based on LOESS, and the power spectral density in order to identify bearing failure. The empirical mode decomposition splits an input vibration signal into components called intrinsic mode functions (IMFs). The first few IMFs contain high and low frequency components of the input signal, while the remaining IMFs contain mainly low frequency components.

Once the IMFs are obtained, depending on the frequency range of the bearing failure frequencies, an IMF is chosen. In this thesis, the first IMF is selected for bearing failure detection. Furthermore, the seasonal trend of the selected IMF is extracted, and the power spectral density of the latter is approximated. By extracting the seasonal trend of the selected IMF, the mode mixing phenomenon incurred by the EMD is mitigated. Recall that mode mixing is the mixture of high and low frequency mode in an IMF. When mode mixing occurs, the generated IMF looses its physical meaningful property.

The results from the proposed new method applied to a concrete case study, show that the new scheme is able to detect bearing failure, as well as their characteristics such as harmonics. In addition, lower amplitude intrinsic mode functions identify lower process such the rotation speed of the machine.

Chapter 4

Conclusion

In section 4.1, the results obtained from the high frequency resonance technique and the scheme presented in this thesis are compared. Section 4.2 gives a summary, followed by a discussion regarding the main results, and recommendations to extend this work.

4.1 Comparison of results

Figure 4.1 and 4.2 show the frequency spectrum obtained from the high frequency resonance technique (HFRT), and the method proposed in this thesis, respectively. The magnitude of each peak in the spectrum given by the HFRT is in G or mm/s^2. Whereas, the magnitude of the peaks obtained in this thesis are in energy square per Hertz. That is, the energy contribution per Hertz. Regardless, for bearing fault detection, a high magnitude defect frequency, corresponds to a severe bearing fault.

Both Figures represent the spectrum of a signal obtained from a bearing with an inner ring defect. The latter can produce a noisy frequency spectrum when the HFRT is applied. This can been seen from Figure 4.1, which shows a relatively noisy frequency spectrum. Although the HFRT applies a series of filtering in order to strip out irrelevant signal components, the resulting frequency spectrum still contains peaks embedded in noise. Consequently, identifying bearing faults with certainty, can be challenging.

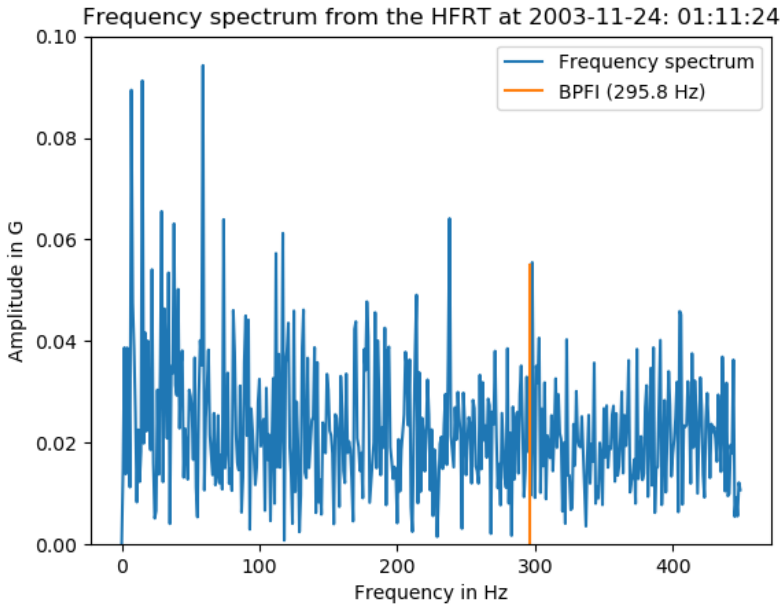

Figure 4.1: Frequency spectrum with an identified ball pass inner race defect frequency obtained from the high frequency resonance technique (HFRT).

