Accounting for nonstationarity in the condition monitoring of wind turbine gearboxes

A Thesis submitted to the University of Sheffield
for the degree of Doctor of Philosophy in the Faculty of Engineering

by

I. Antoniadou

Department of Mechanical Engineering

The University of Sheffield

September 2013

Acknowledgments

First of all, I would like to express my gratitude to my supervisor, Prof. Keith Worden for allowing me to work with him. I would like to thank him for his constant support and guidance, and for being a friend to me apart from a supervisor. I would also like to thank Dr. Graeme Manson, who was my co-supervisor, and Prof. Wieslaw Staszewski, my supervisor in the first year of my Ph.D., for their help in my research. More particularly, I would like to thank Prof. Wieslaw Staszewski for providing me the initial idea of my thesis, and for his support even after leaving to AGH University.

I would really like to thank Dr. Elizabeth Cross, for making the use of *cointegration analysis* a much easier experience than what it would have been without her kind help.

I would also like to express my gratitude to Dr. Charles R. Farrar, for accepting me in the Los Alamos National Laboratories (LANL), and being my mentor during my secondment there, as well as the whole team in the Engineering Division of LANL, for being so friendly and helpful to me.

I would also like to thank Dr. Tomasz Barszcz and EC Grupa for providing me the wind turbine gearbox datasets.

I should also thank Dr. Matthew B. Marshall, Thomas P. Howard and Jack Naumann from the Leonardo Centre for Tribology for allowing me to participate in their experiments with the RICARDO bearing rig.

Furthermore, I would like to thank my colleagues-and occasional office mates-in the Dynamics Research Group: Nikos, Elizabeth, Pete, Charles, Tara, Rob, Vaggelis, Sofia, Claudio, Nevena, Will and James. A special thank you to Nikos for making my Ph.D. years funnier.

Finally, I would like to thank my parents for all their sacrifices and their

encouragement all these years in order that I pursue my goals. I would also like to thank my sister, Eleni and Nikos for their patience and for being there for me.

ABSTRACT

Increasing growth of wind turbine systems suggests a more systematic research around their design, operation and maintenance is needed. These systems operate under challenging enviromental conditions and failure of some of their parts, for the time being, is frequent, although undesirable. Wind turbine gearboxes, more particularly, seem to be so problematic that some wind turbine designs avoid including them. Structural health monitoring and condition monitoring of wind turbines appear to be necessary in order to determine the condition and lifespan of the wind turbine components and the drivetrain respectively. In this way reparative actions could be taken whenever needed resulting in reduction of maintenance costs.

This thesis focuses on the condition monitoring of wind turbine gearboxes, taking into account the varying loads that they endure. Currently, the vibration-based damage detection methods used in real life wind turbine condition monitoring systems are based on conventional methods that generally fail to detect damage at its early stage under the operational conditions observed in wind turbines. Load and speed variations of the drivetrain that are observed commonly in wind turbines influence the vibration signals and can possibly affect potential damage features. This shows a demand for effective methods for early damage detection. Developments in the area of advanced signal processing should be examined and applied in damage detection of wind turbine gearboxes. Methods from time-frequency analysis, time-scale analysis, pattern recognition, multivariate statistics and econometrics are examined in this study in a condition monitoring context.

One important part of the work presented is the development of a simple gearbox model interfaced with realistic wind loading, a model feature that appears to be novel. Other interesting aspects of this thesis are related to the use of the empirical mode decomposition method for time-frequency analy-

sis. The use of Teager-Kaiser energy operator as an alternative technique to Hilbert transform for the estimation of the instantaneous characteristics of the decomposed signals is one of these aspects. The study showed that for some cases and under certain conditions this operator could help to improve the time-frequency analysis. Another aspect is the observation of the change of the number of the intrinsic mode functions produced, for the different load and damage cases, during the decomposition process. This observation was connected theoretically with what is known as the mode mixing problem of the empirical mode decomposition method. For the feature discrimination part of this work, the simplest novelty detection method, outlier analysis, was used in a slightly different manner than in previous studies and the results obtained were compared with a novel adaptive thresholding technique, the 3D phase-space thresholding method. The previously described approaches were applied on the simulated gearbox data but also on real wind turbine gearbox data. Finally, cointegration analysis was proposed as a potential method for removing the effects of the gearbox load variations. This is a novel concept for the condition monitoring of wind turbine gearboxes. An approach which makes it possible to use data from just a single sensor in order to perform cointegration analysis was developed and the process for applying multiscale cointegration using either wavelets or the empirical mode decomposition method was discussed. This final part of the work is an initial step towards applying cointegration to condition monitoring data.

Table of Contents

Chapter 1

Motivation and scope of the thesis

Wind power has gained significant interest in recent years. The main reasons contributing to this are the rise of environmental conciousness, the limited supply of fossil fuels (expected to run out in the next 50 years if the rates of their consumption stay the same) and finally the fast development of technology which helped in decreasing the costs of renewables, making them a realistic commercial option. Figure 1.1 shows part of a wind farm in Aalborg, Denmark.

Figure 1.1: Wind farm in Aalborg, Denmark (2011).

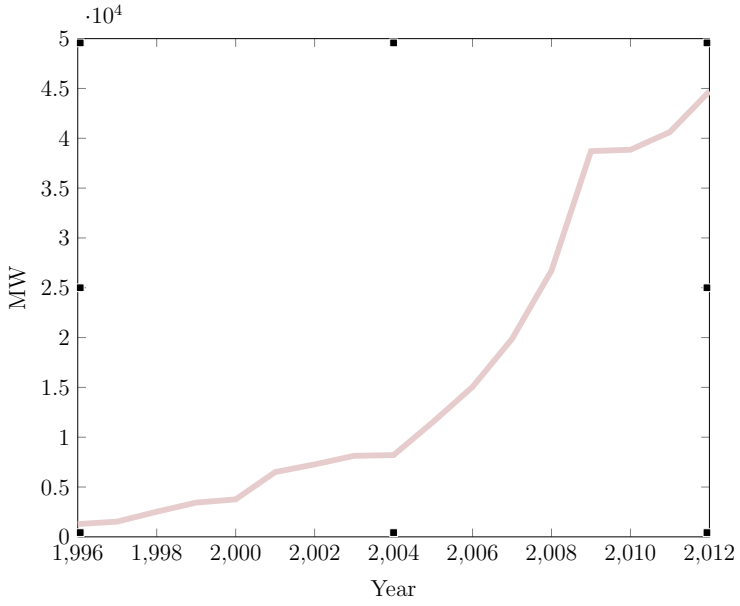


Figure 1.2: Global installed wind energy capacity over the years 1996-2012 in MW [1].

1.1 Motivation

According to the global wind statistics of 2012, estimated by the *global wind energy council* (GWEC) [1], wind power expanded by almost 20% in 2012 around the world to reach a new peak of 282 GW of total installed capacity (Figure 1.2).

The UK now ranks sixth in the world for installed wind power, with 8.5 GW. In Europe, only Germany (31 GW) and Spain (23 GW) have more. China leads the world with 77 GW installed and the US is second with 60 GW. The UK is by far the world leader in offshore wind deployment, installing 0.85 MW in 2012 to bring the total so far to 3 GW as shown in Table 1.1. Wind energy makes a significant contribution to the UK, with an output of 15.5 TWh in 2011 equivalent to the annual electricity demand of 4.7 million homes [2].

In order to make wind energy an attractive alternative option of power generation, the matters of reliability and cost should be addressed [3]. In order to reduce the costs of wind energy, larger wind turbines were manufactured, which on the other hand meant more complex designs. Bigger and more complex designs though, operating under extreme environmental conditions,

Table 1.1: Global offshore installed capacity for 2011 and 2012 (GWEC) [1]

Offshore [MW]	**Total 2011**	**New 2012**	**Total 2012**
United Kingdom	2093.6	854.2	2947.9
Denmark	847.3	46.8	921.1
Belgium	195.0	184.5	379.5
Germany	200.3	80.0	280.3
Netherlands	246.8	0.0	246.8
Sweden	163.7	0.0	163.7
Finland	26.3	0.0	26.3
Ireland	25.2	0.0	25.2
Norway	2.3	0.0	2.3
Portugal	2.0	0.0	2.0
PR China	262.6	127.0	389.6
Japan	25.2	0.1	25.3
Total	4117.3	1292.6	5410.0

common especially for offshore wind turbines, meant bigger challenges in terms of reliability. The oldest wind turbine designs seem to have disappointed expectations concerning their life expectancy [4], and this has been mainly because the initial designs underestimated the extreme conditions under which the wind turbine components operate. These weather conditions include large temperature variations, extreme wind turbulence and lightning. The wind turbine components that have been observed to fail more frequently are the electrical components, the blades and the gearbox. In addition, the gearbox is the most expensive component to repair and the component that can cause the highest downtime in a wind turbine system. All the above mean high operation and maintenance costs, which is not a desirable characteristic in order to make wind power a competitive energy source in the energy market. Better designs of the wind turbine components is of course one answer to the solution of this problem; the other is *structural health monitoring* (SHM) and *condition monitoring* (CM) of the wind turbine systems. For CM, changing environmental and operating conditions of the wind turbine can create difficulties in the signal processing of the vibration signals. This is because wind variations can lead to load variations on the gearbox. CM in this case is more challenging as will be explained further on.

The work produced in this thesis is part of the SYSWIND project, and is funded by the European Commission Seventh Framework Program under the Marie Curie Network Scheme. The aim of this project is to conduct research that could help improve the next generation of wind turbines. Several academic as well as industrial partners and research laboratories are involved in this project, each one of which is responsible for different parts of the work of the project. The full network partners include: Trinity College Dublin, Aalborg University, Risø National Laboratory of Denmark Technical University, LAC Engineering, the University of Patras, the University of Cagliari and the University of Sheffield. The University of Sheffield is responsible for the research related to the SHM and CM of wind turbines.

In the next sections of this chapter the scope of this thesis will be presented and a brief outline of the chapters included will be given.

1.2 Scope of this thesis

The scope of this thesis is to test non-classical vibration-based methods as potential diagnostic tools for condition monitoring of wind turbine gearboxes. The major issue, as discussed in detail later on in this thesis, is the fact that

wind turbine gearboxes do not operate under steady load and speed conditions, which means that their vibration signals are not necessarily stationary all the time. For this reason, the application of more sophisticated signal processing techniques than standard Fourier analysis is necessary.

Time-frequency analysis methods can overcome the problem of analysing signals of a time-varying nature. For this purpose the relatively new *empirical mode decomposition* is proposed and its application on wind turbine gearbox data is demonstrated combined with different amplitude-frequency separation methods. Among these, the *Teager Kaiser energy operator* approach appears to have been applied very few times in the condition monitoring field. The discrimination of damage features is also dealt within a time-frequency perspective using novelty detection methods for this part of the study. The use of an adaptive thresholding method, the *3D-phase space thresholding*, is demonstrated in this part, seeming to be a novel technique in the condition monitoring field and an appropriate choice for the kind of damage features extracted from the time-frequency analysis performed previously.

Finally, a completely different approach is presented as a possible solution to the problem of influence of the varying operating conditions of wind turbines on the damage features. Removing trends induced by operational conditions from damage sensitive data is accomplished by using *cointegration analysis*, a technique that originates in the field of econometrics and has been applied recently in the field of structural health monitoring. Two novel ideas can be found in this part of the study. The solution to the problem of lack of data from different sensors when performing cointegration and the idea of a multiresolution cointegration approach that could be applicable to condition monitoring of wind turbine gearboxes.

1.2.1 Brief outline of thesis

The outline of the chapters of this thesis follows:

- Chapter 2 is an introduction to the basic ideas of condition monitoring. In addition, a literature overview is provided summarising most of the signal processing and machine learning methods that have been used over the years in the condition monitoring field.
- Chapter 3 describes the common configuration of wind turbine systems and their operating modes in order to have a better understanding of the challenging conditions prevailing in wind turbines that might inhibit the successful application of a condition monitoring approach.

Then, one of the main difficulties that one faces in the signal processing part of condition monitoring, related to load variations is presented so as to demonstrate the importance of the use of the more sophisticated analysis methods chosen in this study.

- Chapter 4 gives a detailed description of the main wind turbine gearbox and bearing experimental datasets used in the time-frequency analysis and novelty detection parts of this study. In addition, the simple simulated gearbox model that is used, in order to demonstrate in a better way the kind of results expected from the methods, is thoroughly described. This model allows for certain nonlinear and time-varying characteristics and takes into account varying loads similar to those found in wind turbines.
- Chapter 5 explains the empirical mode decomposition method, the main time-frequency analysis method used in this thesis. The mode mixing problem of this technique is discussed and through a simulation a potential damage feature related to this problem is suggested with concerns and under specific cases. In addition, the Hilbert Transform and the Teager Kaiser energy operator are described; the first as a standard amplitude-frequency separation technique and the second as a rather novel method in this field.
- Chapter 6 applies the empirical mode decomposition method in combination with both of the amplitude-frequency separation techniques presented and a comparison of the methods is made. Damage features in this way are extracted and are available for the feature discrimination that follows in the next chapter.
- Chapter 7 presents in a different manner the way a novelty detection method, outlier analysis, could be applied to the damage features in order to maintain a time-frequency perspective. In addition, 3D-phase space thresholding is introduced as an alternative approach.
- Chapter 8 begins investigations on cointegration. The possibility of cointegration using data from a single sensor is investigated through a simulated example. In addition, multiresolution cointegration is demonstrated on data coming from a benchmark study involving Lamb-wave propagation for detecting damage in a composite plate.
- Chapter 9 concludes the thesis, discusses limitations of the work presented and suggests future work.

Chapter 2

An introduction to condition monitoring

CM is the practice of monitoring parameters indicative of the mechanical condition of rotating machines while in operation. It can be seen as a subfield of the SHM field, that specifically addresses damage detection in rotating and reciprotating machinery, usually as used in the manufacturing and power generation industries [5]. Indeed, condition monitoring methods do not differ significantly from other damage detection methods. A difference of CM when compared to SHM is the fact that the nature of rotating machinery operation influences the way that particular faults are shown in the dynamic responses of the machinery. This is an *a priori* knowledge that can be used in order to achieve a successful CM scheme.

When compared to other maintenance strategies, such as run-to-break down maintenance or preventive maintenance, which is done at regular intervals shorter than the expected "time between failures", on-line CM gives considerable economic advantages, a fact that explains the growing interest in this research area [6].

The most obvious aim of a *CM-based maintenance strategy* is to be able to immediately detect any introduced damage in rotating machinery components, such as gears or bearings. Other objectives for CM would be to determine the fault location and its severity level, after its detection and finally predict the remaining useful life of the component examined, just like the aims of any SHM system. From this point of view the SHM or CM problem is a damage identification problem and can be summarised in an hierarchical

process. Rytter's damage identification hierarchy [7], is usually referred to in describing the SHM objectives (Figure 2.1; although additional levels to this hierarchy have been suggested in the literature [5]). Success at any given level of this hierarchy usually depends upon having already achieved success at lower levels.

Since CM is analogous to SHM, some basic SHM concepts will be first described in the following section, that the author will try to adjust into CM terms.

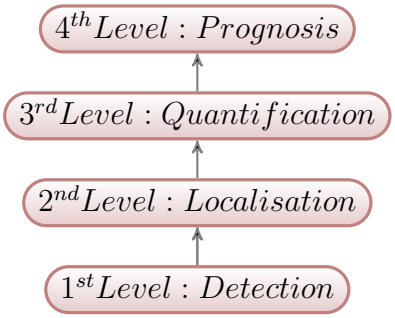


Figure 2.1: Rytter's Hierarchy describing the global objectives of SHM [7].

2.1 The CM pattern recognition paradigm

In general terms, damage is defined as changes introduced into a system, that adversely affect the current performance of that system [8]. These changes can be intentional or unintentional, changes to the material and/or geometric properties as well as changes to the boundary conditions and the system connectivity. The distinction between damage and failure is that failure occurs when the damage progresses to a point where the system can no longer perform its intended function. Damage can be therefore present in all engineering systems at different levels, and can accumulate incrementally over long periods of time, but can also progress very quickly. The above concept of damage implies that damage is not meaningful without a comparison between two different states of the system, one of which represents the initial and often undamaged state.

Pattern recognition is a discipline combining mainly statistical, neural and syntactic approaches that can provide a fundamental framework for carrying

out SHM. A general *statistical pattern recognition paradigm* for an SHM system is summarised in the following sequence of procedures [9]:

- Operational evaluation,
- Data acquisition,
- Feature extraction,
- Statistical modelling for feature discrimination.

These four terms are described briefly in CM terms below:

The *operational evaluation for rotating machinery* answers questions related to the justification of the need of performing CM to the system, the kind of damage in the system being investigated and the description of the operational and environmental conditions of the monitored system. Depending on the application, economic arguments for performing condition monitoring will vary. The definition of damage is usually a straightforward process: the most common types of faults that can be found in rotating machinery are shaft faults, gear faults, bearing faults and faults in the electrical components for the case of electrical machines. Shaft faults include unbalance, misalignment, shaft bending, whirl and cracking, gear faults include spalls, pitting, and cracks and bearing faults are categorised in outer/inner race faults, rolling element faults and cage faults. Finally the main operational limitations of machines are the noisy environments, personnel availability for the necessary measurements, and sometimes the remote location or hazardous environments of such systems.

The *data acquisition* part of the CM process has to do with the selection of the excitation methods, the sensor types, number and locations and the data acquisition hardware. Selection of the placement of the appropriate sensors depends on the type of the machinery and its construction. This thesis focuses on vibration-based condition monitoring, that is why what is described in the following will be explicitly under this perspective. Vibration transducers can measure displacement, velocity or acceleration, that is all three variables in which lateral vibration manifests [6]. Proximity probes, that can give a measure of the relative distance between the probe tip and another surface, accelerometers, velocity transducers, dual vibration probes, and laser vibrometers are such kinds of sensors used to measure the lateral vibration of machines. Torsional vibration transducers also exist, these are useful in cases where failures in machines occur because of excessive torsional vibration. Torsional laser vibrometers and shaft encoders can measure torsional vibration.

Feature extraction is an important part of the process that deals with the identification of data features that allow the distinction between damaged and undamaged states of machinery. A damage-sensitive feature is a quantity that is extracted from the measured and analysed data and is indicative of the presence or not of damage. In the case of rotating machinery, the types of features are often related to the machine element examined, the specific fault and in some cases the level of damage. This is an advantage that can facilitate the condition monitoring process, since the kind of feature observed may give information about the defective element and its location and the level of damage. The basic feature extraction methods used in condition monitoring are described in Section 2.3 of this chapter.

Finally, the *statistical modelling for damage detection* is concerned with the implementation of algorithms that analyse the extracted features in order to quantify the damage state of the structure. The machine learning algorithms used in this step fall into two categories: the supervised learning algorithms, such as group classification and regression analysis, that are implemented when training data are available for all conditions of the machine examined, and the unsupervised learning algorithms, such as outlier analysis or novelty detection methods, used when only data from the undamaged condition of the machine are available. All of these algorithms use the statistical discipline. Such kinds of methods are discussed in Section 2.4 of this chapter.

2.2 Vibration-based condition monitoring

Vibration analysis is one of the most prevalent methods used for condition monitoring [6]. The concept behind vibration analysis is that the vibration signature of the component examined changes in the case of damage occurrence. Generally, when operating in good conditions, machines generate vibrations that are linked to their shaft rotation, the gear meshing of the gear stages of their gearbox, and other periodic or non-periodic events during their operation.

The advantages that vibration analysis offers when compared to other condition monitoring methods are significant. For example, when compared to oil analysis, vibration analysis can give immediate information about the machine's condition, where in contrast with oil analysis, several days elapse between the sample collection and their analysis. Thermography, another condition monitoring method that has been used for damage detection of bearings in rail vehicles [6], is not as effective a method as vibration or

acoustic measurements, since it has been noticed that substantial rise in temperature usually only occurs in the last stages of life of the bearings. Finally, acoustic emission monitoring may need huge amounts of data to be collected in order to capture rare burst events describing damage.

Machine operation involves the generation of forces and motions that produce different vibrational, acoustical, electrical, thermal and other residual processes [10]. This means that each machine can be modelled as a mechanical system having multiple input excitations and multiple outputs incorporating all these processes (a MIMO system), as shown in Figure 2.2. If h_{ij} designates the time-discretised impulse response function describing the causality relationship between an excitation s_j, and a response y_i, on the system examined then the convolution of the force signal with the impulse response function of the transmission path from the source to the measurement point shows the contribution to the response at one measurement point from the source [11]. Assuming linearity and no noise corruption in the response signal, this can be symbolically shown in the time and frequency domain as:

$$y_i = \sum_j h_{ij} * s_j \tag{2.1}$$

$$Y_i = \sum_j H_{ij} S_j \tag{2.2}$$

where the upper case letters represent the Fourier transforms of the lower case symbols, in equation (2.2).

What is obvious from the above is that in this case, where one has a MIMO system, the general response spectrum is not just the product of a single source spectrum and a single FRF. The basic problem of damage detection is to decide whether the diagnosed changes are due to a change at a source or in the system, known generally as a *blind source separation problem* (BSS) [6].

The determination of a relationship between a machine condition and a damage feature is the basis of all damage detection methods used in vibration-based condition monitoring. This relationship can become quite complicated due to the complexity of machinery, thus instead of an overall measurement of vibration level, representing the general machine condition, it has been shown that with the use of advanced signal processing methods, the evaluation of separate machine elements such as shafts, bearings and gearboxes is feasible and more desirable.

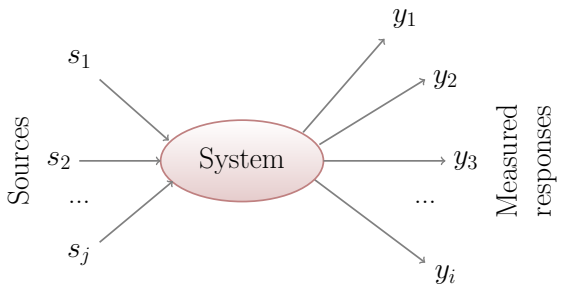


Figure 2.2: The MIMO system describing the condition monitoring problem [6].

2.3 Basic signal processing methods used for feature extraction in condition monitoring

This section contains an overview of the most important signal processing methods that have been used for condition monitoring over the years. These methods can be categorised into time-domain, transformed-domain and time-frequency analysis methods. Reviews of machinery diagnostic and prognostic methods in condition monitoring can be found in [12, 13]. Figure 2.4 at the end of the section summarises the signal processing techniques described.

2.3.1 Time-domain methods

Some of the most simple time-domain methods that have been used for condition monitoring are time-invariant analysis methods, based on probability distributions and density theory. This is a group of methods that use univariate features. The principle of these methods is often to simply use the overall vibration level to describe the general condition of the machine. Estimates used in this case are peak amplitude, peak-to-peak value, root mean square value of the signal, crest factor, and statistical moments such as *variance*, *skewness* and *kurtosis*. Figure 2.3 shows the equations used when analysing Gaussian random signals. Scientific studies employing these parameters can be found in references [14–17].

Another technique, applied specifically to gear vibration signals, is based on the envelope estimation of the signals analysed. In these studies, amplitude and phase demodulation techniques are used to detect fatigue cracks in gears

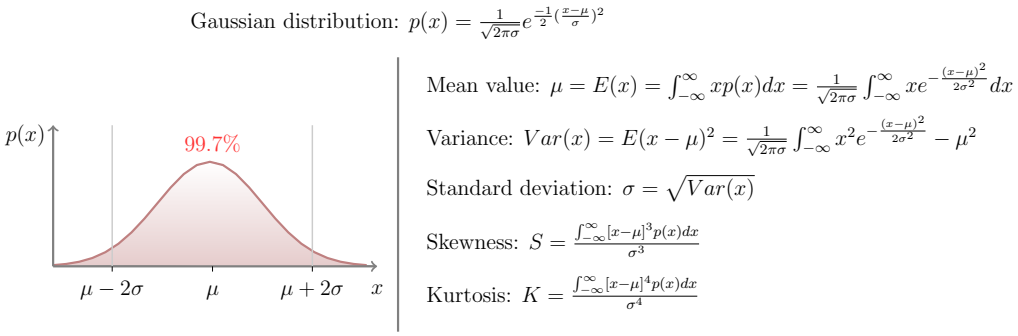


Figure 2.3: Probability density function for the Gaussian distribution.

from the estimated envelope and instantaneous phase [18, 19]. Most commonly in this case, the *Hilbert Transform* (HT) is used in order to extract the instantaneous amplitude and phase of the vibration signal from the analytic signal (phase unwrapping). In addition, for the diagnostics of rolling element bearings, it has been shown that the analysis of the envelope signals is recommended, since the analysis of the raw signals does not always give enough information [20].