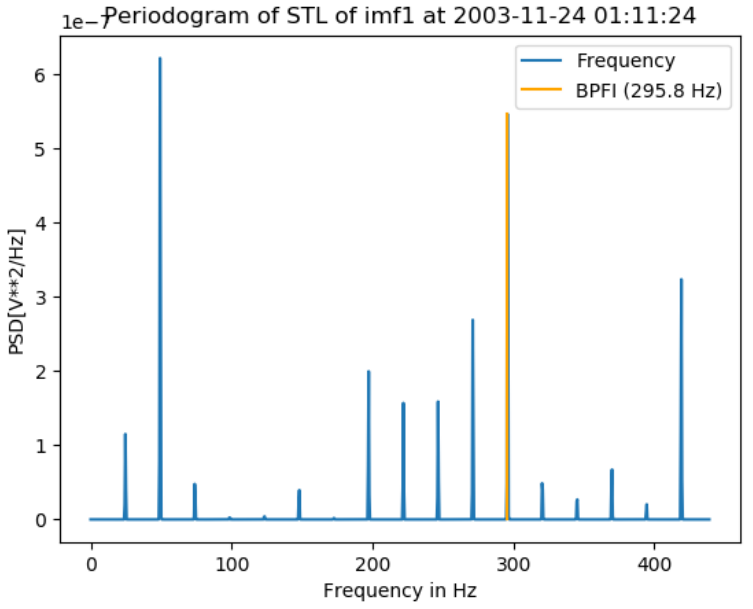

Figure 4.2: Frequency spectrum with an identified ball pass inner race defect frequency obtained from the method posited in this thesis.

In contrast, Figure 4.2 shows a clear spectrum with conspicuous frequency peaks. By applying

the empirical mode decomposition on the target vibration signal, the most relevant information is extracted. Furthermore, the seasonal component of the extracted intrinsic mode function represents a relatively noiseless signal. This means that only part of the signal containing diagnostic information are captured. With enough information of the machine, the origin of most peaks on the frequency spectrum could be (almost) explained. This make bearing fault diagnosis easier and accurate, as we are only restricted to few frequency peaks to analyze.

4.2 Summary

The proclivity of machines towards failure, imposes a monitoring and maintenance scheme, in order to avoid unexpected and catastrophic breakdown. To mitigate such events, most machines are equipped with an array of sensors, collecting continuously data. By applying signal processing methods, incipient failures can be detected.

Being continually subjected to extensive load and stress, bearing failures represent more than 40% of defects in rotating machines. In most industries, Fourier transform is the corner stone of nearly all methods, applied to bearing faults detection. One of the most widely used scheme is the high frequency resonance technique (HFRT). The HFRT filters a bearing vibration signal, in order to remove noise, and isolate desirable components, before generating a frequency spectrum. For a bearing, the failure frequencies are derived from its geometrical components and the rotational speed of the machine on which it is mounted. Once the failure frequencies are available (given by the manufacturer), it suffices to search for them in the frequency spectrum, derived from the Fourier transform.

The high frequency resonance technique although widely used, can produce a very noisy spectrum, in particular when the bearing is subjected to a severe inner race defect. This can render fault detection very challenging. In this thesis therefore, a new method is presented to mitigate this issue. This proposed scheme, consists of decomposing an input bearing vibration signal, into successive high to low frequency components called intrinsic mode functions (IMFs). This is achieved through the empirical mode decomposition (EMD).

Once the intrinsic mode functions (IMFs) have been computed, the first IMF is selected. Furthermore, its seasonal trend is computed through the seasonal trend decomposition method by LOESS (STL). The STL is a computational efficient method that decomposes a signal into its trend and seasonal part. The trend is a monotone function, while the seasonal part is a periodic oscillatory sinusoidal. Moreover, the power spectral density of the seasonal component of the first intrinsic mode function is approximated. This is achieved through the periodogram

method (approximation of the power spectral density). To test the efficiency of the proposed method, a case study was performed. The results show that the new technique was able to identify bearing failures. In addition, the frequency spectrum obtained was relatively noiseless, compared to the one computed through the high frequency resonance technique.

The Hilbert-Huang transform, through the empirical mode decomposition (EMD), boasts itself in generated mono component intrinsic mode functions. However, the EMD suffers from what is known as mode mixing. The latter introduces one or more oscillatory modes into an IMF. This is precisely the reason why in this thesis, the seasonal component of each IMF is extracted through the STL, in order to circumvent mode mixing.

The result obtained are satisfactory in the sens that, bearing failure frequencies are clearly identified. However, only two types of bearing failures were tested: namely faults occurring in the inner and outer ring. The justification is that these are the prevalent faults encountered. As an extension to this work, the new method could be applied to detecting the whole range of bearing defects, in order to further test its full validity. In addition, since the data for this case study was obtained under a control experiment, the new scheme could be tested with data from real industrial processes.

Bibliography

Albrecht, P., Appiarius, J., McCoy, R., Owen, E., and Sharma, D. (1986). Assessment of the reliability of motors in utility applications-updated. *IEEE Transactions on Energy Conversion*, EC-1:39–46.

Antoni, J. (2006). The spectral kurtosis: A useful tool for characterising non-stationary signals. *Mech. Syst. Signal Process*, 20:282–3007.

Antoni, J. (2007). Fast computation of the kurtogram for the detection of transient faults. *Mech. Syst. Signal Process*, 21:108–124.

Antoni, J., Bonnardot, F., Raad, A., and Badaoui, M. E. (2004). Cyclostationary modelling of rotating machine vibration signals. *Mech. Syst. Signal Process*, 18:1285–1314.

Antoni, J. and Randall, R. (2006). The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines. *Mech. Syst. Signal Process*, 20:308–331.

Balderston, H. (1969). The detection of incipient failure in bearings. *Material Evaluation*, pages 121–128.

Board, D. (1975). Incipient failure detection in high-speed rotating machinery. *Proceedings of the 1oth Symposium on Non-Destructive Evaluation, San Antonio, Texas23-25 April*, pages 8–18.

Bogges, N. (2009). *A First Course in Wavelets with Fourier Analysis*. WILEY.

Borghesani, P., Pennacchi, P., Ricci, R., and Chatteron, S. (2013). Testing second order cyclostationary in the square envelope spectrum of non-white vibration signals. *Mech. Syst. Signal Process*, 40:38–55.

Broderick, J., Burchill, R., and Clarck, H. (1972). Design and frabrication of prototype system for early warning impending bearing failure. *NASA CR 123717*.

Burchill, R. (1973). Resonance structure techniques for bearing fault analysis. *Proceedings of the 18th Meeting of Mechanical Failures Prevention Group, National Bureau of Standards NBSIR 73-255*, pages 21–29.

Burchill, R., Frarey, J., and Wilson, D. (1973). New machinery health diagnostic techniques using high-frequency vibration. *Society of Automated Engineers Paper 730930*.

Butterworth, S. (1930). On the theory of filter amplifiers. *Experimental wireless*, pages 536–541.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I. (1990). Stl: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6:3–73.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74:829–836.

Cleveland, W. S. and Develin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83:596–610.

Courrech, J. and Gaudet, M. Envelope analysis-the key to rolling-element bearing diagnosis.

Darlow, M. and Badgley, R. (1975a). Application for early detection of rolling element bearing failure using the high-frequency resonance technique. *American Society of Automated Engineers Paper 75-DET-46*.

Darlow, M. and Badgley, R. (1975b). Early detection of defects in rolling-element bearings. *Society of Automated Engineers Paper 750209*.

Darlow, M., Badgley, R., and Hogg, G. (1974). Application of high frequency resonance techniques for bearing diagnostics in helicopter gearboxes. *Report, US Army Air Mobility Research and Development Laboratory*, pages 74–77.

Fan, W., Wang, X., Wang, L., and Yu, Z. (2016). The application of hilbert-huang transform energy spectrum in brushless direct current motor vibration signals monitoring of unmanned aerial vehicle. *Advances in Mechanical Engineering*, 8(9).

Fangtao, W., Yangyang, Z., Bin, Z., and Wensheng, S. (2011). Application of wavelet packet sample entropy in the forecast of rolling element bearing fault trend. *Multimedia and Signal Processing (CMSP), 2011 international Conference*, pages 12–16.

Feldman, M. (2010). Hilbert transform in vibration analysis. *Mechanical System and Signal Processing*.

Fosso, O. B. and Molinas, M. (2019). Mode mixing separation in empirical mode decomposition of signals with spectral proximity.

Gabor, A. K. (1982). Frequency, time and memory. *Applied optics*, 20.

Gabor, D. (1944). Theory of communication.

Girondin, V., Pekpe, K. M., and Morel, H. (2013). Bearing fault detection in helicopters using frequency readjustement and cyclostationary analysis. *Mech. Syst. Signal Process*, 38:38–55.

Gupta, P. and pradhan, M. K. (2016). A normalized hibert-huang transform technique for bearing fault detection. *5th International Conference of Materials Processing and Characterization*.

Hai, Q., Jay, L., Jing, L., and Gang, Y. (2006). Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of Sound and Vibration*, 289:1066–1090.

He, Q., Wang, J., Liu, Y., Dai, D., and Kong, F. (2012). Multiscale noise tuning of stochastic resonance for enhanced fault diagnosis in rotating machines. *Mech. Syst. Signal Process*, 28:443–457.