The ideas of time-series analysis have also been applied in condition monitoring. Useful features in this case can be generated if one identifies a good time-series model for the CM system examined when it is undamaged. This model can be used to make good predictions of response of that structure and estimate the variance of the error signal (residual), between the model outputs and measured outputs. High variance in this case shows that the system is damaged, since the model failed to make good predictions and a change in the system is implied [21]. This approach has been applied and is known in gearbox condition monitoring as *Time Domain Averaging* method or *Time Synchronous Averaging* (TDA) [22, 23]. The meshing vibration is removed from the averaged signal producing the residual signal used to detect faults, most of the time by estimating the statistical parameters mentioned in the previous paragraph.

2.3.2 Transformed-domain methods

The application of Fourier analysis in condition monitoring for the estimation of the vibration spectrum has been widely used, since the efficiency of the *Fast Fourier Transform* algorithm (FFT) made it a quite popular technique. The spectrum of the damaged gearbox is compared to its spectrum under the

normal (undamaged) condition, and appropriate filtering can also be applied in order to isolate frequency bands believed to be associated with specific kinds of faults. According to Randall [24], there can be observed several characteristic components in a gearbox vibration spectrum:

- tooth meshing frequency harmonics,
- ghost components corresponding to manufacturing errors,
- sidebands due to signal modulations or eccentricity and tooth faults,
- low frequency harmonics related to additive impulses of the shaft speed once per revolution,
- intermodulation components representing sum and difference frequencies of the above components.

Combining the above, one can associate specific types of damage in gearboxes according to the kind of change one observes in the spectrum. Intensive wear may change the toothmeshing harmonics; misalignment and tooth faults are potentially detected by analysing the sidebands of the spectrum (*modulation sidebands*); unbalance or misalignment of the shaft can be monitored by observing the harmonics associated with the shaft speed.

Concerning bearing diagnosis, most frequency-domain approaches rely on the detection of the characteristic rotational frequencies related to specific bearing element faults. For example, in the case of rolling element bearings, the formulae for the *ballpass frequency outer race* (BPFO), the *ballpass frequency inner race* (BPFI), the *fundamental train frequency* (FTF), and the *ball spin frequency* (BSF) are shown in Table 2.1, [25].

Bearing faults in the early stages give sharp impulses that cover wide frequency ranges (upto 100kHz) [6]. These frequencies result from short duration pulses generated whenever a local defect on an element interacts with its mating element. As mentioned in the previous section, the analysis of the envelope signal is preferable, since the raw signal may not expose the information needed, so in this case the envelope spectrum or cepstrum can be used for damage detection.

Cepstrum analysis was introduced in reference [26]. It is the inverse Fourier transform of the logarithmic power spectrum. It has been quite useful in cases where the vibration signals contain families of harmonics and sidebands that characterise the fault. Some applications of the vibration spectrum and cepstrum are given in references [27–29].

Table 2.1: Characteristic frequencies of rolling element bearings

$BPFO = \frac{nf_r}{2}(1 - \frac{d}{D}cos\phi)$
$BPFI = \frac{nf_r}{2}(1 + \frac{d}{D}cos\phi)$
$FTF = \frac{f_r}{2}(1 - \frac{d}{D}cos\phi)$
f_r : Shaft speed
n : number of rolling elements
d : Diameter of the rolling element
D : Race diameter
ϕ : angle of the load from the radial plane

Another approach that can be found in this category is the use of parametric models, such as the *Auto Regressive* (AR) or the *Autoregressive with eXogenous inputs* (ARX) models that can be used to estimate frequency domain features using the harmonic probing algorithm [30]. In the case of AR models, frequency analysis can be employed due to the fact that the transfer function extracted during the process has the same spectrum as the signal being analysed when the excitation is white, which means that the poles of the transfer function represent the frequency components of the signal [31]. In the case of ARX models, the FRFs of the signals can be extracted. Other methods based on parametric modelling, are *adaptive noise cancellation* (ANC), *self-adaptive noise cancellation*, (SANC), and *discrete random separation* methods [32, 33]; they are used to filter unwanted components of the signal.

2.3.3 Time-frequency/time-scale analysis methods

The importance of time-frequency analysis for the signal processing of non-stationary signals is summarised in the following: since the *Fourier Transform* (FT) does not explicitly reflect a signal's time varying nature because it requires integration all over time and usually the vibration signals being analysed change over time, the FT and as a consequence the classical frequency domain approach for signal processing of nonstationary signals is generally inadequate. That is the reason why time-frequency analysis methods have

gained much attention in the condition monitoring field, since faults often manifest in the vibration signals as transient events. Apart from that, in many cases other influences, such as time-varying loads for the case of wind turbine gearboxes for example, may cause changes to the vibration signals over time that might influence damage features. A more detailed analysis of the above topics will be present in the next chapters of this thesis.

The simplest time-frequency method is the *Short Time Fourier Transform*, (STFT). Comparing the signal being analysed with elementary functions that are localised in the time and frequency domains is the basic idea behind the STFT. The *Wavelet Transform* (WT), is another well known method, having the same concept as the STFT. The difference between the two methods is in the different basis/elementary functions chosen during the process that lead to different signal representations. Because of that, wavelets give better localisation properties at high frequencies and are useful for detecting local events in the signals.

In order to be more accurate, a more appropriate signal processing category to which wavelets belong is that of time-scale analysis, since the WT decomposes a signal into different scales rather than actual frequency bands. Wavelet analysis is probably the most popular technique; a review of condition monitoring applications is found in reference [34]. Some examples of WT applications in condition monitoring are given in references [35–39].

Another approach to processing nonstationary signals is the *Cohen class time-frequency distributions*. In this class many different distributions have been introduced with the *Wigner-Ville* (WVD) and the *Choi-Williams* (CWD) distributions being the most famous ones. Unlike the power spectrum produced by the STFT, or the time-dependent spectrum of the WT, the energy distribution in the joint time-frequency domain, such as the WVD, is very complicated [40] as the underlying transforms are nonlinear. Condition monitoring applications of the WVD can be found in [41–43].

Relatively recently, the *empirical mode decomposition* (EMD) method was also proposed [44]. Since then, attention was gained in applying the EMD in the damage detection field [45, 46]. This technique decomposes the signal into a number of meaningful signal components, representing simple oscillatory modes matched to the specific data. This is one of the basic advantages of the EMD when compared to other time-frequency methods. After decomposing the vibration signal, the instantaneous frequency and amplitude of each component can be estimated, most commonly by applying the HT. A review of EMD condition monitoring applications is given in reference [47]. Apart from the original EMD algorithm, new versions of it have been pro-

posed in order to improve its performance. In order for example to overcome a problem of the EMD that has been observed in specific cases of signals where intermittencies exist [48], known as *mode mixing*, the *ensemble empirical mode decomposition* (EEMD) was proposed [49, 50]. Other EMD algorithms that have been proposed are the *complementary ensemble empirical mode decomposition* (CEEMD) [51] and the *bivariate empirical mode decomposition* (BEMD) [52]. Applications of the EMD algorithm and its alternatives to bearings can be found in references [53, 54]; applications to rotors can be found in references [55–57] and applications to gears can be found in [45, 46, 58, 59].

Generally, each technique has advantages and drawbacks, but it seems that the main reason for the popularity of the most recent method, the EMD, is that apart from offering similar results in terms of resolution to other techniques, it is also an adaptive method. *Adaptive time-frequency analysis* is the area of latest interest to signal processing researchers, but with the aim as well of creating a method with a better mathematical background than that of the EMD. One of the latest techniques towards this direction is the *synchrosqueezing transform* (ST) as an EMD-like tool but with a WT mathematical background. Its first application in the condition monitoring field is given in reference [60].

2.3.4 Other methods

Cyclostationary analysis and *spectral correlation* have also been applied in the condition monitoring of bearings and gears [61–63]. Cyclostationary analysis gives the possibility to separate signals with cyclostationary properties of 2^{nd} order [6]. In addition, different cyclic frequencies can be separated from each other as well. By definition, an n^{th}-order cyclostationary signal is one whose n^{th} order statistics are periodic. Vibration signals from gears and bearings are considered cyclostationary and pseudo-cyclostationary, making it convenient to use cyclostationary analysis in order to analyse them. During this process the spectral correlation, which is the two-dimensional FT performed on the two-dimensional autocorrelation function, giving eventually the spectrum of the squared envelope of the signal, as well as the Wigner-Ville Spectrum, can be obtained.

Another category of methods, not described up to now, is that of data compression methods used to reduce the dimensions of data. Such methods are needed because of the limitations of storage but also because of the problems that high dimensionality raises in terms of pattern recognition and machine

learning [64]. Data compression can be achieved through transformation and from this point of view the Fourier transform or the wavelets can be used as data compression methods. *Principal component analysis* (PCA) is a classical method from the field of multivariate statistics, otherwise known as the *Karhunen-Loeve Transform/Hotelling Transform*. It is a data compression method, but also a decomposition method that decomposes a signal spectrally from low frequency bands to high frequency bands. Just like the Fourier analysis, PCA is based on projection onto a set of basis vectors, with the difference that this basis is generated from the data, not fixed and predetermined. Similar methods to PCA are *linear discriminant analysis* (LDA), *independent component analysis* (ICA), *self-organising maps* (SOM) and *discriminant diffusion map analysis* (DDMA) [65].

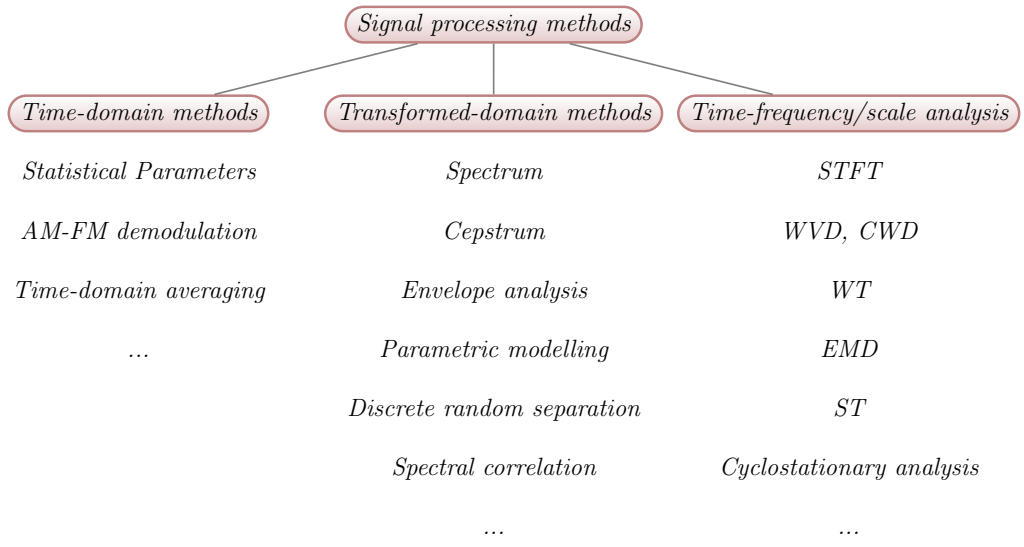


Figure 2.4: Basic condition monitoring methods.

2.4 Pattern recognition and natural computing algorithms

A good tutorial overview of natural computing algorithms can be found in reference [64]. The term *natural* or *soft computing* refers mainly to artificial neural networks, fuzzy logic, machine learning and genetic algorithms, that allow exploration of, or learning from, data.

Pattern recognition, as previously mentioned, deals with the problem of de-

ciding whether the features taken either from raw or analysed signals using the signal processing techniques of Section 2.3 have arisen from a damaged or undamaged structure. Therefore, pattern recognition describes the highest levels of the hierarchy of SHM and CM (Rytter's hierarchy, Figure 2.1).

Supervised learning approaches to pattern recognition used in the condition monitoring field include: Bayesian classification methods, nearest-neighbour search, artificial neural network classifiers [66, 67] and more recently support vector machines [68]. *Unsupervised learning approaches* have received less attention in the condition monitoring field [8]; an example can be found in reference [69] where one-class support vector machines were used. In practice, however it is not easy to apply supervised learning techniques due to the lack of training data required for the models corresponding to damaged systems, that is why unsupervised learning/novelty detection approaches might probably be of more use for condition monitoring.

The following summarises some basic ideas of these methods:

- *Artificial Intelligence approaches:* include *artificial neural networks* (ANNs), *fuzzy logic* (FL) and *evolutionary algorithms* (EAs).

  ANNs are computational models whose original purpose was the learning of the human brain structure and the imitation of its processes. In mechanical systems research terms the ANNs can solve four problems:

  – Autoassociation: reconstruction of a signal from noisy or incomplete data.

  – Regression: input-output mapping.

  – Classification: assignment of input data to given classes.

  – Detection: find abnormalities in the input data.

  The first two problems are related to modelling using neural networks and the last category includes the novelty detection problem. There are various ANNs but the most common is the *feedforward neural network* (FNN), in which the information moves in only one direction, from the input nodes through the hidden nodes, to the output nodes. The *multi-layer perceptron* (MLP) is an FNN (Figure 2.5) and there have been several applications of it in machinery fault diagnostics, examples are references [70–72]; these are supervised learning approaches.

  FL is a form of logic theory. Fuzzy set theory allows for approximate values and inferences as well as incomplete or ambiguous data (fuzzy data) as opposed to only relying on certain data (binary yes/no

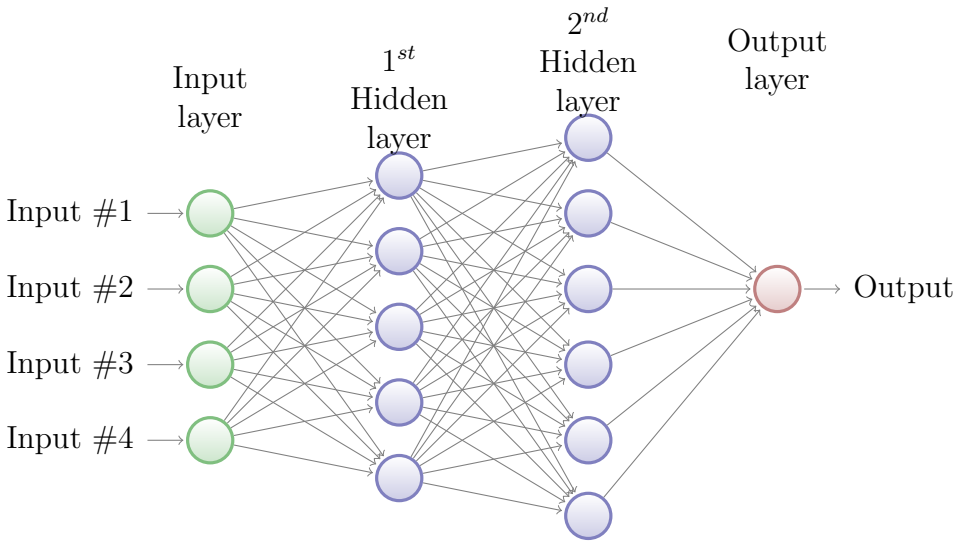


Figure 2.5: An example of a MLP. In this case, two hidden layers exist.

choices). An example of applying FL is given in [73], where frequency spectra representing various element bearing faults are classified.

EAs are natural optimisation methods that use mechanisms inspired by biological evolution and physics. The simulated annealing algorithm is the simplest of the optimisation heuristics, and genetic algorithms are the most popular of the EAs. Some applications of EAs are given in references [74–76].

- *Statistical approaches:* include *hypothesis testing* [77, 78], *statistical process control* [79], *cluster analysis*, *nearest-neighbour algorithms*, *support vector machines* (SVM) and *hidden Markov models* (HMM). Among these, cluster analysis methods have been used most often. In CM, cluster analysis as a term, describes multivariate statistical analysis methods that are used to group signals characterising different fault categories according to the similarity of the features they possess. The result of cluster analysis is a number of heterogeneous groups with homogeneous contents [12]; there are substantial differences between the groups, but the signals within a single group are similar. The easiest way of signal grouping is based on calculating distance measures or similarity measures between signals. Discriminant functions in statistical pattern recognition are used to derive these measures, the most common of which are the *Euclidean distance*, the *Mahalanobis distance*

[80, 81], the *Kullback–Leibler* distance and the *Bayesian distance*.

As discussed in reference [64] the future machine learning methods in mechanical systems and signal processing will be probably focused in the application of the Gaussian Process method, Variational Bayes, Markov-Chain Monte-Carlo and Deep Belief Networks.

2.5 Conclusions

The basic ideas of CM were discussed in this chapter, and the most common methods that have been used over the past years have been described briefly in a literature overview. So far, the subject of CM was discussed in general terms for all kinds of machinery that may exist in all industrial and other environments. This thesis however, will be focused on the condition monitoring of wind turbine gearboxes mainly and wind turbine bearings, to a lesser extent, and the solution of major problems faced in the feature extraction and pattern recognition part of CM. Despite the fact that CM is a mature field now, and that it would be a relatively easy task in laboratory environments, for the case of wind turbines, certain factors related to environmental and operational variations can make CM much more complex and difficult. These factors will be described in the next chapter.

Chapter 3

Wind turbines and CM in varying load conditions

In order to have a more clear understanding of the conditions prevailing in wind turbine systems a basic description of the wind turbines and their operating modes will follow.

3.1 Wind turbine structural overview

Wind turbines are devices that convert kinetic energy from the wind (wind energy) into mechanical energy, used to produce electricity. Wind turbines can be divided into two categories: the vertical axis and the horizontal axis, according to their rotor designs. The horizontal axis wind turbines are the most popular, and there exist two main designs for them: the up-wind and down-wind designs. Figure 3.1 shows the basic wind turbine components for an horizontal axis wind turbine.

These components are briefly described in the following [82]:

- The *rotor* is the rotating mechanism attached to the *blades* of the wind turbine.
- The *blades* are usually made from *glass reinforced plastic* (GRP) and incorporate lightning protection measures. In variable-blade-pitch wind turbines, the blade-pitch angle can be changed to change the torque delivered to the rotor.

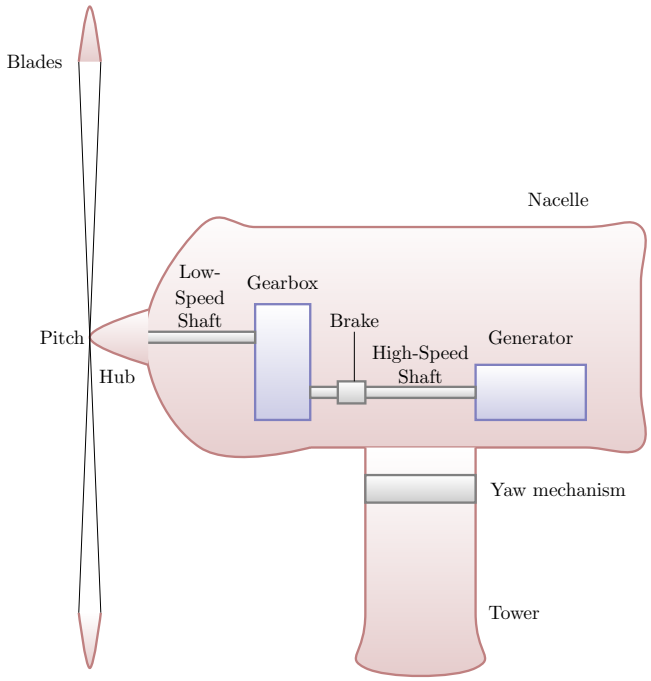


Figure 3.1: Horizontal axis wind turbine structural overview.

- The *hub* connects the rotor to the low speed shaft. This is where the blade-pitch actuator system is located.
- The *nacelle* is the "box" within which the main components are housed such as the shafts, the gearbox, the brakes, the generator, and the electrical components. In addition, the nacelle provides the structural connection between the tower and the rotor.
- In most wind turbine designs the rotor is connected to the generator through a *transmission system* that transmits the mechanical power of the rotor to the electrical generator. The gearbox increases the rotational speed of the shaft, in such a way that the speed is more suitable for driving the electrical generator. A wind turbine gearbox photograph is given in figure 3.2.

  Some wind turbine designs do not include gearboxes (gearless drive trains), and are known as direct drive wind turbines. These designs aimed to overcome reliability problems related to gearbox failures, often observed in wind turbines. To make up for a direct drive generator's slower spinning rate, the diameter of the generator's rotor is increased, hence containing more magnets which lets it create a lot of power when

Figure 3.2: A wind turbine gearbox.

turning slowly. To reduce the generator weight, permanent magnets in the generators' rotor can be used, while conventional turbine generators use electromagnets fed with electricity from the generator itself. The main issues of gearless wind turbines are that they are often heavier, that extremely rare and expensive elements are needed for the permanent magnet generators (Neodymium), and their reliability, although still not known for the offshore wind turbines, is not necessarily improved due to the lack of the gearbox, since their electrical components are very sensitive to the challenging conditions of their operational environment.

- Most wind turbines include a mechanical *brake* which can be used to stop the rotor. It can be applied during an emergency shutdown of the wind turbine.

- The *generator* converts the mechanical energy into electricity. In variable-speed wind turbines a converter is used to interface the generator to the AC grid.

- The *tower* is usually made of steel and its main purpose is to provide structural support to the nacelle and the rotor. A tall tower allows long blades, resulting in a large coverage area of the rotor and provides favourable aerodynamic conditions. The nacelle which is mounted on the top of the tower can be rotated by the *yaw mechanism*. In this way the rotor can be aligned with the wind. In upwind wind turbine designs, an active yaw mechanism is essential as the orientation of the rotor does not self-align with the wind.

Table 3.1: Wind Turbine Classes

Class	**Average Wind Speed [m/s]**	**Turbulence [%]**
IA	10	18%
IB	10	16%
IIA	8.5	18%
IIB	8.5	16%
IIIA	7.5	18%
IIIB	7.5	16%
IVA	6	18%
IVB	6	16%

The wind turbine dimensions (tower height, rotor diameters, blade length) are determined by the class of the wind turbine. Wind turbines are classified by the wind speed they are designed for, from class I to class IV, with A or B referring to the turbulence, as shown in the Table 3.1 [83].

Turbulence is a term that describes the fluctuations in wind speed, that originate from meteorological phenomena or large terrain elements [84]. Most of the times turbulence is described by the *turbulence intensity* which is the ratio that is given in percentage according to the following equation:

$$T_{intensity} = \frac{\sigma_{10min}}{v_{10min}} \tag{3.1}$$

where σ_{10min} is the standard deviation of the wind speed during 10 minutes and v_{10min} is the mean wind speed during 10 minutes.

3.1.1 Wind turbine operation modes

Wind turbines are designed to generate maximum power over a wide range of wind speeds. There is a maximum wind speed, commonly at around 60 m/s (216 km/h, 134 MPH), called survival speed, above which the wind turbines cannot survive. If the rated wind speed is exceeded the power has

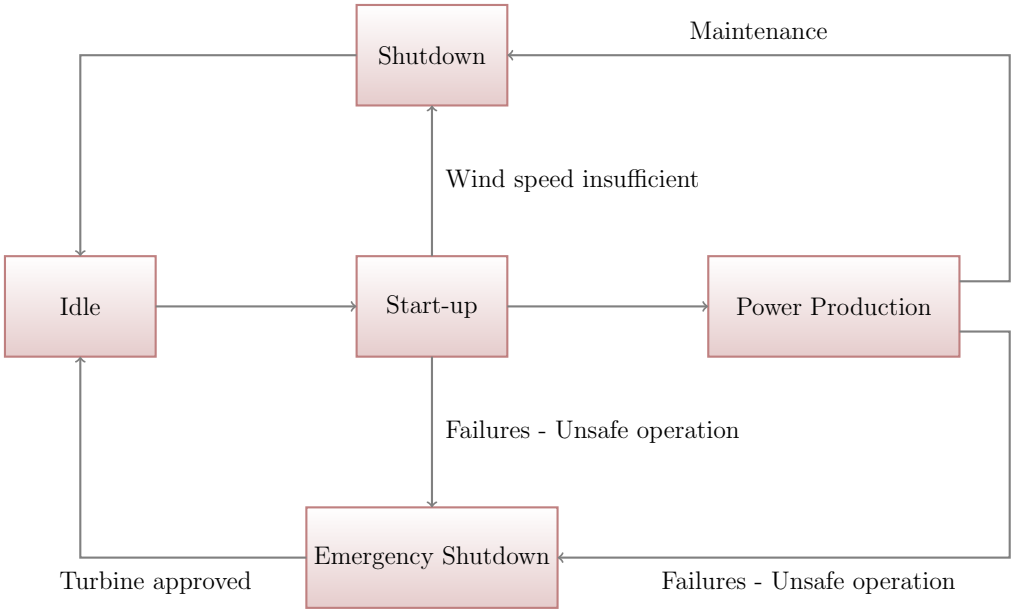


Figure 3.3: Operation modes [85].

to be limited. The operation of a wind turbine is controlled by the sequence controller, whose objective is to guide the wind turbine operation based on the wind conditions. The wind turbine has the following states of operation [85]:

- Idle
- Start-up
- Power production
- Shutdown
- Emergency shutdown

The start-up, the shutdown and the emergency shutdown states are transition states between the idle and power production states. Figure 3.3 illustrates the sequence of these operation modes.

In variable speed wind turbines, the decision of whether the wind turbine will operate and which mode of operation it should follow is done according the wind speed measurements. Figure 3.4 shows the power curve for an ideal (lossless) wind turbine. At very low wind speeds, there is insufficient torque exerted by the wind on the turbine blades to make them rotate and the wind turbine does not generate power (Region 1: No generation). The speed at

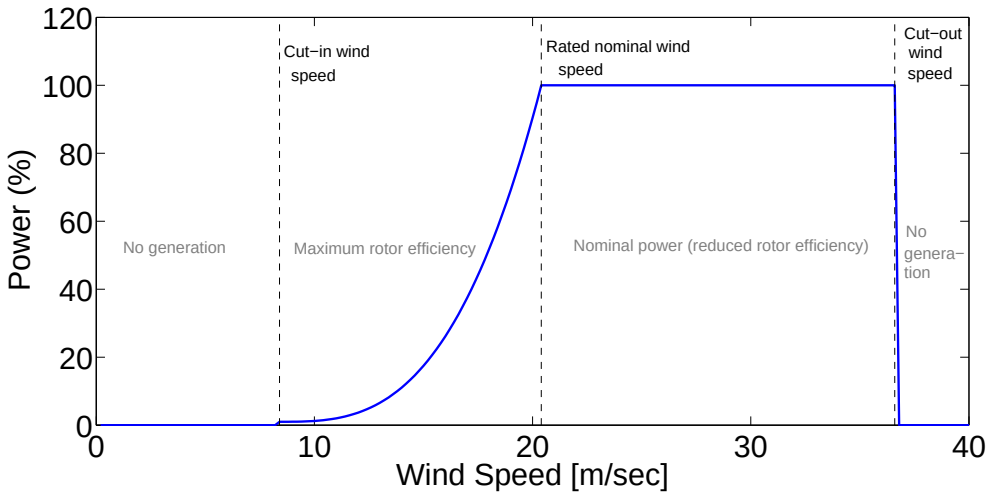


Figure 3.4: Typical wind turbine Power Curve.

which the turbine first starts to rotate and generate power is called the cut-in speed. As the wind speed rises above the cut-in speed, the level of electrical output power rises rapidly (Region 2: Maximum rotor efficiency). However, there is a point where the power output reaches the limit that the electrical generator is capable of. This limit to the generator output is called the rated power output and the wind speed at which it is reached is called the rated output wind speed (Region 3: Nominal power (reduced rotor efficiency)). With the increase of wind speed above the rated nominal wind speed, the forces on the turbine structure continue to rise and, at some point, there is a risk of damage to the rotor, that is why a braking system is employed to bring the rotor to a standstill (Region 4: No generation); this is called the cut-out speed. Above this wind speed the wind turbine doesn't operate.

The wind speed variations can be significant even during a single day of operation of a wind turbine. The wind is an intermittent resource, and even in the same wind farm, the turbines experience different wind shapes and wind turbulence resulting in the production of different wind power. In order to assess the frequency of wind speeds at a particular location, a probability distribution function is often fitted to the observed data. Different locations will have different wind speed distributions. Usually the Weibull model is used to mirror the distribution of hourly/ten-minute wind speeds at many locations [83, 86].

Due to wind variations, but also the operational modes transitions of the wind turbines, wind turbine gearboxes may undergo load and speed variations.

The latter are relatively slower because of the huge inertia of the system, so in this thesis what will be mostly discussed is the first kind of variations which the author will sometimes refer to as operational variations. Varying or nonstationary loads may affect the vibration signals in a way that is described in the following sections of this chapter.

3.2 Nonstationarity and condition monitoring

Machine components generate characteristic vibration signals that separate them from the other components and distinguish their health conditions, as mentioned in the introduction.

Classical methods used in machinery diagnostics such as spectral analysis are based on the assumption that the vibration signals are stationary. More advanced approaches though, aim to solve problems related to the nonstationarity that is often observed in real world signals. An overview of these methods was given in Chapter 2. It was mentioned that, sometimes classic approaches are sufficient enough for the analysis, but in order to have the best results possible one should be aware of the kind of signals that one analyses and choose the most appropriate method for one's case. For most purposes, describing a signal as a mathematical function is too broad [87]. One can impose a little more order than this by classifying signals into different groups. In the following section there will be a brief description of the types of signals that exist, according to reference [87].

3.2.1 Classification of signals

Signals are functions of one or more independent variables that describe the behaviour or nature of some phenomenon.

One can categorise the signals into *deterministic* or *random*. Deterministic signals can be predicted for arbitrary times [88]. Random or probabilistic signals, on the other hand, cannot be predicted by closed analytical forms even though the physics of the system examined may be known. These signals can be described by probability laws (stochastics, statistics).

Deterministic signals can be categorised into *periodic* and *aperiodic*. *Periodic* signals complete a pattern within a measurable time frame, called a period,

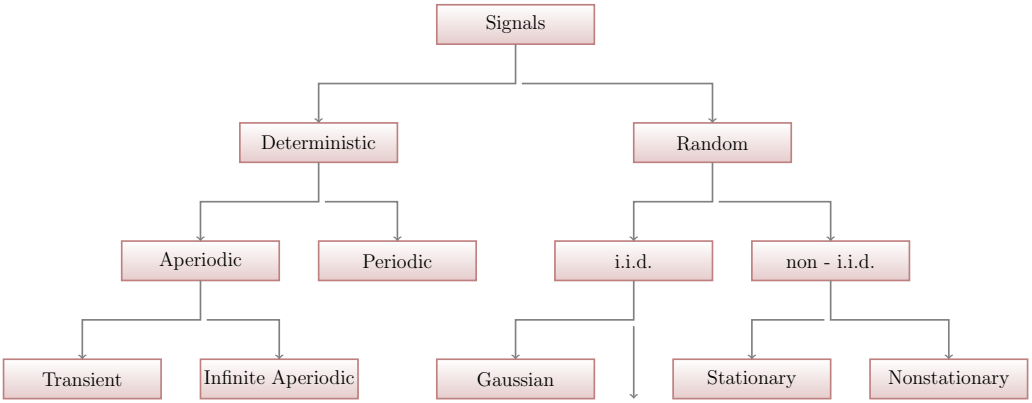


Figure 3.5: Classification of signal types.

and repeat that pattern over identical subsequent periods. *Aperiodic* signals are signals that are not periodic and can be divided into *transient* or *infinite aperiodic*. *Transient* signals are those existing for a finite range of time. Examples include impulse (hammer) excitation response of systems, explosion and shock loading, acoustic emissions, crack propagation (when the crack is the width of the structure, the structure fails).

Random signals can be divided into *independent and identically distributed* - i.i.d., for which one needs only to specify one distribution $p(x)$ describing the stochastic process of the random signal, and *non independent and identically distributed* - non i.i.d.. The most important i.i.d. processes arise when the distribution is Gaussian. The non i.i.d. processes can be divided into *stationary* and *nonstationary*. If the process's joint probability distribution does not change when shifted in time or space the signal is called (strictly or strongly) *stationary*. Consequently, parameters such as the mean and variance, if they exist, also do not change over time or position. If only the mean and standard deviation are constant, the signal is called weakly stationary. Figure 3.5 summarises the above signal categories.

Vibration signals of rotating machinery

Vibration signals generated by rotating machines [6] are not necessarily stationary in general. Vibrations tend to change with speed and load of the machine. A constant speed and load will typically generate stationary signals if the machine remains in normal condition. Certain types of damage however are shown as transient events in the vibration signals, such as gear tooth cracks or bearing cracks. Of course, machine run-up, shutdown and

braking conditions (transient events), produce nonstationary signals also. In addition, for the case of wind turbine gearboxes in particular, the type of input load and speed that they undertake are varying due to the highly turbulent wind conditions but also because of the operational wind turbine control which intends to maximise energy extraction and prevent turbine damages due to extreme wind conditions or grid connection faults [89]. The next section intends to describe the difficulties that may arise when trying to perform condition monitoring in wind turbine gearboxes that are related to these load variations.

3.2.2 The problem of varying load conditions in wind turbine gearboxes

Condition monitoring of wind turbine gearboxes faces a major challenge: the input of the system examined, the gearbox, can be a nonstationary signal, resulting in a nonstationary vibration. In addition, the dynamics describing a gearbox may be nonlinear, due to clearances found at its mating components (tooth backlash for example).

Signal processing of gearbox vibration signals under these conditions requires the use of more complicated tools than just a spectrum analysis. This is the reason why, some studies have initiated research on the influence of varying loads on condition monitoring. Some examples of these studies are given in the following. One of these studies, considering load variation in condition monitoring is found in reference [90]. In this paper the pseudo-WVD was used to produce the instantaneous power spectrum of the vibration data of a gearbox test rig in order to perform damage detection. In the recent past, condition monitoring of a planetary gearbox of a bucket wheel excavator used in lignite mines taking into consideration the varying load was examined in [91]. A band-pass filter was applied and the FFT spectrum as well as the WVD of the acceleration signal around the first harmonic of the planetary gearbox mesh frequency was examined. In this study it was also observed that the acceleration signal shows some effects related to the external load variation. In another study [92], the same datasets were used in order to compare the PCA method with the *canonical discriminant analysis* (CDA), method results. In addition, in reference [93] the same data sets were used again and this time the feature used for condition monitoring was the sum of the 10 amplitudes of the meshing components in the envelope spectrum.

Finally two other studies [94, 95], used gearbox rig experimental data in

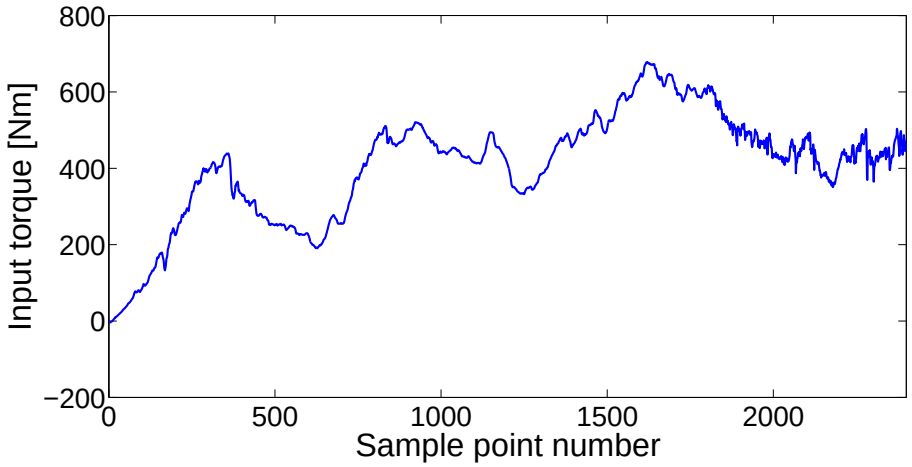


Figure 3.6: A simulated example of the input loading of a wind turbine gearbox.

order to test the effects of a sinusoidally-varying load when performing condition monitoring and in one of them a time-frequency analysis using the Morlet continuous wavelet transform was used and in the other a parametrical modelling analysis approach was proposed. The above studies confirm the problems related to varying load conditions when performing condition monitoring in gearboxes. Other such studies, performed in laboratory environment using similar methods can be found in references [96, 97].

Wind turbine gearboxes undergo much more challenging load variations than those described in the previous studies. This justifies the importance of research on this subject, in order to improve on-line wind turbine gearbox condition monitoring. In the previous section it was explained that wind turbine gearboxes load conditions are varying for two reasons:

- the wind variations during the day and the wind turbulence,
- and the effect of the operational wind turbine control system as described earlier.

References [89, 98] show measurements of wind turbine gearbox loads and speeds. An example of a typical input load of a wind turbine gearbox is given in Figure 3.6. This diagram was produced using the simulation model developed in reference [99]. This model not only takes into account the wind turbulences and control systems of the wind turbine but also the wake influences of the wind turbines in a wind farm according to the wind direction chosen for the simulation. The signal depicted is obviously nonstationary.

The gearbox vibration signals are directly influenced by these nonstationary effects.

Summarising what was explained previously, the gearbox vibration signals under these loads may be nonstationary. There might be amplitude and frequency modulations, as described in the next chapters of this thesis, that can only be identified using time-frequency analysis methods, and not just a classical Fourier analysis. Finally, the frequency bands related to damage might be influenced by the effects of the varying loads. These effects might not necessarily render the entire vibration signal nonstationary at all times, but could still affect the signal's frequency bands of interest, the ones that damage sensitive features are expected to create. In this case, potential damage features might be altered in an uncommon way, or even false alarms might be created, questioning the reliability of the condition monitoring method chosen.

3.3 Conclusions

In this chapter, the structural overview and the operating modes of a generic wind turbine were described briefly. It was explained that because of these operating modes and the quite significant wind variations, the loads of the wind turbine gearboxes are varying. In a CM context, this means that more sophisticated analysis methods should be chosen. In the next chapter, the experimental and simulated datasets used in the time-frequency analysis part of this thesis will be presented.

Chapter 4

Experimental and simulated datasets

The proposed condition monitoring approaches in this thesis are tested on both experimental and simulated datasets. Acceleration data from a real wind turbine gearbox in operation, and measurements from a bearing during a lab experiment constitute the experimental datasets that are going to be described. The gearbox model whose synthetic data will be used for the analysis in the next chapters is also going to be presented in detail here.

4.1 Experimental data description

4.1.1 Gearbox datasets

The experimental gearbox vibration data analysed in this study comes from an NEG Micon NM 1000/60 wind turbine in Germany. The measurements of the experimental data have been taken by members of the company EC Grupa, a Polish engineering company that maintains the wind turbine system from which the gearbox vibration datasets were obtained. The author did not participate in this process and therefore would like to thank Dr. Tomasz Barszcz and Prof. Wieslaw Staszewski for sharing the data. The gearbox is described by the kinematic scheme shown in Figure 4.1.

Generally wind turbine gearboxes are designed according to standards in order to meet the appropriate reliability criteria. ANSI/AGMA/AWEA 6006-

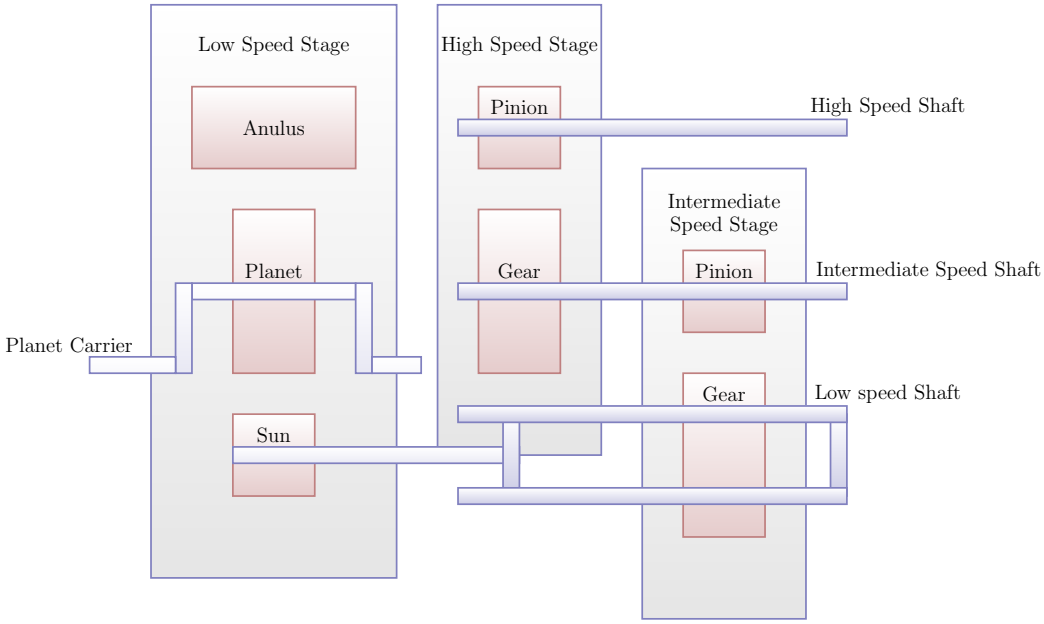


Figure 4.1: An example of a three-stage wind turbine gearbox.

A03: Standard for Design and Specification of gearboxes for wind turbines is one of the most commonly used [100]. Gearbox designs vary, although at present, the three-stage gearboxes with one- or two-stage planetary gear, and the remaining stages being helical or spur gears, are generally more common. Other designs include differential gearboxes and planet coupled gearboxes. The gearbox examined in this case (Figure 4.1), falls into the most common category, since it consists of three gear stages: one planetary gear stage and two spur gear stages.

The measurements come from a single accelerometer chosen by EC Grupa to be carrying more information concerning the damage of the gearbox. The sampling frequency of the measurements was 25000 Hz.

Acceleration signals from this gearbox were obtained at three different dates: 31/10/2009, 11/2/2010 and 4/4/2010. The first dataset was described as the one to be used as a reference. This is sensible because we expect damage to increase. If it transpired that the system was undamaged at the first test, then one would see clear signatures of damage in the later data. If it transpired that damage was already present during the first test, one can still use it as a reference and look for increased signatures of damage. The second dataset was considered to be one describing an early tooth damage of the gearbox and the third one was the dataset of the vibration signal with progressed tooth damage in the gearbox.

The kinematic scheme of the gearbox shows that it has two parallel spur gear stages and one planetary gear stage. When a gearbox has two or more mesh stages, signal processing of its vibration signals becomes more challenging because there are multiple shaft speeds and meshing frequencies apart from noise. This means, for the case of gear tooth damage, that one should examine the specific frequencies associated with the meshing frequencies and their harmonics of the specific stages that include the damaged gears. Without analysing the data, one can have an initial idea of the frequency bands that will be affected by damage. In this case, one should expect damage features at the highest frequency components (excluding noise) of the signal since the second parallel gear stage includes the fast generator shaft and because generally, it is known that it is in the harmonics of the meshing frequency of the damaged gear pair that damage features occur.

Figure 4.2 shows the time domain signal of the dataset obtained on 4/4/2010. The gearbox examined has a 28-tooth gear (smaller wheel) that meshes with an 86-tooth gear (bigger wheel) at the parallel gear stage II, which is the one at which damage was observed.

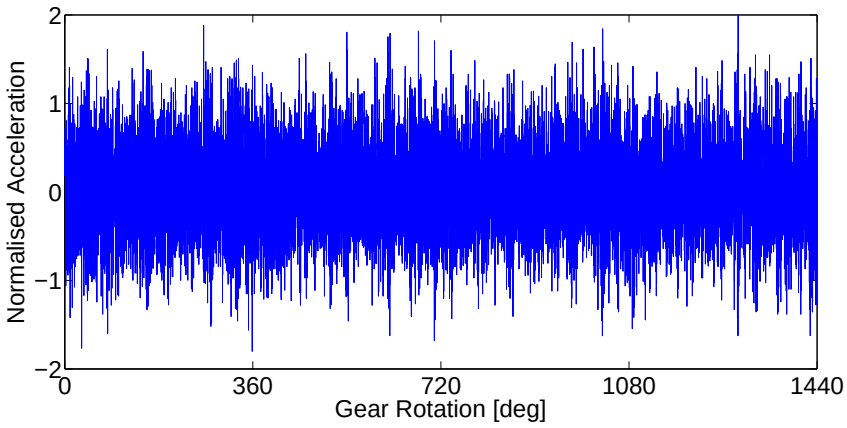


Figure 4.2: Time history signal of acceleration (data set 4/4/2010).

The frequency components of the signal as shown in the acceleration spectrum (Figure 4.3), are the following:

- 15.07, 30.14 Hz: relative meshing frequency and second harmonic of the planetary gear,
- 89.535, 179.07 Hz: relative meshing frequency and second harmonic of the 1^{st} parallel gear stage,

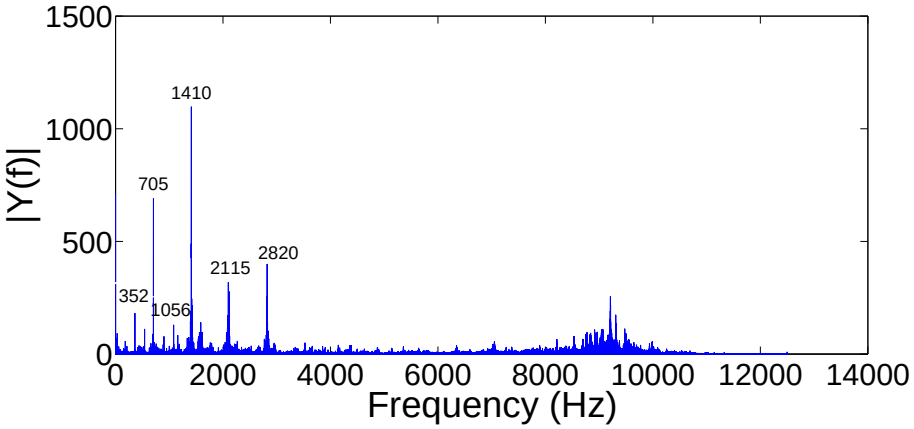


Figure 4.3: Spectrum of acceleration (data set 4/4/2010).

- 352, 705, 1056, 1410, 2115, 2820 Hz: relative meshing frequency and harmonics of the 2^{nd} parallel gear stage,
- Noise.

4.1.2 Bearing datasets

The bearing datasets were measured during a joint experiment with The Leonardo Centre for Tribology at the University of Sheffield that is participating in the Ricardo 'Multilife' wind energy bearing project. This project is aiming to increase service life of the 'Multilife' rolling element bearing by up to 500%. It is funded by the *UK northern wind innovation programme* (NWIP).

The '$MultiLife^{TM}$' mechanism is a concept developed by engineers at Ricardo UK Ltd, and aims to increase service life of wind turbine planetary stage gearbox bearings by up to 500 percent. The plan of the acoustic emission experiments is the following: the nominally stationary inner raceway is periodically rotated by a defined angular displacement so as to move 'fresh' inner raceway material into the loaded zone of the bearing, thus distributing damage around the entire raceway circumference.

The test rig used for this work houses $NU2244$ type cylindrical roller bearings, typical of the size utilized in the planetary stage of 10 MW wind turbine gearboxes. The rig is designed to ensure a comparable mode of operation to reality; such that the bearings are rotated via their outer raceways, at a

rotational speed similar to that expected of the planetary stage bearings in service (100 RPM) and a one directional radial load is applied via a hydraulic ram applying load through tension arms mounted to a non-rotating inner shaft. The bearing outer race is belt driven and the test bearing sits within a larger spherical roller bearing to permit this rotation.