Huang, N. E., c. Chern, C., Huang, K., Salvino, L. W., Long, S. R., and Fan, K. L. (2001). A new spectral representation of earthquake data: Hilbert spectral analysis of station tcu129, chi-chi, taiwan, 21 september 1999. *Bulleting of the Seismological Society of America*, 91:1310–1338.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Zheng, H. H. S. Q., Yen, N.-C., Tung, C. C., and Liu, H. H. (1998). The empirical mode decomposition and the hilbert huang spectrum for nonlinear and non-stationary time series analysis. *The royal society*, 454:909–995.

Huang, N. E. and Wu, Z. (2008). A review on hilbert-huang transform: Method and its applications to geophysical study. *Reviews of Geophysics*, 46:1208–1217.

Jianbo, Y. (2012a). Health condition monitoring of machines based on hidden model and contribution analysis. *Instrumentation and Measurement, IEEE Transactions*, 61:2200–2211.

Jianbo, Y. (2012b). Local and nonlocal preserving projection for bearing defect classification and performance assessment. *Industrial Electronics, IEEE Transactions*, 59:2363–2376.

Jiang, R., Cheng, J., Dong, G., Liu, T., and Xiao, W. (2013). The weak fault diagnosis and condition monitoring of rolling element bearing using minimum entropy deconvolution and envelope spctrum. *Proc. Inst. Mech. En. Part C J. Mech. Eng*, 227:1116–1129.

Khadersab, A. and Shivakumar, S. (2018). Vibration analysis tecniques for rotating machinery and its effect on bearing faults. *2nd international Conference on Materials Manufacturing and Design Engineering*, 20:247–252.

Konar, P. and Chattopadhyay, P. (2011). Bearing fault detection on induction motor using wavelet and support vector machines (svm). *Applied soft Computing*.

Lee, J., Qiu, H., Yu, G., and Lin, J. (2007). Ims, university of cincinnati bearing. bearing data set, nasa ames prognostics data repository. (http://ti.arc.nasa.gov/project/prognostic-data-repository), nasa ames research center, moffett field, ca.

Lei, Y., He, Z., and Zi, Y. (2011). Eemd method and wnn for fault diagnosis of locomotive roller bearings. *Expert Syst. Appl*, 38:7334–7341.

Li, H., Zhang, Y., and Zheng, H. (2009a). Hilbert-huang transform and marginal spectrum for detection and diagnosis of localized defects in roller bearings. *Journal of Mechanical Science and Technology*, 23:291–3001.

Li, H., Zhang, Y., and Zheng, H. (2009b). Hilbert-huang transform and the marginal spectrum for detection and diagnosis of localized decefts in roller bearings. *Journal of Mechanical Science and Technology*, pages 291–301.

Lin, J. and Qu, L. (2000). Feature extraction based on morlet wavelet and its application for mechanical diagnosis. *J. Sound Vib*, 234:135–148.

McFadden, P. and Smith, J. (1984a). Model for the vibration produced bby a single point in a rolling element bearing. *Journal of Sound and Vibration*, pages 69–82.

McFadden, P. and Smith, J. (1984b). Vibration monitoring of rolling element bearings by the high frequency resonance technique-a review. *Tribology International*, 17:3–10.

Mejia, D. T., Medjaher, K., Zerhouni, N., and Gerard, T. (2011). Hidden markov models for failure diagnostic and prognostic. *Prognostics and System Health Management Conference (PHM-Shenzhen)*, pages 1–8.

Mejia, T., Alejandro, D., Kamal, M., Noureddine, Z., and Gerard, T. (2010). A mixture of gaussians hidden markov model for failure diagnostics. *Automation Science and Engineering (CASE), 2010 IEEE Conference*, pages 338–343.

Mejia, T., Alejandro, D., Kamal, M., Noureddine, Z., and Gerard, T. (2012). A data driven failure prognostics method on mixture of gaussians hidden markov models. *Reliability, IEEE Transaction*, 61:491–503.

Mobius (2014). Vibration training course book category 2. mobius institute.

Mortada, M. A. and Yacout, S. (2011). cbmlad-using logical analysis of data in condition based maintenance. *Computer research and development (ICCRD), 2011 3rd International Conference*, pages 30–34.