The Leonardo Centre for Tribology at the University of Sheffield is participating in this project with the main research purpose, among others, to study and detect fatigue crack propagation in the inner race of the 'Multilife' rolling element bearing using acoustic emission and ultrasound measurements, mostly. For this reason, the test rig is currently located at one of the labs of the Department of Mechanical Engineering of the university. So participating in the experiments of the rig was of great interest.

The rig, that includes a stationary shaft, two pulley wheels and a hydraulic ram for the implementation of loads, a rig bearing and the test bearing is shown in Figure 4.4 (a). Figure 4.4 (b) shows the outer race and 16 the rollers of the rig. A schematic figure of the bearing is also given in Figure 4.5.

(a) (b)

Figure 4.4: (a) Ricardo 'Multilife' rolling element bearing (b) and its rolling elements.

The experimental datasets used further on in this thesis were taken on two different dates: 26/11/2012 and 28/11/2012. They are acceleration measurements obtained from a PCB piezotronics triaxial accelerometer, with model number $356B21$. The accelerometer was placed on the top of the bearing housing. Here only the x-axis datasets will be analysed. The sampling frequency was 10000 Hz. In addition, during the experiment, for both datasets, a radial load of 1000 kN was applied to the test bearing. On the first date the

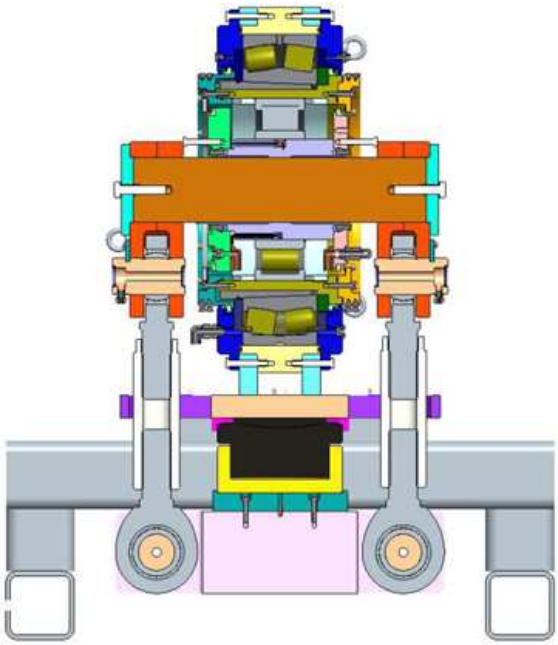


Figure 4.5: Ricardo bearing schematic.

rig was run under healthy conditions (no damage was introduced). On the second date however (28/11/2012), damage was introduced to the rig using the *electrical discharge machining* method (EDM), otherwise known as wire erosion. A notch, 0.2 mm deep and almost 0.2 mm wide, was introduced at the inner race of the bearing at the point of maximum load (bottom-dead-centre), as depicted in Figure 4.6, and the rig was run up to failure (around 300000 rotations).

The speed of the shaft was around 110 RPM for the first dataset and 100 RPM for the second dataset. Because the measurements were captured during a fatigue experiment the rotational speeds used were low and not the same for each dataset. When the rotational speed is low the changes on one revolution of the shaft can be influential and create difficulties in damage detection. The main bearings of wind turbine drivetrains operate under quite low speeds that vary over time. This is basically one of the main reasons why there is often not great diagnostic information in the raw spectrum of these signals, since smearing effects in the spectra could mask the repetition frequencies related to damage. In this case, the rotational speeds during the experiment were slowly increasing during the experiment.

Generally, if the spectrum of the vibration of a healthy bearing contains any information at all, then it is information related to the shaft rotation

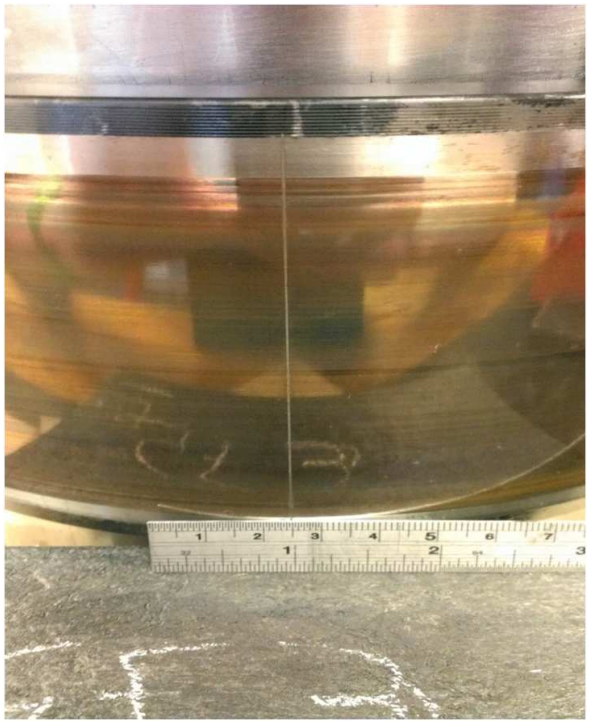


Figure 4.6: EDM notch introduced at the inner race.

speed and its harmonics, which is shown as *Zone I* in Figure 4.7. Any other frequencies might indicate noise, or frequencies related to other rotating machinery operating at the same time with the bearing examined. A rolling element bearing fault could appear at the inner, the outer race and/or on the rolling elements. During its early stages the damage on the surface is most of the time only local, e.g. pits or spalls. The vibration signal in this case includes repetitive impacts of the moving components on the defect. These impacts could create "repetition" frequencies that depend on whether the defect is on the inner or the outer race, or on the rolling element [101]. Apart from *Zone I* in this case, *Zones II*, *III* and/or *IV* might appear in the frequency spectrum of the vibration (Figure 4.7). Most of the times, only the vibration spectra of bearings with early faults contain information of damage, since with time these faults can eventually be smoothed and not give as sharp impulses. So for early faults, the repetition impulses might create initially an increase of frequencies in the ultrasonic frequency range, *Zone IV*, and maybe excite the resonant frequencies of the bearing parts later on, *Zone III*, as well as the repetition frequencies of *Zone II* (BSF, BPFO, BPFI) [39]. It has been observed in previous studies though that, many times the vibration of a damaged bearing might not carry the desired information, and that in

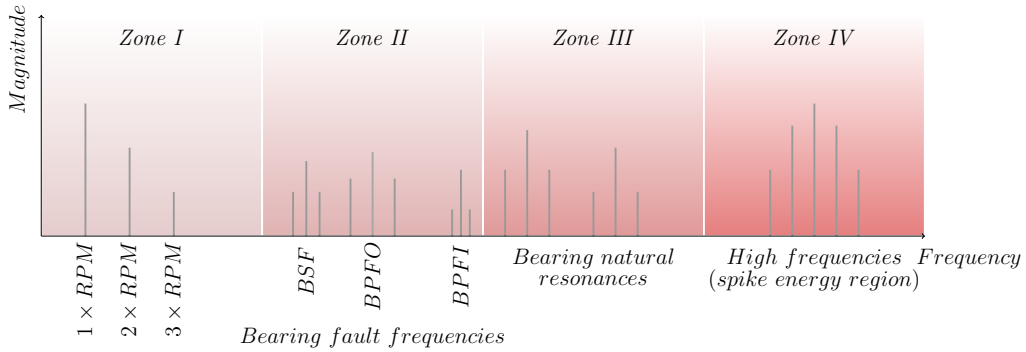
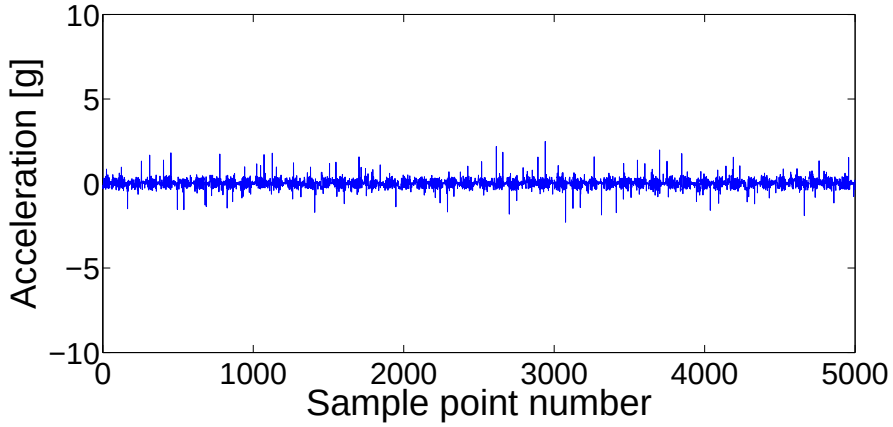


Figure 4.7: Frequency content of a signal of a damaged bearing.

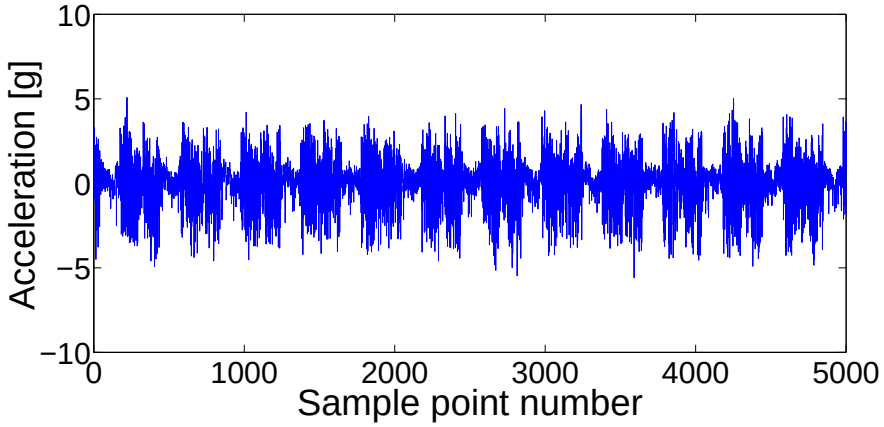
this case envelope analysis might be of more use for damage detection.

The test rig did not include other components, like gears, that could generate additional frequencies in the spectrum and generally it was observed that no significant amount of noise was in the vibration signals, so bandpass filtering in this particular case was not important. The acceleration diagrams for both datasets are given in Figure 4.8 and their frequency spectra are given in Figure 4.9.

The frequency spectrum of the first dataset is flat while the spectrum of the second dataset has more than 10 or so harmonics in the frequency bands of $6000 - 7000$ Hz and $2000 - 4000$ Hz as well as a peak at the frequency of 50 Hz and two harmonics at 100 and 150 Hz. The figures show normalised values of the frequency which were obtained by dividing the actual values with the sampling frequency. The harmonics of the higher frequency ranges are all separated by 50 Hz each. One could admit that this is a very apparent indication of damage to the rig even using this simple approach.

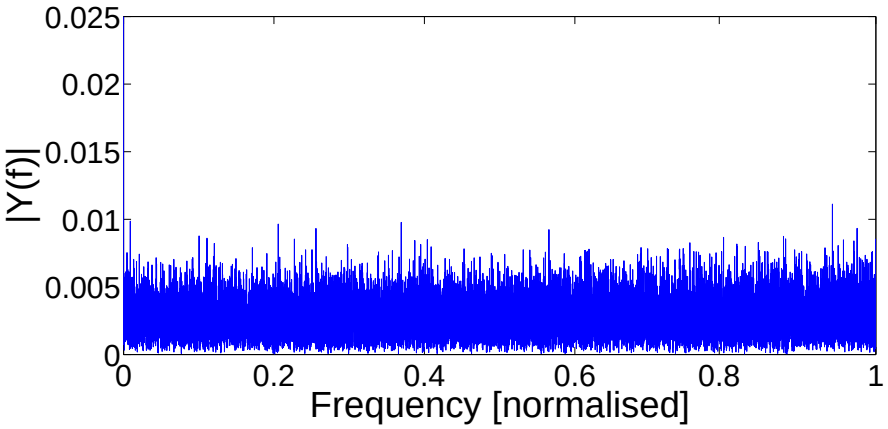


(a)

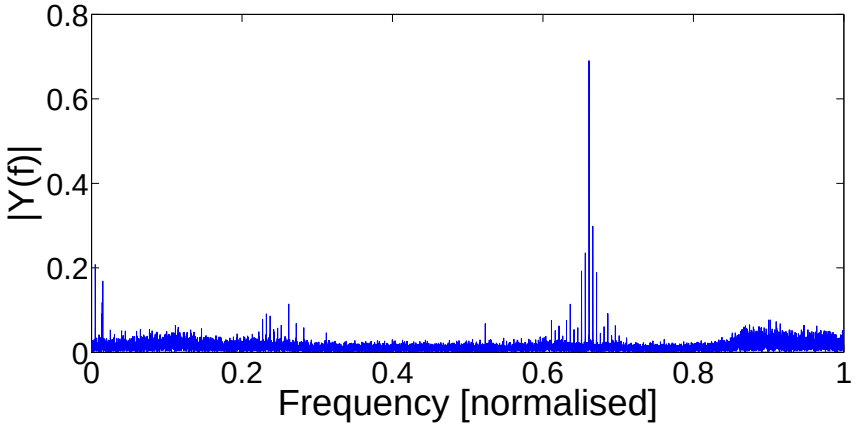


(b)

Figure 4.8: Time history of the acceleration: (a) dataset 26/11/2012 and (b) dataset 28/11/2012.



(a)



(b)

Figure 4.9: Frequency spectrum of the acceleration: (a) dataset 26/11/2012 and (b) dataset 28/11/2012.

4.2 Simulated data description

Many scientific teams have examined the subject of dynamic gear behaviour and various mathematical gear models have been developed in the past. These models can be categorised into linear time-varying and nonlinear time-varying. Linear gear models are arguably too simplified, and for the sake of accuracy, nonlinear models should be used. This is because the geometrical characteristics of the gear teeth affect the dynamics of the gear systems. The varying gear tooth contact ratio causes a variation of the gear meshing stiffness. Gear backlash, which is introduced either intentionally at the design stages or caused by wear, makes the equations of motion of gear systems nonlinear. This is the reason why many of the analytical gear models developed served in the study of nonlinearities in gears, parametric friction instabilities, nonlinear frequency response of gears and bifurcation [102–109]. Reviews of gear dynamic models are given in the references [110–112].

A generic geared-rotor bearing system consisting of a spur gear pair mounted on flexible shafts, supported by rolling element bearings [104], is shown in Figure 4.10 (a). The governing equations of motion can be given in matrix form as:

$$[\overline{M}]\{\ddot{\overline{q}}(\bar{t})\} + [\overline{C}]\{\dot{\overline{q}}(\bar{t})\} + [\overline{K}(\bar{t})]\{f(\overline{q}(\bar{t}))\} = \{\overline{F}(\bar{t})\} \tag{4.1}$$

where $[\overline{M}]$ is the time-invariant mass matrix, $\{\overline{q}(\bar{t})\}$ is the displacement vector, $[\overline{C}]$ is the damping matrix (assumed to be time-invariant), $[\overline{K}(\bar{t})]$ is the stiffness matrix, considered to be periodically time-varying, $\{f(\overline{q}(\bar{t}))\}$ is a nonlinear displacement vector that includes the radial clearances in bearings and gear backlash and $\{\overline{F}(\bar{t})\}$ the external torque and internal static transmission error excitations. In the equations, the overline is used to describe dimensional parameters in order to be distinguished from their nondimensional version (for which the same symbols are used but without the overline).

The system of Figure 4.10 (a) can be described by the *two-degree-of-freedom* (2 DOF) nonlinear model shown in Figure 4.10 (b), which has been used in previous studies [104, 106]. Its validity has also been compared with experimental results [104]. The gear mesh is described by a nonlinear displacement function $B(\overline{x}(\bar{t}))$ with time-varying stiffness $\overline{K}(\bar{t})$ and linear viscous damping $\overline{C}$. The bearings and shafts that support the gears are assumed to be rigid. The input torque fluctuation is included but the output torque is considered to be constant: $\overline{T}_1(\bar{t}) = T_{te}(\bar{t}) + \overline{T}_{1var}(\bar{t})$ and $\overline{T}_2(\bar{t}) = \overline{T}_{2m}$, with $\overline{T}_{2m}$ being the mean output torque, $T_{te}(\bar{t})$ is related to transmission error excitations and $\overline{T}_{1var}(\bar{t})$ the external input torque fluctuation. In this simulation $\overline{T}_1(\bar{t})$

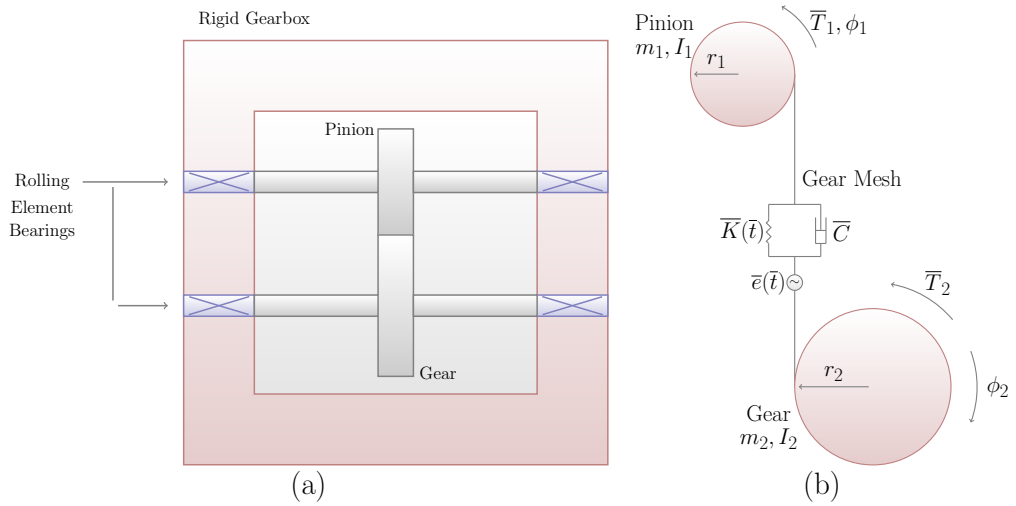


Figure 4.10: (a) A generic geared rotor bearing system and (b) its 2DOF nonlinear gear model.

includes a torque produced using the FAST design code [113], and will be discussed in detail in a following section. The main purpose of introducing this input to the model is to add loads similar to those that wind turbine gearboxes experience; previous models of gears have not included such a characteristic.

The equations of motion describing the gear model given in the general form are the following, where $\phi_i(\bar{t})$ is the torsional displacement, r_i is the base radius, and I_i is the mass moment of inertia of the i_{th} gear:

$$I_1\ddot{\phi}_1(\bar{t}) + r_1\overline{C}\dot{\overline{x}}(\bar{t}) + r_1 B(\overline{x}(\bar{t}))\overline{K}(\bar{t}) = \overline{T}_1(\bar{t}) \tag{4.2}$$

$$I_2\ddot{\phi}_2(\bar{t}) - r_2\overline{C}\dot{\overline{x}}(\bar{t}) - r_2 B(\overline{x}(\bar{t}))\overline{K}(\bar{t}) = -\overline{T}_2(\bar{t}) \tag{4.3}$$

The gear meshing frequency is given by:

$$\overline{\Omega}_{mesh} = n_1\overline{\Omega}_1 = n_2\overline{\Omega}_2 \tag{4.4}$$

where n_1 and n_2 stand for the number of teeth of each gear and $\overline{\Omega}_i$ is the rotating frequency of the i_{th} gear.

The *meshing stiffness* and the *static transmission error* are assumed to be periodic functions of time and can be expressed in a Fourier series form:

$$\overline{K}(\bar{t}) = \overline{K}(\bar{t} + \frac{2\pi}{\overline{\Omega}_{mesh}}) = \overline{K}_m + \sum_{j=1}^{\infty} \overline{K}_j cos(j\overline{\Omega}_{mesh}\bar{t} + \phi_{kj}) \tag{4.5}$$

$$\overline{e}(\bar{t}) = \overline{e}(\bar{t} + \frac{2\pi}{\overline{\Omega}_{mesh}}) = \sum_{j=1}^{\infty} \overline{e}_j cos(j\overline{\Omega}_{mesh}\bar{t} + \phi_{ej}) \tag{4.6}$$

The transmission error is the difference between the actual position of the output gear and the position it would occupy if the gear drive were manufactured perfectly. The transmission error in two meshing gears consists mainly of pitch error, profile error and run-out error.

The mesh stiffness on the other hand, varies due to the transition from single to double and from double to single pairs of teeth in contact.

The nonlinear displacement function related to the *gear backlash* nonlinearity is given by:

$$B(\overline{x}(\bar{t})) = \begin{cases} \overline{x}(\bar{t}) - b_g, & \overline{x}(\bar{t}) \geq b_g \\ 0, & -b_g < \overline{x}(\bar{t}) < b_g \\ \overline{x}(\bar{t}) + b_g, & \overline{x}(\bar{t}) \leq -b_g \end{cases} \tag{4.7}$$

where $2b_g$ represents the total gear backlash. The gear tooth backlash function controls the contact between teeth and allows for the fact that occasionally contact is lost.

The difference between the *dynamic transmission error* and the *static transmission error* is described as:

$$\overline{x}(\bar{t}) = r_1\phi_1(\bar{t}) - r_2\phi_2(\bar{t}) - \overline{e}(\bar{t}) \tag{4.8}$$

The equation (4.8) helps in describing the original model with a single equation of motion:

$$m_c\ddot{\overline{x}}(\bar{t}) + \overline{C}\dot{\overline{x}}(\bar{t}) + \overline{K}(\bar{t})B(\overline{x}(\bar{t})) = \overline{F}_m + \overline{F}_{te}(\bar{t}) + \overline{F}_{var}(\bar{t}) \tag{4.9}$$

where:

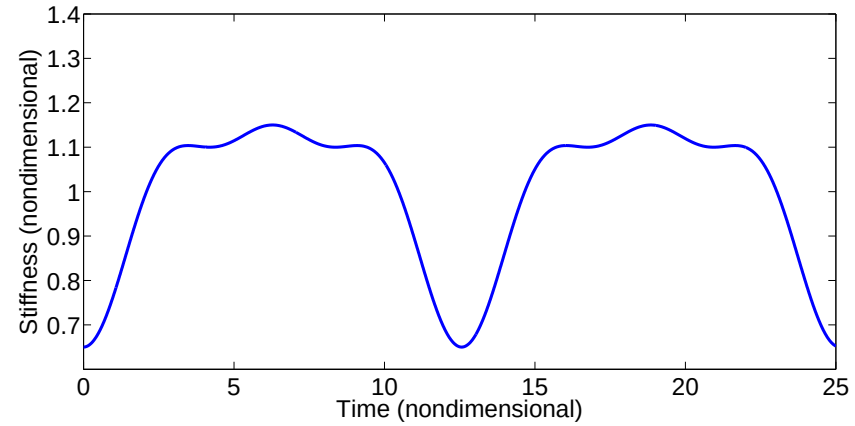
$$m_c = \frac{I_1 I_2}{I_1 r_2^2 + I_2 r_1^2} \tag{4.10}$$

$$\overline{F}_m = \frac{\overline{T}_{1m}}{r_1} = \frac{\overline{T}_{2m}}{r_2} \tag{4.11}$$

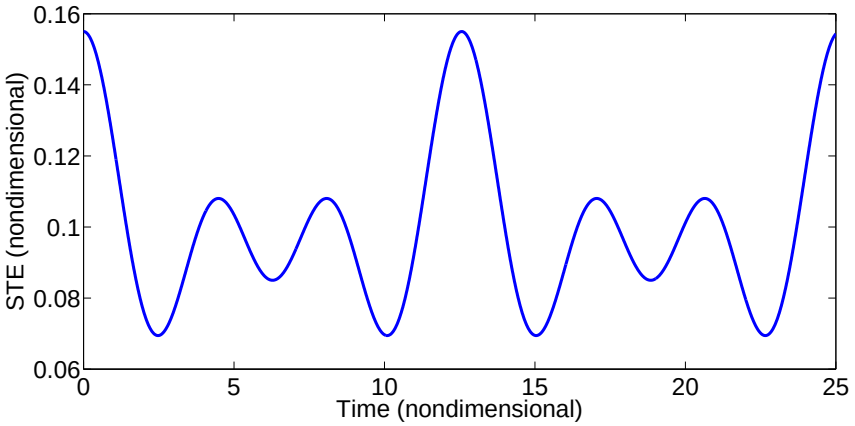
$$\overline{F}_{te}(\bar{t}) = -m_c\ddot{\overline{e}}(\bar{t}) \tag{4.12}$$

Now concerning the equations related to the FAST loads input to the model: $\overline{F}_{var}(\bar{t}) = \frac{r_1 I_2 \overline{T}_{1var}(\bar{t}) - r_2 I_1 \overline{T}_2(\bar{t})}{I_1 r_2^2 + I_2 r_1^2}$, $\overline{T}_{1var}(\bar{t}) = \overline{T}_{FAST} - \overline{T}_{FASTmean}$, $\overline{T}_{1m} = \overline{T}_{FASTmean}$. A specific section later on discusses these loads.

The equations of motion of the model can be written in a dimensionless form by letting: $x(\bar{t}) = \frac{\bar{x}(\bar{t})}{b_g}$, $w_n = \sqrt{\frac{\overline{K}_m}{m_c}}$, $z = \frac{\overline{C}}{2\sqrt{m_c \overline{K}_m}}$, $K(\bar{t}) = \frac{\overline{K}(\bar{t})}{\overline{K}_m}$, $F_{te}(\bar{t}) = \frac{\overline{F}_{te}(\bar{t})}{m_c b_g w_n^2} = \frac{-m_c \ddot{\bar{e}}(\bar{t})}{m_c b_g w_n^2}$, $F_{var}(\bar{t}) = \frac{\overline{F}_{var}(\bar{t})}{m_c b_g w_n^2}$ and $t = w_n \bar{t}$. Furthermore, the nondimensional excitation frequency is: $\Omega_{mesh} = \frac{\overline{\Omega}_{mesh}}{w_n}$. Figure 4.11 (a) shows the diagram of the gear meshing stiffness (nondimensional values) and Figure 4.11 (b) the diagram of the *static transmission error* (STE), used in the gear model simulations.



(a)



(b)

Figure 4.11: (a) The gear meshing stiffness function and (b) the STE function.

The nondimensional form of the original equation of motion is:

$$\ddot{x}(t) + 2z\dot{x}(t) + K(t)B(x(t)) = F_m + F_{te}(t) + Fvar(t) \tag{4.13}$$

where the backlash function can be expressed as:

$$B(x(t)) = \begin{cases} x(t) - 1, & x(t) \geq 1 \\ 0, & -1 < x(t) < 1 \\ x(t) + 1, & x(t) \leq -1 \end{cases} \tag{4.14}$$

Figure 4.12 shows the form of backlash function.

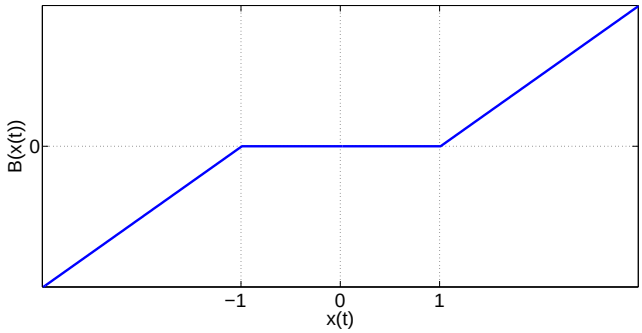


Figure 4.12: The backlash function.

The model was encoded in *Matlab* functions and was solved with the *ode45* differential equation solver, with a fixed step of 0.015. The parameters used for the simulation are chosen according to reference [104] and are described in Table 4.1.

In the simulations presented in this thesis for the case where no varying load is applied to the model, single-sided tooth impact was observed (tooth separation without back collision, backlash region $-1 < x(t) < 1$) for 20.6% of the simulation samples. For the case where the simulated varying load was applied, single-sided tooth separation was observed again for 23.5% of the simulation samples. In neither case was double-sided impact observed (backlash region $x(t) < -1$). The time varying meshing stiffness, due to the Fourier series form expression chosen to describe it and the non-dimensional values given in Table 1 ($k_1 = 0.2$, $k_2 = 0.1$ and $k_3 = 0.05$), has a $\pm 35\%$ variation around the mean. Due to the stiffness variations, the natural frequency varies also, in the range $0.8 < w_n < 1.16, (w_n = \sqrt{\frac{\overline{K}_m}{m_c}})$. The dimensionless meshing frequency is given in Table 1 and is 0.5; so a 0.015 time step corresponds to a sampling frequency with 67 points per cycle of the highest frequency of interest and is certainly large enough to prevent aliasing.

Changing certain values according to the reference [104] could help in order to have stronger nonlinear behaviour, but that was not the primary concern of the study.

Table 4.1: Simulation parameters

$I_{1,2} = 0.00115\ [kg\ m^2]$
$m_c = 0.23\ [kg]$
$r_{1,2} = 0.094\ [m]$
$number\ of\ teeth = 16$
$\overline{K}_m = 3.8 \cdot 10^8\ [\frac{N}{m}]$
$b_g = 0.1\ [m]$
$z = 0.05$
$\Omega_{mesh} = 0.5$
$F_m = 0.1$
$F_{te1} = 0.01$
$F_{te2} = 0.04$
$F_{te3} = 0.02$
$k_1 = 0.2$
$k_2 = 0.1$
$k_3 = 0.05$

Despite the fact that this is a very simple wind turbine gearbox model, since it consists of only one spur gear stage and it ignores the bearing vibrations, when in reality wind turbine gearboxes usually consist of three gear stages with one or two being planetary stages and the rest, parallel spur gear stages, it is sufficient for the current study. The reason is that the time frequency method that will be used further on has the ability to isolate specific frequency components of the signal. More gear stages in a vibration signal would mean more frequency components at different frequency bands. Damage at a specific gear stage would therefore be shown in the vibration signal associated with the meshing frequency and its harmonics of the gear stage examined. What is important in this gear model, and the main purpose of the simulations is to show how time varying loads similar to those observed in wind turbines influence the vibration signals and how the time-frequency analysis methods proposed help in this case to overcome the challenges produced by these influences.

Gear tooth faults

Tooth damage causes a reduction in gear tooth stiffness, thus it can be modelled here by reducing the meshing stiffness function periodically, with the period of the rotation of the gear with the damaged tooth. Tooth stiffness is a key parameter in gear dynamics concerning the determination of factors such as dynamic tooth loads and the vibration characteristics of the geared system. The changes due to the tooth damage appear in the vibration signals as amplitude and phase modulations [114–117]. Briefly, the mechanism is the following: when the defected tooth accepts the load, the crack opens gradually with the increase of load up to the point where the load reaches its maximum level. Then, the load is gradually transferred to nearby teeth and the crack steadily closes. This process occurs once per shaft revolution.

Concerning the simulation here, a gear fault was introduced into the model by reducing the meshing stiffness $\overline{K}(\bar{t})$ to 99.7% of the nominal gear meshing stiffness for 5 degrees of the shaft rotation, periodically, for every rotation of the damaged gear. This was done in previous studies also, see reference [96]. Figure 4.13 shows the reduction in stiffness for a tooth with crack. Figure 4.14 shows the acceleration responses of the simulation for the undamaged and damaged cases, for steady load conditions. The Fourier spectra of these diagrams are shown in Figure 4.15. It is clear that any differences due to damage are quite subtle.

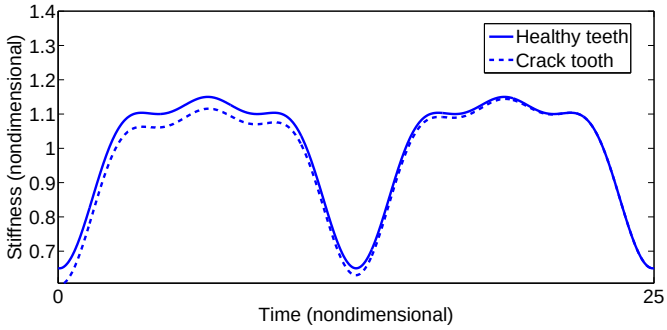
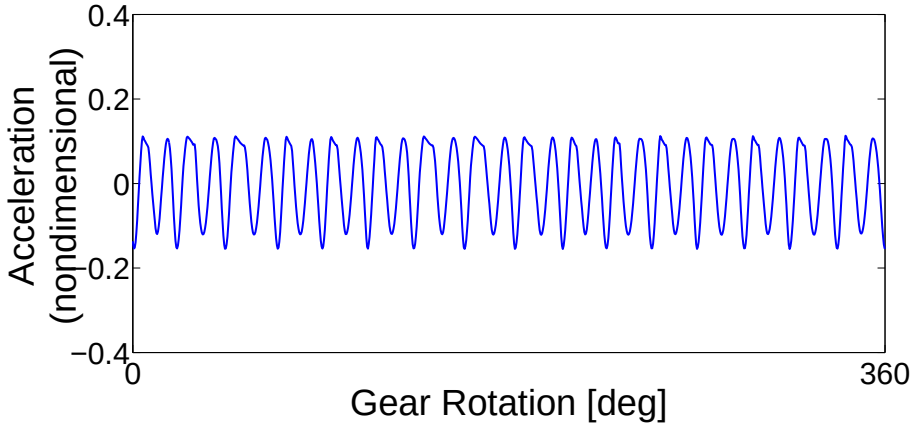


Figure 4.13: The meshing stiffness for healthy teeth and for a tooth with crack (0.3% stiffness reduction).

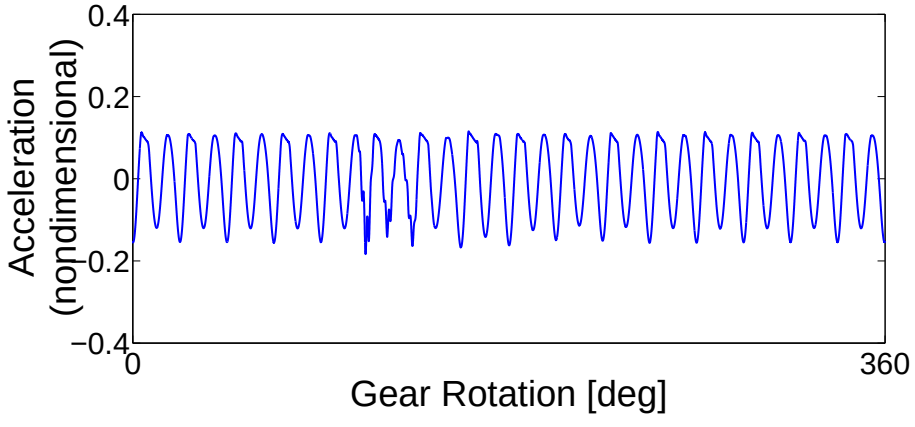
Simulation of time-varying loads

To simulate the load conditions corresponding to those of real wind turbines, a series of wind turbine aerodynamics codes, developed by the *national renewable energy laboratory* (NREL, US) is used in this work [113]. FAST is a design code written in FORTRAN that can be used to simulate wind turbine systems under a variety of operating conditions, including conditions related to wind speed variation and turbulence, and generator start-ups and shutdowns. Figure 4.16 shows the load cases examined in this work. Turbulent wind conditions are considered, with wind speed 12 m/s.

The load variation is shown in the time-acceleration plot as an increase in the acceleration amplitude when the sudden load changes occur. Such a load creates a nonstationarity in the acceleration signal. Concerning the values of this input torque simulated in FAST, $\overline{T}_{1var}(\bar{t})$, they correspond to a different system produced in FAST. For this reason it was important to do a normalisation before inserting the load to the simulated gear model. The equations of this normalisation are described in Section 4.2. The acceleration response for the undamaged and damaged cases and varying load conditions is shown in Figure 4.17. The influence of the time-varying load on the acceleration diagrams can be observed since the signals have become evidently less smooth and more nonstationary. Moreover, the damage is not as visible in the acceleration signal as shown in Figure 4.18. In addition, the Fourier spectra of these signals for the damaged and the undamaged cases are quite similar. Under these circumstances, it is even harder to detect damage in its early stages using conventional vibration monitoring techniques. That is why a time-frequency approach is proposed in this paper.

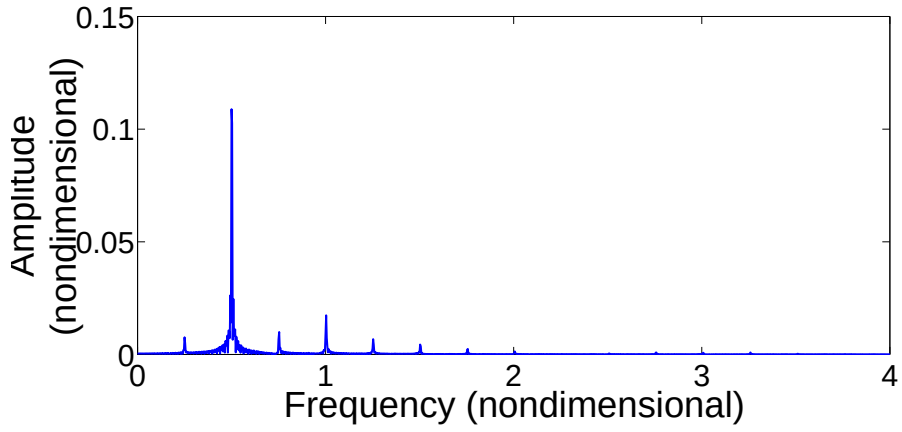


(a)

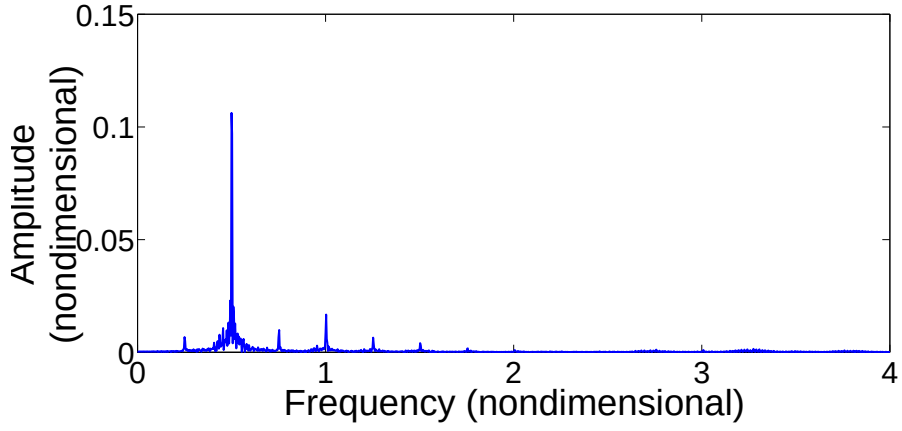


(b)

Figure 4.14: The acceleration diagram of the simulation: (a) undamaged and (b) damaged for the steady load conditions.



(a)



(b)

Figure 4.15: The Fourier Spectra of the simulation (a) undamaged and (b) damaged cases for $\overline{T}_{1var}(\bar{t}) = 0$ (steady load conditions).

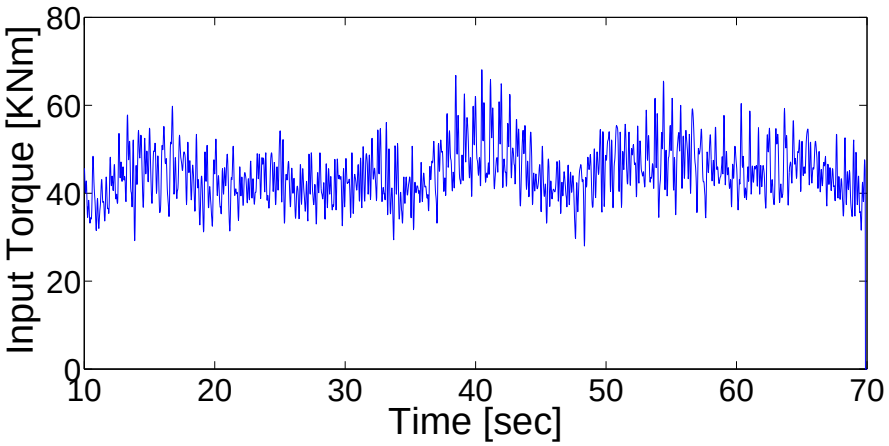
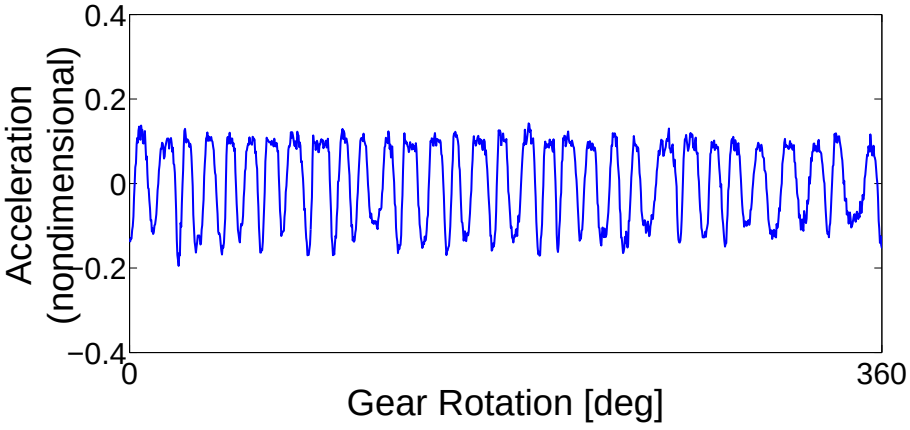


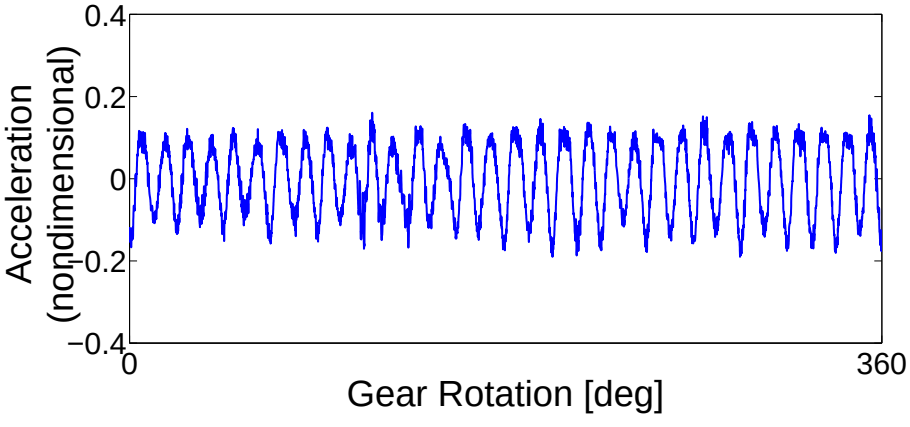
Figure 4.16: The high speed shaft torque for turbulent wind conditions (wind speed=12m/sec, FAST simulation).

4.3 Conclusions

The experimental and simulated datasets that will be used in the largest part of this work were presented in this chapter. In the following chapter the main signal processing method used will be presented in detail.

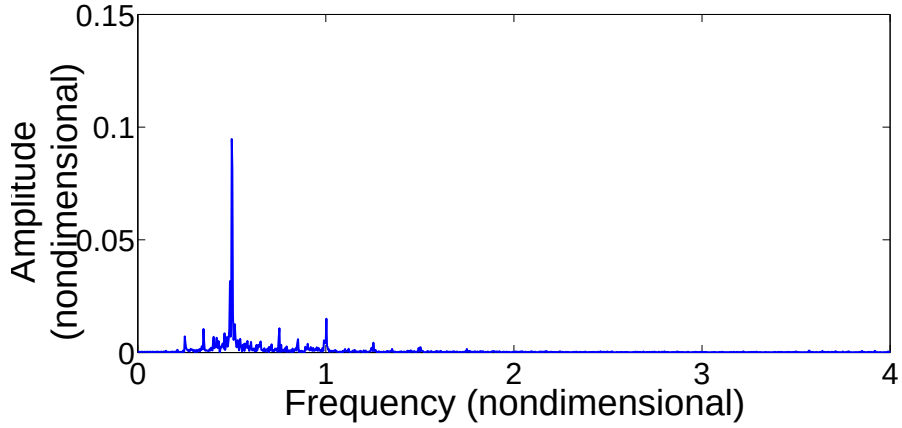


(a)

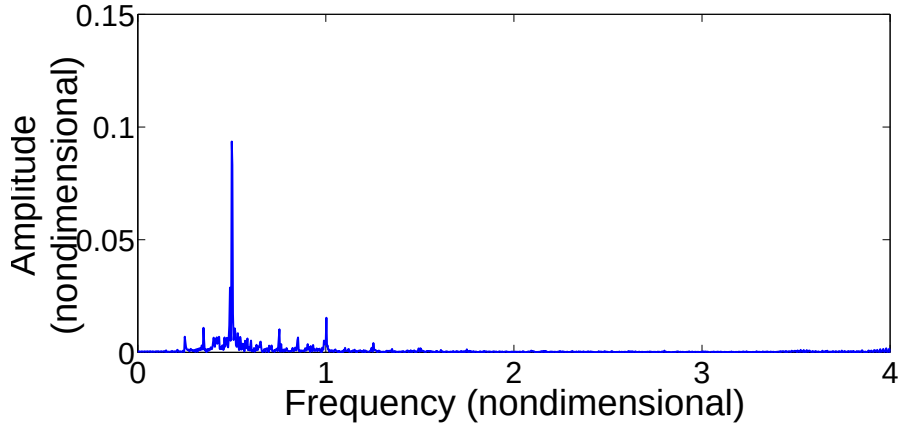


(b)

Figure 4.17: The acceleration diagram of the simulation: (a) undamaged and (b) damaged for the varying load conditions.



(a)



(b)

Figure 4.18: The Fourier Spectra of the simulation (a) undamaged and (b) damaged cases for the varying load condition.

Chapter 5

Time-frequency analysis in CM

The signal processing that is required in the case of wind turbine gearboxes should overcome the challenges of the sometimes existent nonstationarities in the gearbox vibration signals and the amplitude and frequency modulations produced by the load (and sometimes speed) variations.

In addition, because wind turbine gearboxes consist of multiple stages, and their vibrations signals are much more complex (having more frequency components), the signal processing methods used should be able to separate the signal's parts that are associated with the damaged stage.

Time-frequency analysis methods in this case have an advantage over conventional methods, e.g. the Fourier spectrum. First of all they act as filter banks, and second they are more appropriate when one wants to follow the time-varying properties of the signals, i.e. frequency or amplitude variations. Well known time-frequency analysis or time-scale analysis methods that have been applied for condition monitoring are the STFT [118], the WVD [41], the WT [36], and most recently the EMD [46]. Generally, each technique has advantages and drawbacks, but adaptive time-frequency analysis is the area of latest interest to signal processing researchers, explaining the increasing popularity of the EMD method.

For the *instantaneous amplitude* (IA) and *instantaneous frequency* (IF) estimation of the signal components there have been several publications mainly concentrating on the WVD and its variations, where the IF is defined through the mean moment of different components at a given time [119]. Other than

this approach, the IF obtained using the HT separation algorithm has gained much attention, and was used in combination with the EMD method extensively, [58, 59, 120–124]. Alternative methods for the IF estimation have recently emerged [125], with the *Teager-Kaiser energy operator* (TKEO) approach gaining interest from some scientific teams because of its simplicity as an algorithm.

5.1 Basic time-frequency/scale analysis methods

For any vibration signals being analysed that change over time, the FT does not explicitly reflect a signal's time-varying nature. The simplest way to overcome this is to compare the signal with elementary functions that are localised in the time and frequency domains; that is the basic idea behind the STFT and the WT, two well known now time-frequency methods. Different basis/elementary functions chosen lead to different signal representations. For the STFT, localised, harmonically related sinusoidal basis functions are used, which means that the functions contain both frequency and time information. In this transform the different elementary functions are obtained by frequency modulation, and have the same envelope, whereas on the contrary, the WT basis functions are achieved by time scaling and have different envelopes. So the time and frequency resolutions of the WT basis functions change with a scale factor. This results in a uniform bandwidth of the STFT functions in the frequency domain, and in a bandwidth of the WT basis that increases as frequencies increase. While the time-frequency resolutions of the STFT are uniform in the entire time-frequency domain, they vary in the WT [126]. At high frequencies, there is better time resolution and worse frequency resolution and at low frequencies there is better frequency resolution and worse time resolution.