OF, E., Camci, F., and Jennions, I. (2012). Major challenges in prognostics: Study on benchmarking prognostics datasets. *European Conference of Prognostics and Health Management Society*, 59:2363–2376.

Osman, S. and Wang, W. (2013a). An enhanced hilbert-huang transform technique for bearing condition monitoring. *Measurement Science and Technology*, 22.

Osman, S. and Wang, W. (2013b). A new hht technique for bearing health condition monitoring. *Proceedings of the International Conference on Mechanical Engineering and Mechatronics, Toronto, Ontario, Canada*.

Osman, S. and Wang, W. (2014). A normalized hibert-huang transform technique for bearing fault detection. *Journal of Vibration and Control*, 22.

Peng, Z., Tse, P. W., and Chu, F. (2004). A comparison study of improved hilbert-huang transform and wavelet transform: Application to fault diagnosis for rolling bearing. *Mechanical and Signal Processing*, 19:974–988.

Qiu, H., Lee, J., Lin, J., and Yu, G. (2006). Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *J. Sound Vib*, 289:1066–1090.

Rai, A. and Upadhyay, S. (2016). A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. *Tribology international*.

Rai, V. and Mohanty, A. (2006). Bearing fault diagnosis using fft of intrinsic mode functions in hilbert-huang transform. *Mechanical Process ans Signal Processing*, 21:2607–2615.

Randal, R. B. and Antoni, J. (2010). Rolliing element bearing diagnostics-a tutorial. *Mechanical System and Signal Processing*.

Randall, R. B. and Antoni, J. (2011). Rolling element bearing diagnostics-a tutorial. *Mechanical System and Signal Processing*, 25:485–520.

Rego, D. M., Oscar, F. R., and Amparo, A. B. (2011). Power wind mill fault detection via one-class nu-svm vibration signal analysis. *Neural Networks (IJCNN), The 2011 International Joint Conference*, pages 511–518.

Sallo, Z. G. and Grif, H. S. (2019). Hilbert-huang transform and marginal spectrum for detection and diagnosis of localized defects in roller bearings. *Procedia Manufacturing*, 32:591–595.

Sawalhi, N., Randall, R. B., and Endo, H. (2007). The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. *Mech. Syst. Signal Process*, 21:2616–2633.

scipy community, T. (2019). *https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.hilbert.html*.

Sergey, P. and Zigmund, B. (2012). Remaining useful life estimation for system with non-trendability behaviour. *Prognostics and Health management (PHM), 2012 IEEE Conference*, pages 1–6.

Soualhi, A., Medjaher, K., and Zerhouni, N. (2015). Bearing health monitoring based on hilbert-huang transform, support vector machine, and regression. *IEEE TRANSATIONS ON INTRUMENTATION AND MEASUREMENT*, 64.

Stoica, P. and Moses, R. (2004). Spectral analysis of signals. *Experimental wireless*.

Tan, J., Cheng, X., Wang, J., Chen, H., Cao, H., Zi, Y., and He, Z. (2009). Study of frequency-shifted and re-scaling stochastic resonance and it application to fault diagnosis. *Mech. Syst. Signal Process*, 23:811–822.

Vasudevan, K. and Cook, F. (2000). Empirical mode skeletonization of deep crustal seismic data: Theory and applications. *Journal of Geophysical Research Atmospheric*, 105:7845–7856.

Xiaoan, Y. and Minping, J. (2018). A novel optimized svm classification algorithm with multi-domain feature and its application to fault diagnosis of rolling bearing. *Neurocomputing*, 313.

Yacout, S. (2012). Logical analysis of maintenance and performance data of physical assets, id34. *Reliability and Maintainability Symposium (RAMS), Proceedings-Annual*, pages 1–6.

Yan, R. and Gao, R. X. (2006). Hilbert-huang transform and marginal spectrum for detection and diagnosis of localized defects in roller bearings. *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT*, 55:2320–2329.

Yu, D., Cheng, J., and Yang, Y. (2005). Application of emd method and the hilbert spectrum to the fault diagnosis of roller bearings. *Mech. Sys. Signal Process*, 19:259–270.

Zhang, B., Zhang, S., and Li, W. (2019). Bearing performance degradation assessment using long short-term memory recurrent network. *Computer in Industry*, 106.

Zhao, M., Lin, J., Xu, X., and Li, X. (2014). Multi-fault detection of rolling element bearings under harsh working condition using imf-based adaptive envelope order analysis. *Sensor*, 14:20320–20346.