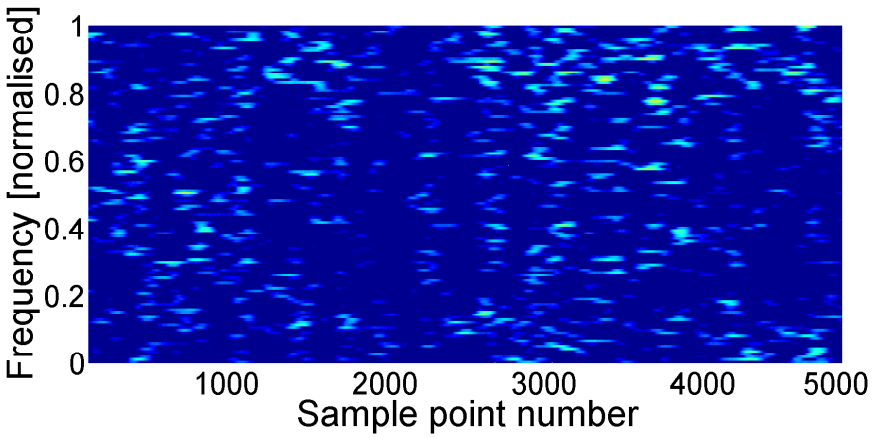
Since the STFT and WT time-frequency analysis methods are now well established, their theoretical background is not going to be presented in detail in this section. The reader who is interested in the details of these basic methods, can find some theory in the Appendix A of this thesis, as well as in the reference [126]. It might be of significant interest though, given this initial time-frequency analysis on the experimental datasets, to discuss briefly the potential features one can obtain for the case of condition monitoring of wind turbine bearings and gearboxes when performing time-frequency analysis.

As mentioned in previous chapters, rolling element bearing faults produce signals that are a series of high-frequency bursts, excited by near-periodic impacts. Because the diagnostic information is contained in the repetition frequency of the fault, a direct analysis of the vibration signal might not be sufficient at all times. Generally it is much easier to detect damage in bearings when it is at early stages, since the impulses in the vibration are sharper in this case. Figure 5.1 shows the STFT time-frequency representations of the acceleration of the bearing datasets presented in Chapter 4. A window of 256 points was used here with overlap 250 points. The scales are the same for both cases, and it is interesting that damage is quite apparent in the second case, in the frequency range between 6000 and 7000 Hz. The repetition phenomenon is very nicely depicted in this case in Figure 5.1 (b).

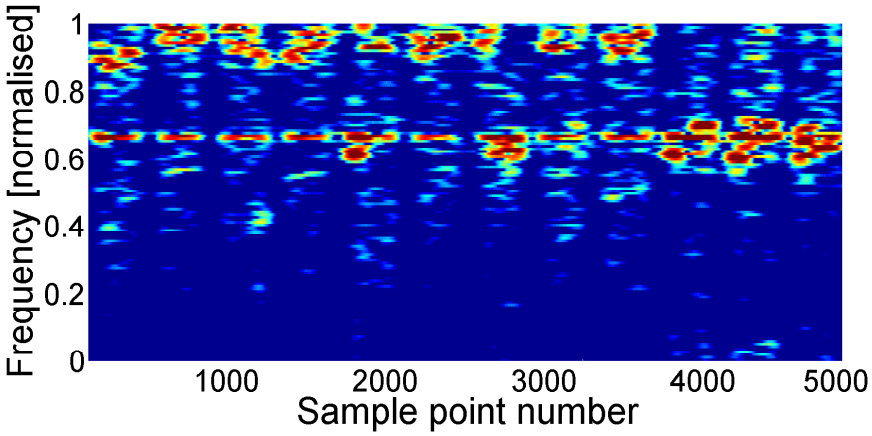
Bearing signals can be modelled as an amplitude modulation of the carrier signal at the resonance frequency by a near-periodic series of exponential pulses, which means that the most suitable approach would be to first extract the signal envelope and frequency analysed to reveal these repetition frequencies; this can be done with a amplitude-frequency demodulation algorithm such as the HT. Some of these algorithms are presented in Section 5.3 of this chapter. Time-frequency analysis of the instantaneous amplitude of the vibration signal can then be performed. A theoretical explanation behind the superiority of the analysis of the squared envelope compared to the acceleration signal is related to the fact that the spectrum of the squared envelope is equivalent to the integrated spectral correlation of the signal over the frequency band of interest, as shown in [101].

The vibration signal of a gear pair is characterised by the gear pair meshing frequency, $\overline{\Omega}_{mesh}$, as given in equation (4.4). This frequency and its harmonics, $k\overline{\Omega}_{mesh}$ where $k = 1, 2, 3...$ will appear in the frequency spectrum and as a consequence in the time-frequency representation of the vibration signal of the gear pair as well. Another characteristic frequency is the rotating frequency of each gear, whose relationship with the meshing frequency is given by the same equation (4.4).

If one of the two gears has tooth damage then, in the time-frequency analysis, there should appear damage features at the time instants where this tooth engages. Assuming that there are not great fluctuations in the rotating speed of the gears then the damage feature should have an almost periodic appearance in the time-frequency representation, with the period of the rotation of the damaged gear. In addition, this damage feature should appear at the frequency "regions", *frequency bands*, of the meshing frequency and its harmonics.



(a)



(b)

Figure 5.1: STFT representations of the acceleration of the bearing datasets.

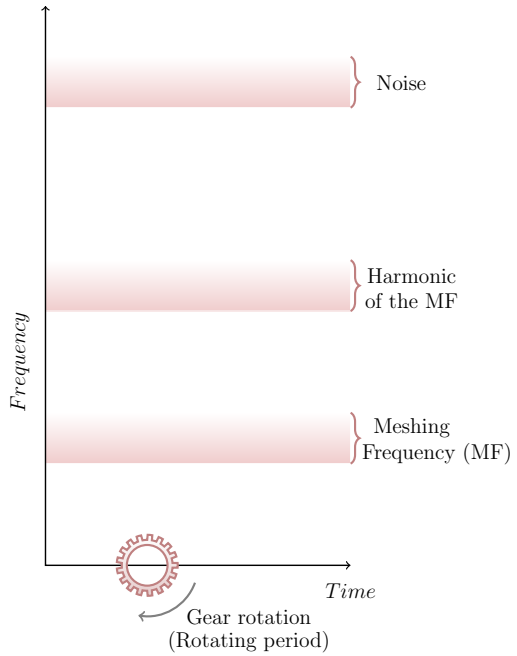
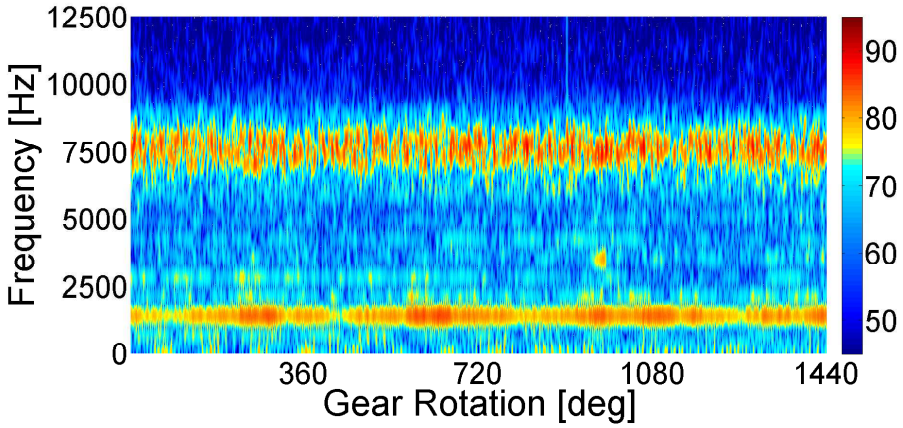


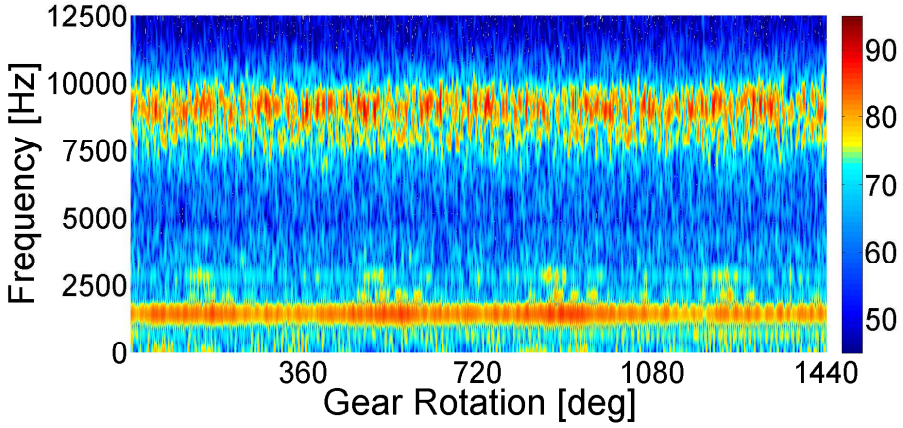
Figure 5.2: Explanation of the time-frequency representation of the vibration of a gear pair.

Figure 5.2 is given at this point in order to demonstrate in a simple way what is explained above. Apart from these basic frequency components in the vibration signal of the gear pair, other frequency components might exist as well, such as noise or ghost components. It is important to keep in mind though, that any damage features associated with tooth damage in the gears should have the frequency signature described in the previous paragraph.

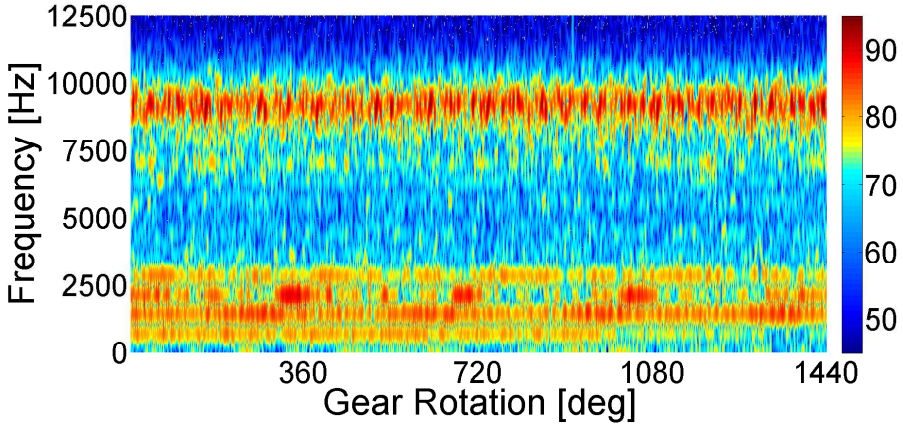
In order to have an example of what was explained previously, the STFT was applied to the wind turbine gearbox datasets described in Section 4.1 of this thesis. Figure 5.3 shows the STFT spectrograms of the three datasets. For the estimation of each spectrogram, the Welch method was used with a Hamming window of 64 points and an overlap equal to 60 points. Damage appears in the spectrogram of the third dataset at the frequencies around 2500Hz, that are related with the second harmonic of the relative meshing frequency of the second parallel gear stage. Damage can be seen as an increase of the power of the signal at the specific frequencies happening almost periodically, close to the rotating period of the damaged gear. The other two datasets, and specifically the second dataset that describes the gearbox at an early damaged condition have no significant damage indications with this method.



(a)



(b)



(c)

Figure 5.3: Time-frequency representation using the STFT for the three gearbox datasets examined, where (a) is dataset 31/10/2009, (b) dataset 11/2/2010 and (c) dataset 4/4/2010.

5.2 Adaptive time-frequency analysis

The EMD method is probably the most well known adaptive time frequency analysis method; in this section the theory behind it will be discussed.

5.2.1 The Empirical Mode Decomposition method

The EMD method decomposes the time-domain signal into a set of signal components (oscillatory functions) in the time-domain called *intrinsic mode functions* (IMFs). Each IMFs is associated with a frequency band of the signal, so the EMD method is a filter bank method, and can be used for isolating unwanted or interesting components of the signals being analysed. By definition, an IMF should satisfy the following conditions [44]: (a) the number of extrema and the number of zero crossings over the entire length of the IMF must be equal or differ at most by one, and (b) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. The EMD method is well known now, but the theory is added here for completeness.

The EMD decomposition procedure for extracting an IMF is called the *sifting process*. Figure 5.4 shows the estimation of the upper and lower envelopes and mean value of a signal to be analysed with the EMD method. These are the first steps of the sifting process which is briefly described in the following:

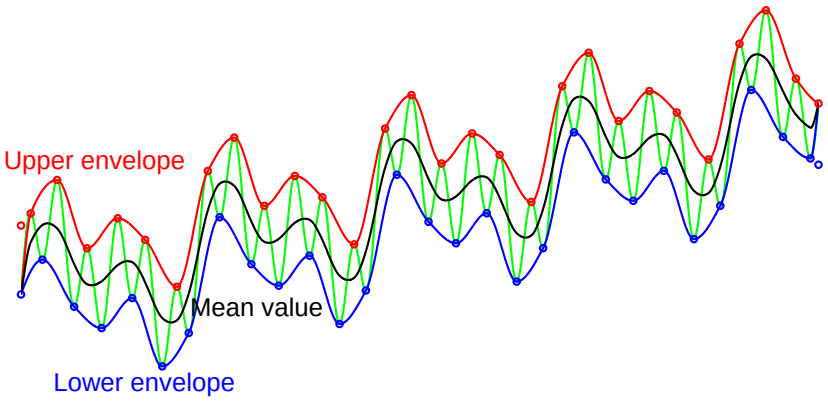


Figure 5.4: Illustration of the sifting process of the EMD method.

1. The local extrema and the local minima of the signal $x(t)$ are found.
2. All the local extrema of the signal are connected to form the upper envelope $u(t)$, and all the local minima of the envelope are connected to form the lower envelope $l(t)$. This connection is usually made using a cubic spline interpolation scheme.
3. The mean value $m_1(t)$ is defined as:

$$m_1(t) = \frac{l(t) + u(t)}{2} \tag{5.1}$$

and the first possible component $h_1(t)$ is given by the equation:

$$h_1(t) = x(t) - m_1(t) \tag{5.2}$$

The component $h_1(t)$ is accepted as the first component only if it satisfies the conditions to be an IMF. If it is not an IMF, the *sifting process* is followed until $h_1(t)$ satisfies the conditions to be an IMF. During this process $h_1(t)$ is treated as the new data set, which means that its upper and lower envelopes are formed and the mean value of these envelopes, $m_{11}(t)$, is used to calculate a new component $h_{11}(t)$ hoping that it satisfies the IMF criteria:

$$h_{11}(t) = x(t) - m_{11}(t) \tag{5.3}$$

The *sifting process* is repeated until the component $h_{1k}(t)$ is accepted as an IMF of the signal $x(t)$ and is denoted by $C_1(t)$:

$$C_1(t) = h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t) \tag{5.4}$$

4. The first IMF is subtracted from the signal $x(t)$ resulting in the residual signal:

$$r_1(t) = x(t) - C_1(t) \tag{5.5}$$

During the *sifting process* the signal $x(t)$ is decomposed into a finite number N of intrinsic mode functions and as a result N residual signals are obtained. The process ends when the last residual signal $r_N(t)$, is

obtained and is a constant or a monotonic function. The original signal $x(t)$ can be exactly reconstructed as the sum:

$$x(t) = \sum_{j=1}^{N} C_j(t) + r_N(t) \tag{5.6}$$

The EMD method is purely an empirical procedure not a mathematical transformation. The method's main advantage compared to most previous methods of data analysis is that it is an adaptive method based on and derived from the data. In deterministic situations the EMD method proves really efficient as a decomposition method. In stochastic situations involving noise the EMD method is basically a dyadic filter bank resembling those involved in wavelet decomposition [48, 127]. The difference between them is that the EMD is a signal-dependent time-variant filtering method that creates modes and residuals that can intuitively be given a "spectral" interpretation, different to a pre-determined subband filtering method, like the wavelet transform [48]. So when the method works well, it has the advantage of decomposing the signal in a smaller number of meaningful signal components, when it does not, problems known as *mode mixing* can occur [49]. *Mode mixing* is defined as any IMF consisting of oscillations of dramatically disparate scales, caused most of the times by intermittency of the driving mechanisms. So in that case different physical processes can be represented in one mode. Still, as will be shown, even with this problem the EMD method proves efficient for damage detection.

Apart from a classical time-frequency analysis approach, which can be applied when the EMD is used in combination with an *amplitude-frequency separation algorithm*, the EMD method can also be used as a simple *band pass filtering* method. *Band pass filtering* is used when one needs to isolate the frequencies within specific bands, which might be of greater interest. In a condition monitoring application, with the use of such a method, one would be able to exclude parts of the vibration signal not associated with the particular component examined, in this case the gearbox. Also for the case of multistage gearboxes, where the vibration signals are influenced by the meshing frequencies and harmonics of the different stages, the application of filter banks is useful.

The mode mixing problem of the EMD

What will follow is not a theoretical explanation of mode mixing but rather an attempt to show when this phenomenon is encountered and how it is exhibited when implementing the EMD algorithm, using some simulation examples. Since there exists no theoretical explanation of the EMD, it is quite difficult to theoretically explain the mode mixing problem as well, but one can find studies [49], that try to give a more in-depth description of the phenomenon, and propose solutions to it. As a matter of fact, the simulations that follow are similar to some of the examples given in [49], but adjusted in order to fit the analysis of the kind of signals that will be encountered later when analysing the gearbox datasets. In that paper, in an attempt to explain the mode mixing problem, data with a fundamental part as a low-frequency sinusoidal wave with unit amplitude were simulated and at the three middle crests of the low-frequency wave, high-frequency intermittent oscillations with an amplitude of 0.1 were added riding on the fundamental. It was shown there that because of the steps followed in the sifting process, the upper envelope resembled neither the upper envelope of the fundamental (which is a flat line at unity) nor the upper one of the intermittent oscillations (which is supposed to be the fundamental outside intermittent areas). Rather, the envelope is a mixture of the envelopes of the fundamental and of the intermittent signals that led to a distorted envelope mean. Consequently, the initial guess of the first IMF is the mixture of both the low-frequency fundamental and the high-frequency intermittent waves.

Here, a similar simulation will be shown. The lower frequency fundamental wave will be a sum of sinusoids, described by the equation:

$$x_{fundamental} = 0.5t + sin(\pi t) + sin(2\pi t) + sin(6\pi t) \tag{5.7}$$

while the intermittencies, will be Gaussian modulated sinusoidal pulses (*Matlab* function *gauspuls.m*). The reason for using such kinds of intermittencies in the simulation is the fact that gear tooth damage and damage in bearings most of the time appears as impulses in the acceleration signals. At first the simulation was performed without adding any noise to the signals. The first steps of the sifting process are shown in Figure 5.5, where it is obvious that what was claimed in [49] is true. When intermittencies exist in the simulation, the mean envelope produced is influenced by both the envelopes of the fundamental and of the intermittent signals, indeed.

What will be of more interest in the current study though, is the number of the IMFs produced in the simulations shown, and the frequency content of

each IMF. The *power spectral densities* (PSDs) of the signals of Figure 5.5, is given in Figure 5.6. It is apparent there that the signal contains a lower frequency fundamental, for both simulations. The existence of intermittencies creates some higher frequency components.

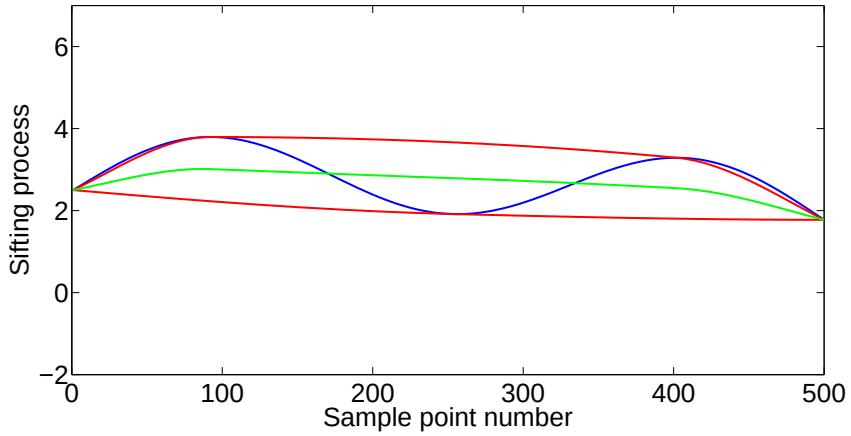
After having applied the EMD algorithm to the previously described datasets, two IMFs were produced for the first case and six IMFs for the second case. This is shown in Figure 5.7. The number of IMFs produced is influenced by the amount of runs of the algorithm needed by the sifting process in order to produce a residual signal (constant or monotonic function). Since mode mixing is the reason that, during the first steps of the sifting process, the highest and the lowest frequency parts of the signal are not separated, but carried to the next modes, one could claim that mode mixing could be an additional cause of the production of a higher number of IMFs during the decomposition, which is actually confirmed by the simulation results given in this case.

The result of mode mixing is shown in Figure 5.8, where the PSDs of the first three IMFs are given. It is obvious here that these modes have coinciding frequency bands.

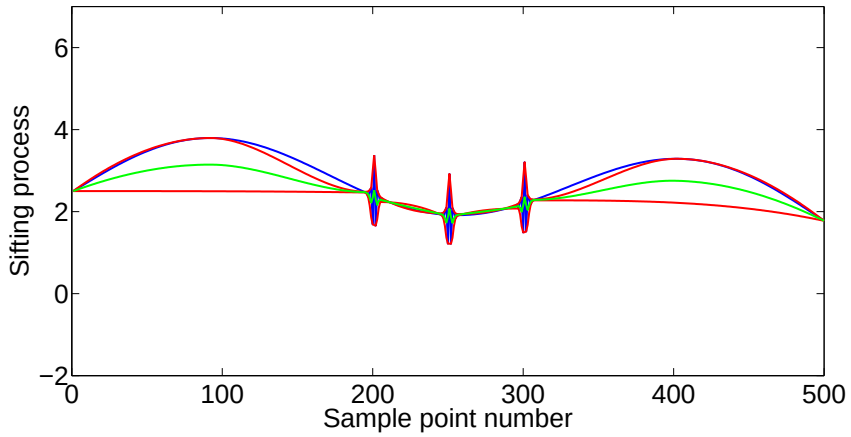
This problem seems to be improved when using a *noise assisted data analysis* approach to the EMD, the EEMD, and several studies have focused on the performance of the EMD algorithm in the presence of noise, [48, 49]. If the decomposition is insensitive to added noise of small amplitude and bears only little changes, the decomposition is generally considered *stable* and satisfies a condition of physical uniqueness; otherwise, the decomposition is *unstable* and does not satisfy physical uniqueness. From this point of view, the EMD is considered an unstable technique, since due to mode mixing any pertubation can result in a new set of IMFs as reported by [128].

For deterministic signals, or even signals that may contain a reasonable amount of noise, any additional signal components should result in additional IMFs (since the algorithm is supposed to decompose signals into meaningful signal components). If the signal does not contain any noise and has intermittencies then an even higher number of IMFs may be expected. If the signal contains a certain amount of noise, because noise seems to solve the mode mixing problem then additional components would result in additional IMFs again, but this time, mode mixing might or might not occur depending on the amount of noise the signal contains and the kind of intermittency. This will be shown in the next simulation.

When analysing Gaussian noise with the EMD, it is worth mentioning that



(a)



(b)

Figure 5.5: Illustration of the first steps of the sifting process for (a) a signal containing no intermittencies and (b) a signal containing intermittencies. The blue line shows the signal, the red lines the upper and lower envelopes and the green line the mean value of the envelopes.

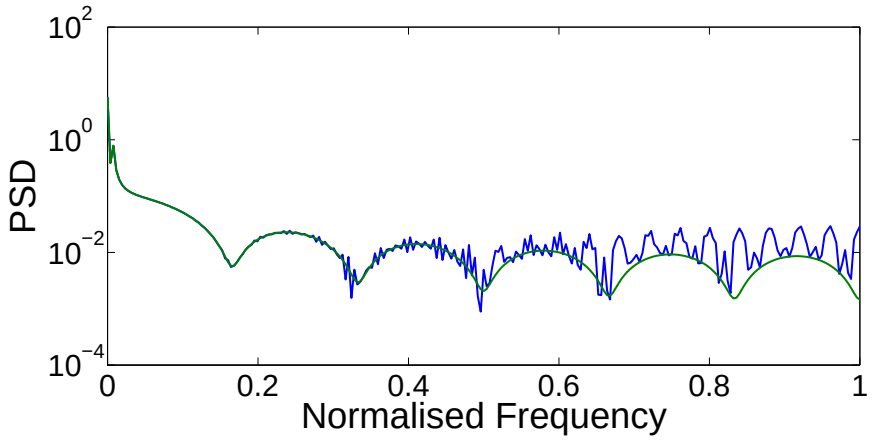


Figure 5.6: The PSD of the signals of Fig. 5.5.

as already proved in [48], the EMD method acts as a dyadic filter bank. In that case a specific number of IMFs with overlapping frequency bands will be created. It is generally believed that noise seems to fix the mode-mixing problem.

In order to have an idea of how the addition of noise would influence the previous simulation, the same datasets are used again in the next example (Figure 5.9), this time after having added Gaussian noise with a signal-to-noise ratio of 25. The number of IMFs produced in this case is five for the first dataset that does not contain the intermittencies and eight for the dataset that contains the intermittencies. The existence of noise in the datasets has generated a higher number of IMFs, but again for the signal that contained the Gaussian modulated sinusoidal pulses (intermittencies), even more IMFs were produced. So it is obvious from this simulation, that the addition of impulses in a signal that contains a little noise will add again additional IMFs when performing the EMD method.

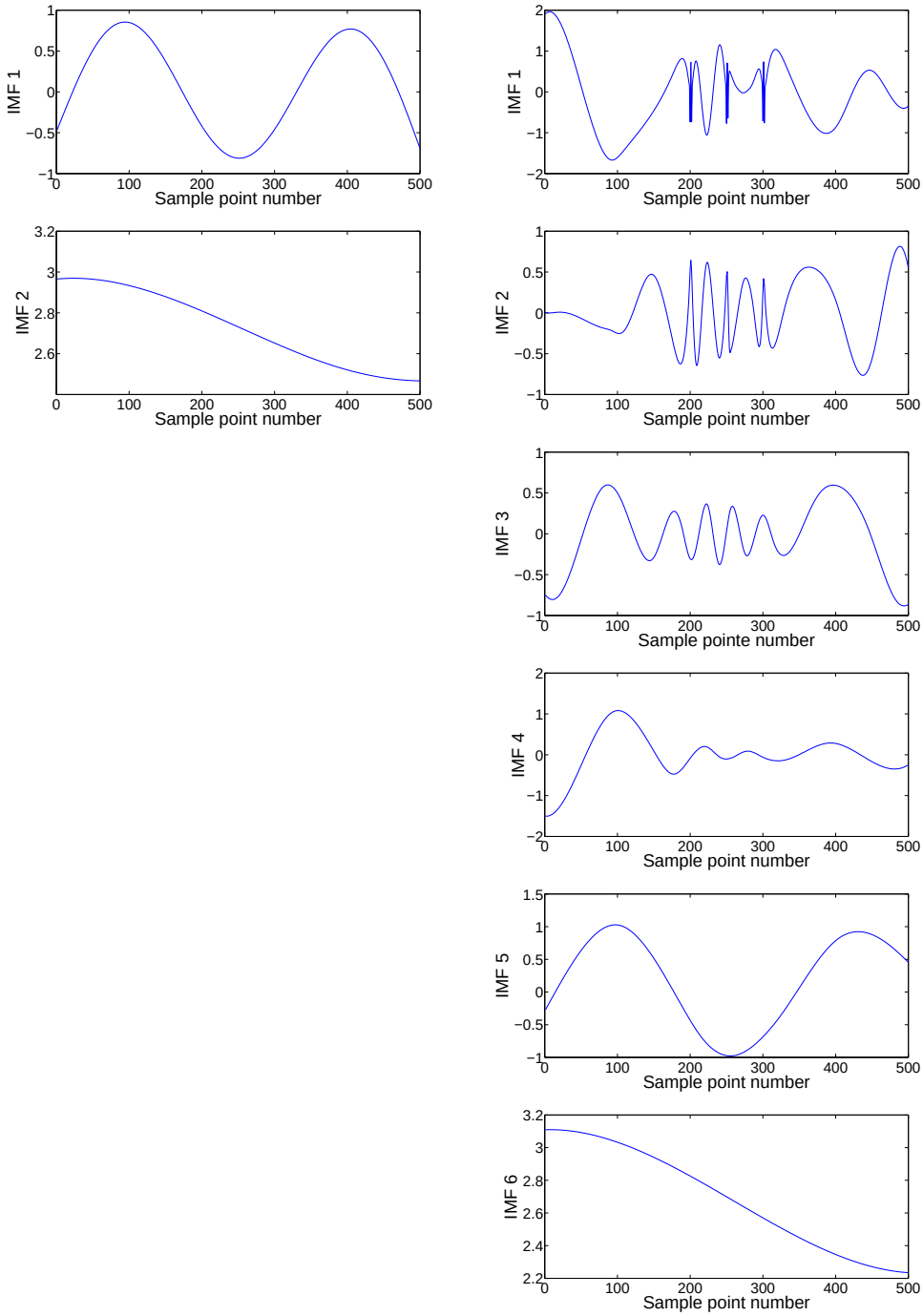


Figure 5.7: IMFs of the signals shown in Figure 5.5. More particularly: (a) IMFs of signal with no intermittencies (blue line of 5.5 (a)) and (b) IMFs of signal with intermittencies (blue line of 5.5 (b)).

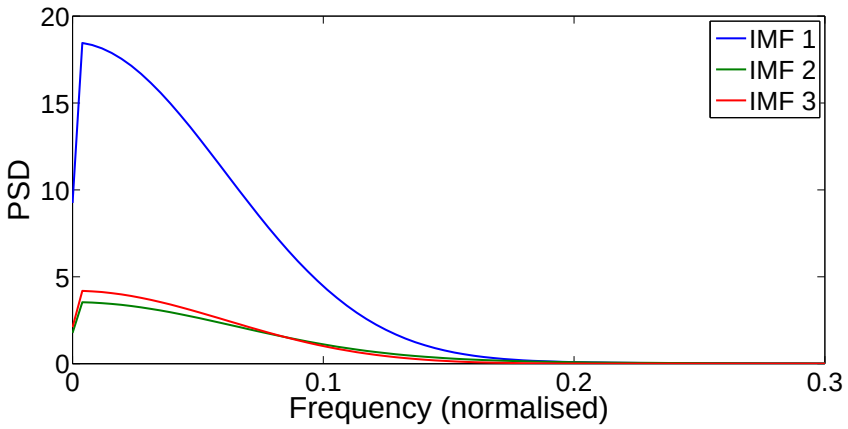


Figure 5.8: The PSD of the first three IMFs of the second signal (that contains intermittencies).

5.3 Amplitude-frequency separation algorithms

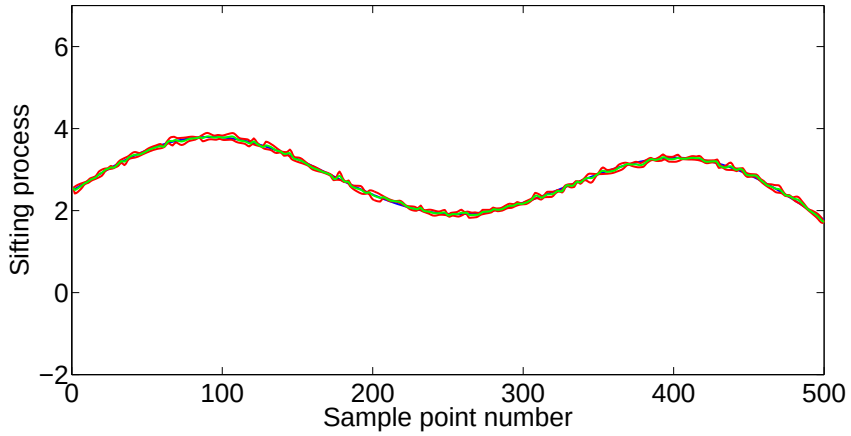
The terms *amplitude-frequency separation algorithms* or *amplitude-frequency demodulation algorithms* are generally used to describe algorithms used in signal processing to extract the IA and the IF of a signal being analysed. These methods are based on concepts different to that of the classical Fourier theory frequency concept, which is given in the Appendix A. The need of a more accurate representation of nonstationary signals led to the need for estimation of their instantaneous properties, *instantaneous phase*, *frequency* and *amplitude*. Among these algorithms, the HT is probably the most popular and based on this transform the concept of *instantaneous frequency* was established [119, 129].

5.3.1 The concept of instantaneous frequency

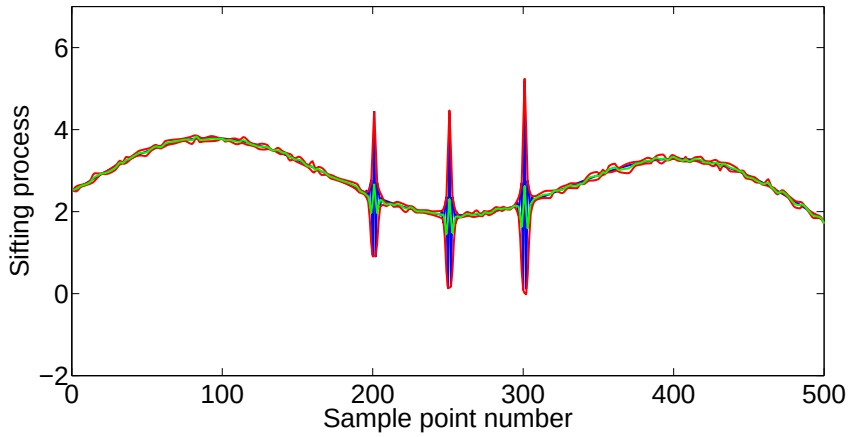
The simplest definition of frequency can be given for cyclical processes, such as oscillations. In this case frequency f, is the inverse of the period T (time taken for one cycle):

$$f = \frac{1}{T} \tag{5.8}$$

This definition assumes that the frequency is constant and that there is a whole cycle of motion.



(a)



(b)

Figure 5.9: Illustration of the first steps of the sifting process for (a) a signal containing Gaussian noise and no intermittencies and (b) a signal containing Gaussian noise and intermittencies. The blue line shows the signal, the red lines the upper and lower envelopes and the green line the mean value of the envelopes.

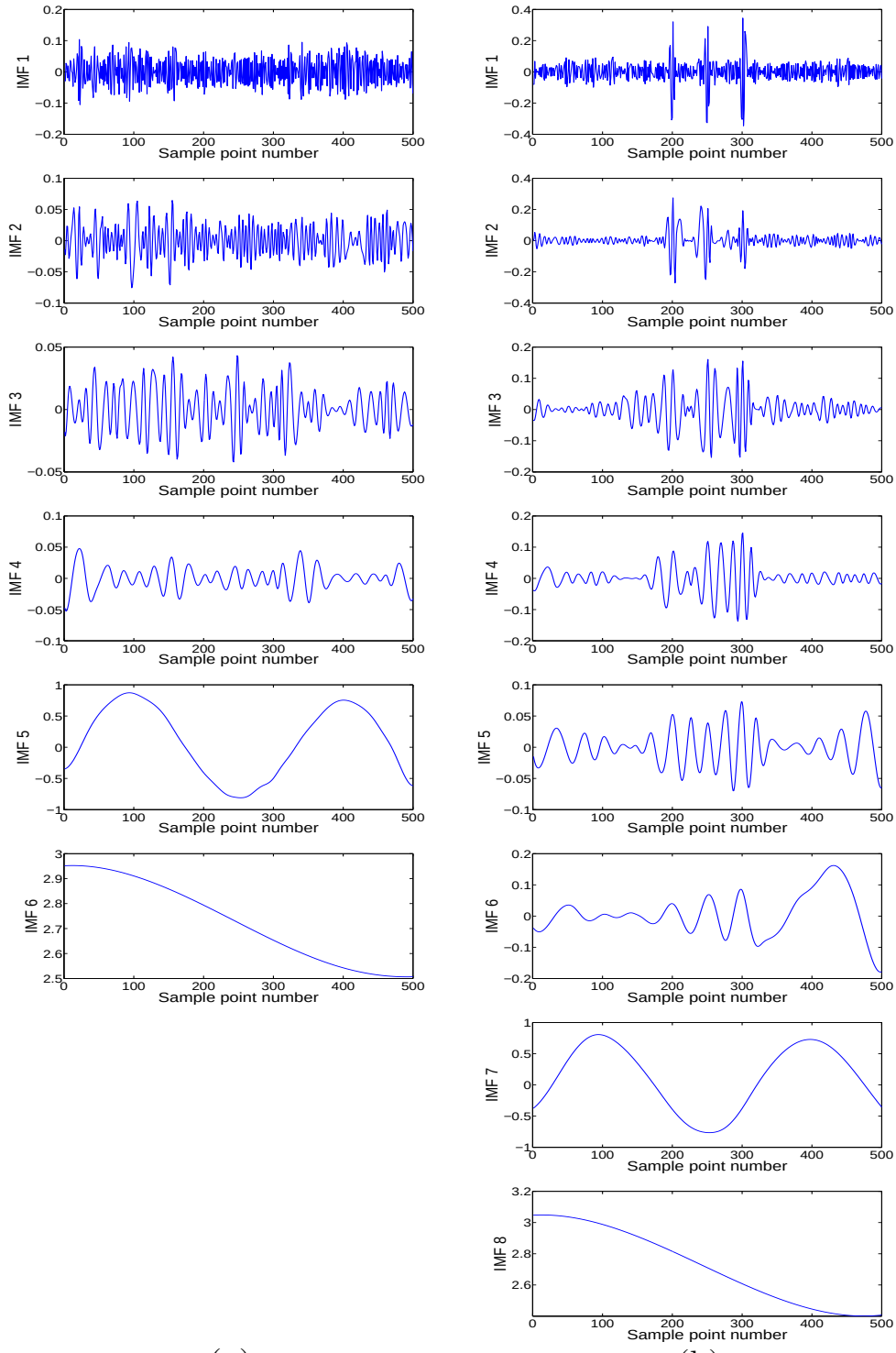


Figure 5.10: IMFs of the signals shown in Figure 5.9: (a) IMFs of signal with no intermittencies (blue line of 5.9 (a)) and (b) IMFs of signal with intermittencies (blue line of 5.9 (b)).

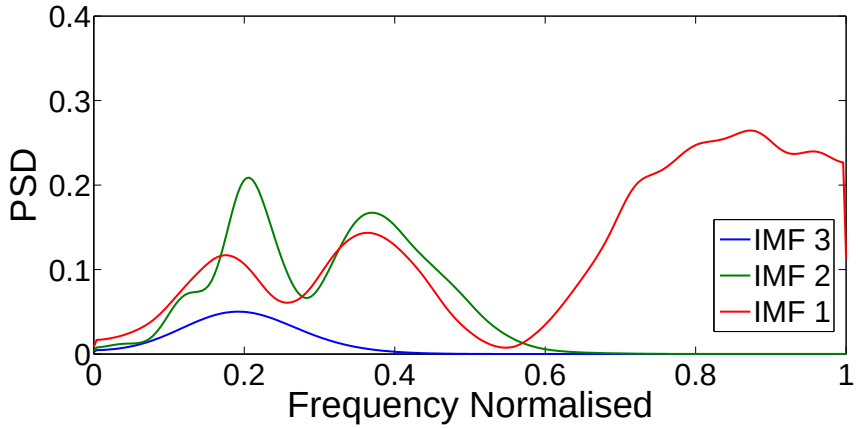


Figure 5.11: The PSD of the first three IMFs of the second signal (that contains intermittencies).

Another classic definition of frequency, is the Fourier based definition. There is an extensive section on the Fourier analysis in the Appendix A of this thesis. What is important to mention here though, is that according to this definition the several coexisting values of frequency that may exist in a signal can be estimated, but these frequency values are constant over the whole time span analysed with the Fourier analysis method. This means that the method is meaningful only for data representing linear and stationary processes [129].

It is mostly because of the classic Fourier frequency definition and the uncertainty principle associated with it, that the concepts of the IF was in the beginning quite controversial. Despite this, the need of a frequency as a function of time has been justified in both mathematical and practical terms in reference [129]. Simply stated: the need of the existence of an instantaneous frequency concept is related with the need of representing nonstationary signals, whose frequency is constantly changing, accurately.

5.3.2 The Hilbert transform

The HT is a linear integral operator [130]. It can be used to derive the *analytic representation* of the signal $x(t)$. The analytic representation of a signal facilitates the estimation of the IA and IF of a signal. If $\hat{x}(t)$ is the HT of a signal $x(t)$, it is given by the equation:

$$\hat{x}(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t-\tau} \, \mathrm{d}\tau = x(t) * \frac{1}{\pi t} \tag{5.9}$$

Given $\hat{x}(t)$ one can define the analytic signal, introduced by [131]:

$$z(t) = x(t) + i\hat{x}(t) \tag{5.10}$$

where i is the imaginary unit. The equation (5.10) written in its exponential form gives:

$$z(t) = A(t)e^{i\theta(t)} \tag{5.11}$$

where $A(t)$ is the IA of the signal, and $\theta(t)$ the instantaneous phase. The amplitude envelope (or IA) and instantaneous phase can be estimated by the following equations:

$$A(t) = \sqrt{x(t)^2 + \hat{x}(t)^2} = |z(t)| \tag{5.12}$$

$$\theta(t) = arctan(\hat{x}(t)/x(t)) = Arg(z(t)) \tag{5.13}$$

and the IF can be calculated by:

$$f(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt} \tag{5.14}$$

Combining the above with the EMD method, the original signal $x(t)$ can be expressed as:

$$x(t) = Re\{\sum_{j=1}^{N} A_j(t) e^{i \int \omega_j(t) dt}\} \tag{5.15}$$

where j is the index of the IMFs and A_j and ω_j the instantaneous amplitude and instantaneous angular frequency of the j^{th} IMF. The above equation enables one to represent the instantaneous amplitude and instantaneous frequency of the signal in a three-dimensional plot. This time-frequency representation is designated as the Hilbert spectrum.

The FT of equation (5.9) gives:

$$\hat{X}(\omega) = X(\omega)(-jsgn(\omega)) \tag{5.16}$$

Equation (5.16) shows that the HT can be estimated in a more simple way: by transforming the signal into the frequency domain, and shifting the phase

of positive frequency components by $-\pi/2$ and of negative components by $+\pi/2$ and then transforming back to the time-domain. So two important properties of the HT are:

- the HT preserves the domain in which the signal is defined,
- the HT shifts the phase of the signal by 90°.

The HT is estimated in the following analysis using the *Matlab* function *hilbert.m*. The algorithm used in order to compute the HT uses the following steps.

- The FFT, of the input sequence is calculated, storing the result in a vector x.
- A vector h (Hilbert window) is created. The elements of this vector ($h(i)$) have the values:
  - 1 for $i = 1, (\frac{n}{2}) + 1$
  - 2 for $i = 2, 3, \ldots (\frac{n}{2})$
  - 0 for $i = (\frac{n}{2}) + 2, \ldots, n$

  where n are the sample points.
- The element-wise product of x and h is calculated.
- The inverse FFT of the sequence obtained in the third step is calculated and the first n elements of the result are returned.

Estimating the IA after these steps is simple, since it is the amplitude of the complex HT. The IF is the time rate of change of the instantaneous phase angle.

5.3.3 Teager-Kaiser energy operator and energy separation algorithms

The TKEO can estimate the "energy" of a signal, and it is defined as:

$$\Psi_c[x(t)] = [\dot{x}(t)]^2 - x(t)\ddot{x}(t) \tag{5.17}$$

where $x(t)$ is the signal and $\dot{x}(t)$ and $\ddot{x}(t)$ are its first and second derivatives respectively. In the discrete case, the time derivatives of the equation (5.17) can be approximated by time differences:

$$\Psi_d[x(n)] = x(n)^2 - x(n+1)x(n-1) \tag{5.18}$$

The TKEO offers excellent time resolution because only three samples are required for the energy computation at each time instant. The operators Ψ_c and Ψ_d were developed by Teager during his work on speech production modelling [132, 133] where he described the nonlinearities of speech production and showed a plot of "the energy creating sound", without giving though the algorithm to calculate this "energy". Later, Kaiser presented the algorithm developed by Teager in his work [134, 135].

An alternative approach to that of the HT separation algorithm for the estimation of IA ($A(t)$) and IF of the signal ($f(t)$), was developed in [136] and uses the TKEO to estimate initially the required energy for generating the signal being analysed and then to separate it into its amplitude and frequency component using an *energy separation algorithm*.

The energy separation algorithm is described by the following equations:

$$f(t) = \frac{1}{2\pi}\sqrt{\frac{\Psi[\dot{x}(t)]}{\Psi[x(t)]}} \tag{5.19}$$

$$|A(t)| = \frac{\Psi[x(t)]}{\sqrt{\Psi[x(t)]}} \tag{5.20}$$

These equations estimate exactly the instantaneous frequency and amplitude envelope of a sinusoidal signal, and the approximation errors for the cases of *amplitude modulated* (AM), *frequency modulated* (FM) and AM-FM signals are small [137]. There have been developed several *discrete time energy separation* algorithms, here Desa-1 is given [136]:

$$arccos(1 - \frac{\Psi_d[y(n)] + \Psi_d[y(n)]}{4\Psi_d[x(n)]}) \approx f_i(n) \tag{5.21}$$

$$\sqrt{\frac{\Psi_d[x(n)]}{1 - (1 - \frac{\Psi_d[y(n)]+\Psi_d[y(n+1)]}{4\Psi_d[x(n)]})^2}} \approx |A(n)| \tag{5.22}$$

where $x(n)$ is the signal and $y(n) = x(n) - x(n-1)$ is its backward asymmetric difference and if T is the sampling period:

$$F_i = f_i T \tag{5.23}$$

The frequency estimation part assumes that $0 < F_i(n) < \pi$ which means that the algorithm can estimate frequencies up to $1/2$ the sampling frequency.

This is due to the fact that the sine function and its inverse have a unique correspondence between 0 and $\pi/2$. Since the sine argument is $F_i/2$, it follows that the frequency can be uniquely determined for any F_i between 0 and π. A simple simulated example in order to show the performance of both the HT and the TKEO/Desa-1 approach, is given here. The instantaneous frequency and amplitude envelope of a chirp signal (Figure 5.12), are calculated using both methods. The results (Figures 5.13 and 5.14), show that both methods have almost the same resolution with the energy separation method doing a little better.

The reason that the results of the HT show a worse end-effects problem is related to a circular convolution issue [138]. This problem could probably be solved if one chose to zero-pad or truncate the data analysed, this could also improve the resolution of the method although it would make the algorithm a little slower. Since the analysis of the experimental data did not show similar problems, as will be shown later, because the variations in the signals were not as dramatic as the variation on the simulated chirp, this problem is not of major importance in the current study.

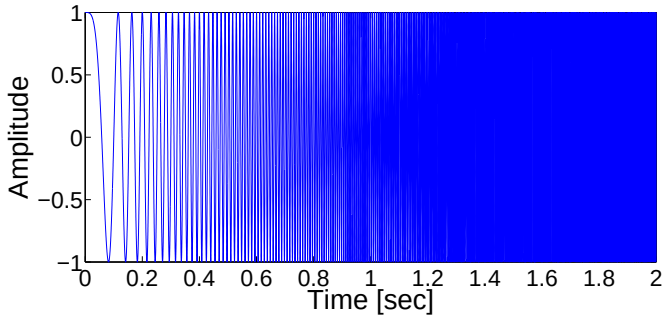
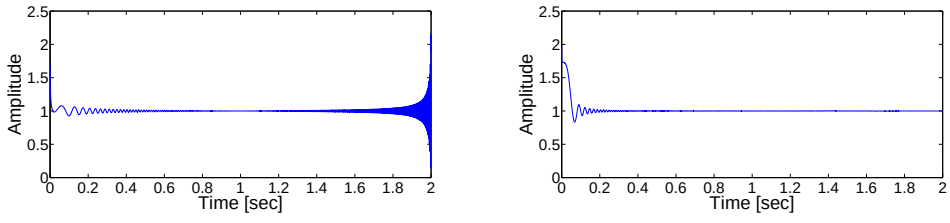


Figure 5.12: Chirp signal.



(a) (b)

Figure 5.13: The amplitude envelope of the chirp signal using: (a) the HT and (b) the TKEO and Desa-1.

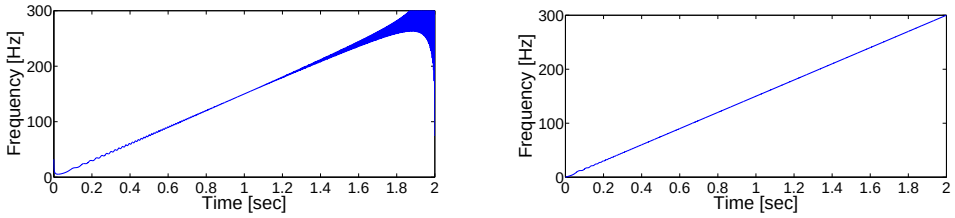


Figure 5.14: The instantaneous frequency of the chirp signal using: (a) the HT and (b) the TKEO and Desa-1.

The main problem that occurs during the estimation of the instantaneous characteristics of the signal being analysed, using both methods, is that a differentiation exists in the process. This is the reason why signals with low signal to noise ratio cannot be analysed efficiently, since differentiation amplifies noise. So applying an appropriate filter before the estimation of the instantaneous characteristics, in order to denoise the signal, is important. The TKEO can only be applied to monocomponent signals, so a filter bank method should first be applied to the signal analysed in order to extract its monocomponents and analyse them separately with the TKEO/Desa-1 approach. In this case, the EMD method is the filter bank method applied. In addition, one should pay attention to having signals that are smooth enough for both of the amplitude-frequency separation approaches, since differentiation may produce spikes at the points where the signal is not smooth. For this reason, a cubic spline interpolation in the EMD sifting process is preferred, as opposed to a linear interpolation which would not produce smooth IMFs, and also a smoothing filter should be applied to the IMFs produced before the estimation of their instantaneous characteristics.

5.3.4 Other amplitude-frequency separation methods

The *generalised zero crossing* method (GZC), is one of the most fundamental methods for computing local frequencies [129] and it has long been used to compute the mean period or frequency for narrow band signals. This approach is again only meaningful for monocomponent functions. As an approach it is quite robust and accurate.

Finally another amplitude-frequency separation method is the *direct quadrature* method (DQ), in which the quadrature of the signal is estimated using the equation:

$$sin\phi = \sqrt{1 - F^2(t)} \tag{5.24}$$

and then there are several ways in order to estimate the phase and frequency of the signal. More details on these last two methods can be found in reference [129], which presents a thorough study on the matters of instantaneous frequency.

5.4 Conclusions

In this chapter the theoretical background of the basic time-frequency analysis methods that will be later applied on the datasets of this study was presented in detail. The EMD algorithm and a major problem known as *mode mixing* were thoroughly described and some examples were also given in order to help in the understanding of the analysis that is going to follow in the next chapters.

Chapter 6

Application of the EMD to the datasets

In the previous chapter, the EMD method was described in detail, as well as the two main amplitude-frequency demodulation algorithms, the HT and the TKEO, that are going to be used in combination with the EMD method in order to perform time-frequency analysis. In addition, the author attempted to describe the kind of damage features that one should expect in such signals.

In this chapter, the experimental and simulated datasets described in Chapter 4 will be analysed using the methods presented in the previous chapter, in order to test their performance. The work presented in this chapter will focus on the application of the TKEO, which is suggested in this thesis as a recently-emerged algorithm, alternative to the time-frequency analysis using the EMD method and HT.

Another important subject that will be addressed in this chapter is the kind of influence of the varying loads in the vibration signals, in other words what frequency bands of the vibration are influenced, how the EMD method corresponds to this influence, and whether the frequency bands associated with damage are influenced by the load variations.

6.1 Simulation data results

Figure 6.2, displays the results of the EMD on the simulated vibration signals given in Figure 6.1 which correspond to three different cases. The first case is a simulation under steady load conditions and without any tooth damage in the gear model described in detail in Section 4.2. The second case is a simulation under time-varying load, produced in FAST and without any tooth damage in the gear model. Finally, the third case is the simulation under time-varying load and with an early damaged tooth. The diagrams each represent three gear revolutions. One can observe that the first case produced has the smallest number of IMF signal components. The application of time varying load in the model created effects in the signal, that were recognised by the EMD algorithm as more signal components than the previous case, that is the reason why this time one more IMF was extracted. Finally, the introduction of damage created three more IMFs than the second case. The reason that three more IMFs were produced in this case, instead of just one, is related to the *mode mixing* problem discussed in Section 5.2.1. Basically, IMFs 2, 3 and 4 contain the frequency region of the harmonics of the meshing frequency and include damage features, that should probably be found in just one IMF and would be separated from the rest of the signal components if the EMD algorithm worked perfectly. Damage is shown in these IMFs as periodic pulses with the period of the revolution of the damaged gear. So damage is an intermittency in the vibration signal.

The best IMF representing the effects of damage visually (the pulses are more apparent), is IMF 3. The instantaneous characteristics of this IMF can give probably the best features for damage detection, although the analysis of the other two IMFs would give sufficient features also. The occurrence of the EMD's *mode mixing* problem, in the way that is shown in this simulation suggests that further research might be of interest on whether the number of IMFs, for the case of gearbox signals that don't contain significant amount of noise, can serve as one more indicator of damage, from the point of view that a higher number of IMFs might be produced in this case not only due to new frequencies in the signal related to damage and also due to mode mixing.

This is an interesting observation that hasn't been found in other studies, that have treated the mode mixing as a disadvantage and have proposed improved methods of the algorithm for condition monitoring. Strictly speaking, mode mixing is indeed a problematic feature of the algorithm in signal processing terms, since it might reduce the resolution of the time-frequency analysis. Still, even under these circumstances, the EMD produces, if not better, at

least equal results with other time-frequency methods, such as wavelets. All the above insinuate that there might be a potential in using this trait of the EMD as an indicator of damage, since damage is a form of intermittency in the vibration signals that could create a significantly higher number of produced IMFs.

Another observation is that the kind of time varying load used in these simulations has an effect on the first IMF of the signals. A major part of the load variation effect is decomposed in the first IMF of the decomposition and therefore identified mainly as *noise* or a high frequency component by the EMD method. It is important to mention here, that the first IMF of each case represents a different frequency band. Basically in the second and third case examined, a highest frequency component exists, influenced by the time-varying load, and represented by the first IMF, that does not exist in the first case, that is the steady load undamaged case. The second IMF of the second case (Figure 6.2 (b)) and the fourth IMF of the third case (Figure 6.2 (c)) are the equivalent to the first IMF of the first case (Figure 6.2 (a)); they represent the same frequency band, around the harmonics of the meshing frequency. In a similar manner, the IMFs that follow for the three cases examined represent the same frequency bands of the signal. This fact is shown in Figure 6.3 (a) where the PSDs of the 1^{st} IMF of the first simulation, the 2^{nd} IMF of the second simulation and the 4^{th} IMF of the third simulation are compared and it is obvious that they fall in the same frequency bands. In the same manner the PSDs of the next two IMFs of each case are compared in Figures 6.3 (b) and (c) leading to the same conclusion. The effects of the time-varying load influence these frequency bands as well, still this will not create any problem in the damage detection part. The reason for this is mainly because the EMD algorithm manages to create an IMF representing the frequencies of damage effects in the vibration signal-in this case the IMF 3 of the third simulation-and isolates it from the rest of the signal components.

Finally, one should remember when proceeding in the analysis of the IMFs, that the first IMF represents the highest frequency component of the signal, and the ones that follow represent lower frequency components. So noise, which in this case has not been added to the signals, the harmonics of the meshing frequency of the gear model, the meshing frequency of the model and lower frequency components are represented in this particular order by the IMFs if they do exist in the signal. The same index of IMF does not necessarily represent the same time-scale of the signal for the different cases examined. In addition, generally it is the first IMFs that will have damage indicators, since they represent the highest frequency components, and are

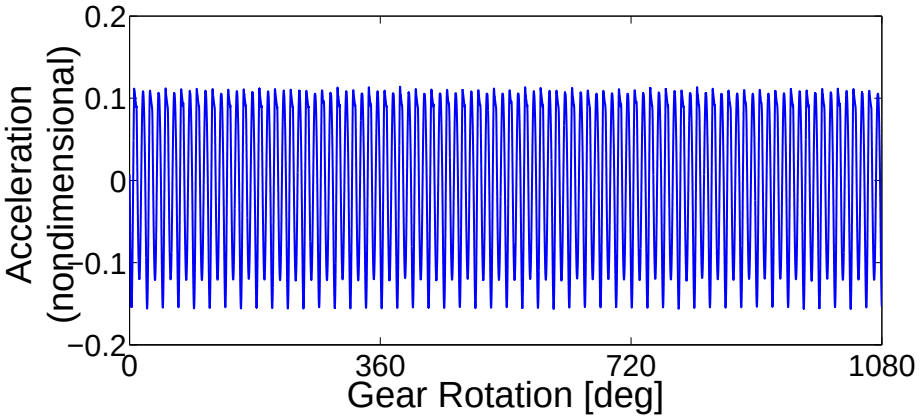
more suitable for damage identification.

The amplitude-frequency separation of the third IMF of the third (damaged) case, using both approaches, the HT and the TKEO/Desa-1 approach is presented in this section. Both methods seem to have similar resolution in the frequency (Figure 6.4) and amplitude (Figure 6.5) diagrams. What is shown here is that the frequency of the third IMF, that was identified as the most sensitive to damage, drops at the point where the damage occurs, while the amplitude increases. The increase of the IA can be explained due to the increase of the vibration levels when the damaged tooth engages. The IF, on the other hand drops because the meshing stiffness drops and the compliance increases at that time. This happens because the damaged tooth is more flexible (more compliant) due to damage. IF could be potentially used for diagnostics. The only restriction in this case could be that the estimation of frequency might be badly influenced by noise, but there exist simple solutions for this problem (filtering, which is what the EMD method does, and smoothing).

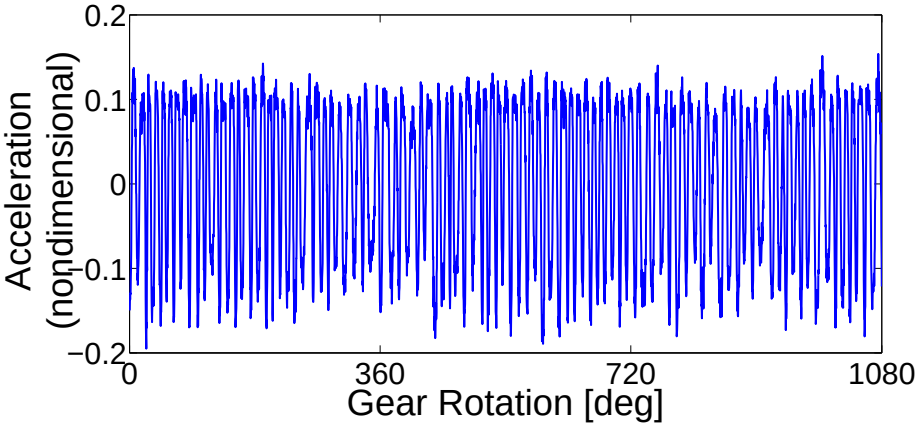
A measure of the power of the third IMF was estimated as well (Figure 6.6), according to the equation:

$$P = \frac{1}{2} {A_i}^2(t) \tag{6.1}$$

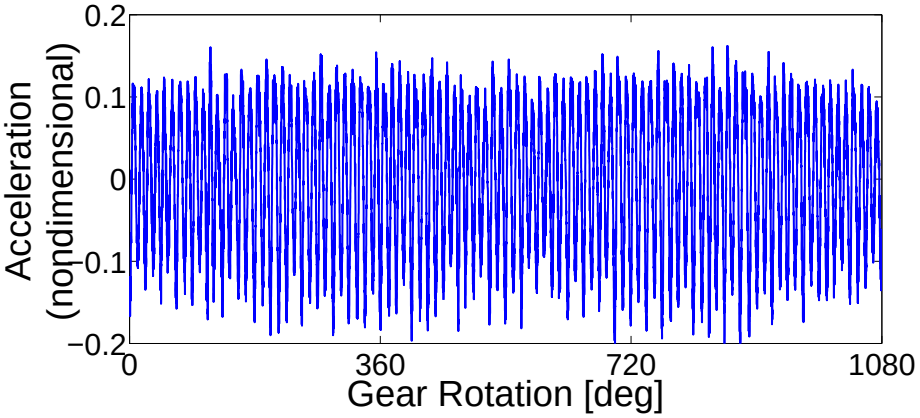
where A_i is the instantaneous amplitude of the i^{th} IMF. This measure has also been used in reference [59] where it was referred to as *instantaneous energy*. It should be noted here that the power is estimated using the same equation for the TKEO/Desa-1 approach and has nothing to do with the Teager "energy" that can be also estimated according to equation (5.17). In this way it is acceptable to make comparisons between the results that both methods produce.



(a)

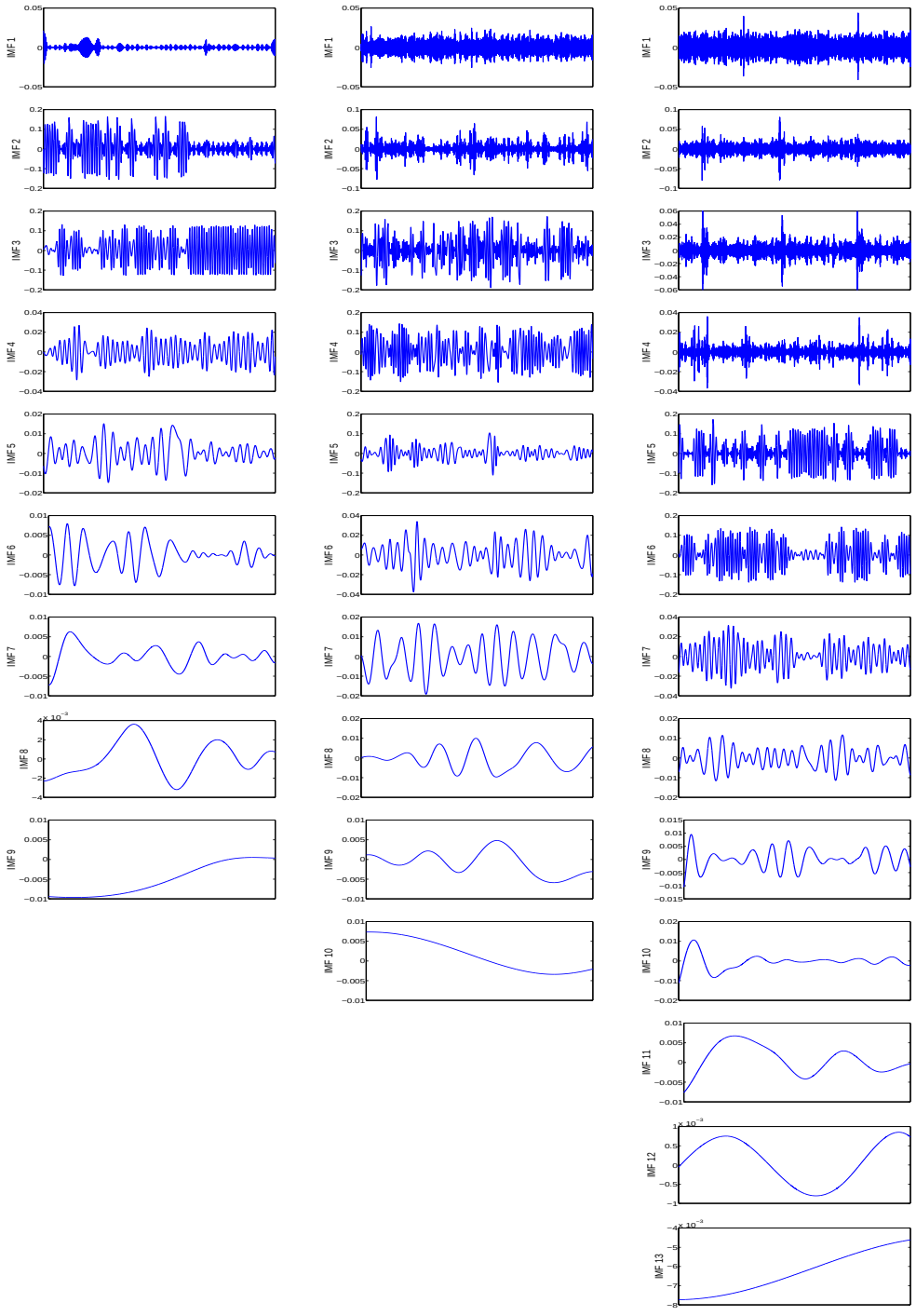


(b)



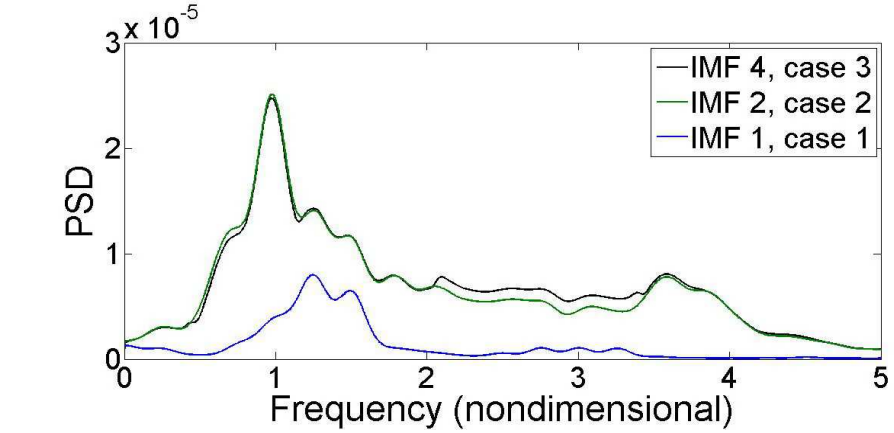
(c)

Figure 6.1: Acceleration diagrams of the simulations: (a) undamaged case - steady load, (b) undamaged case - time-varying load, and (c) early damage in one tooth - time-varying load.

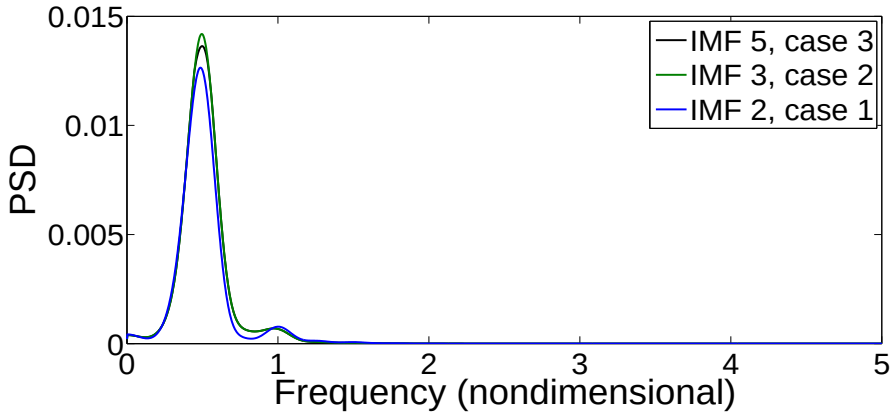


(a) (b) (c)

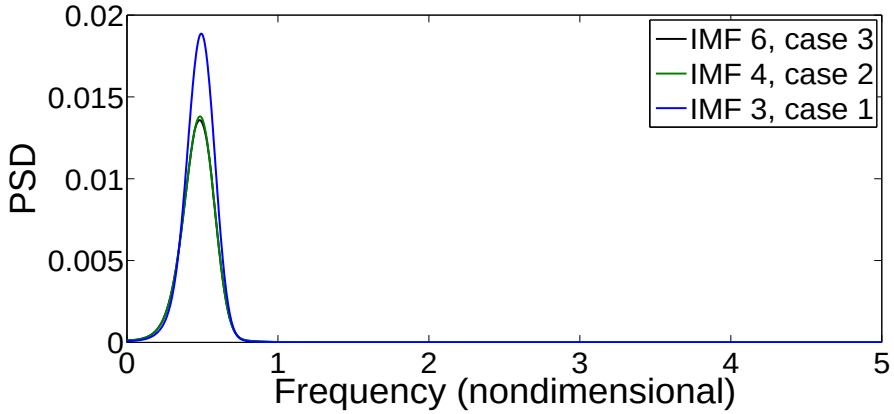
Figure 6.2: Intrinsic mode functions of simulations: (a) undamaged case - steady load, (b) undamaged case - time-varying load, and (c) early damage in one tooth - time-varying load.



(a)

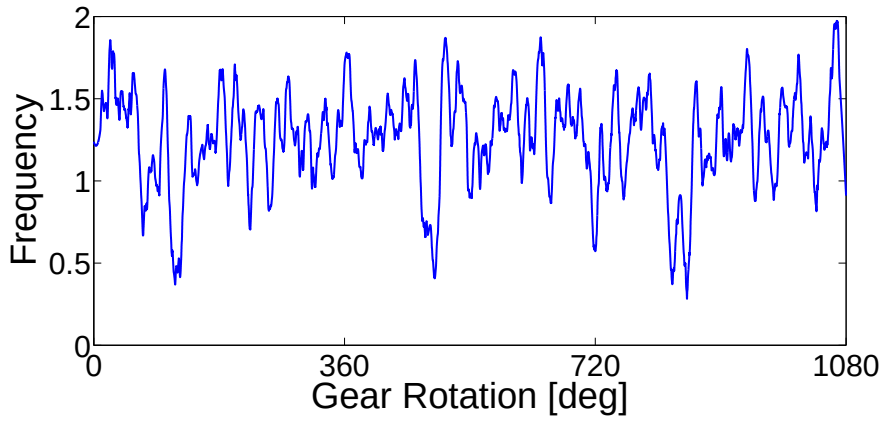


(b)

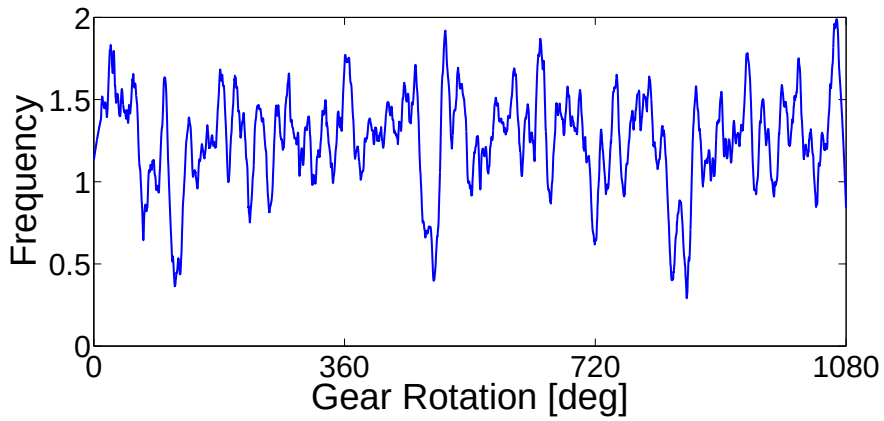


(c)

Figure 6.3: PSD using the Welch method of the: (a) 1^{st} IMF of the 1^{st} case, 2^{nd} IMF of the 2^{nd} case, and 4^{th} IMF of the 3^{rd} case, (b) 2^{nd} IMF of the 1^{st} case, 3^{rd} IMF of the 2^{nd} case, and 5^{th} IMF of the 3^{rd} case, (c) 3^{rd} IMF of the 1^{st} case, 4^{th} IMF of the 2^{nd} case, and 6^{th} IMF of the 3^{rd} case,

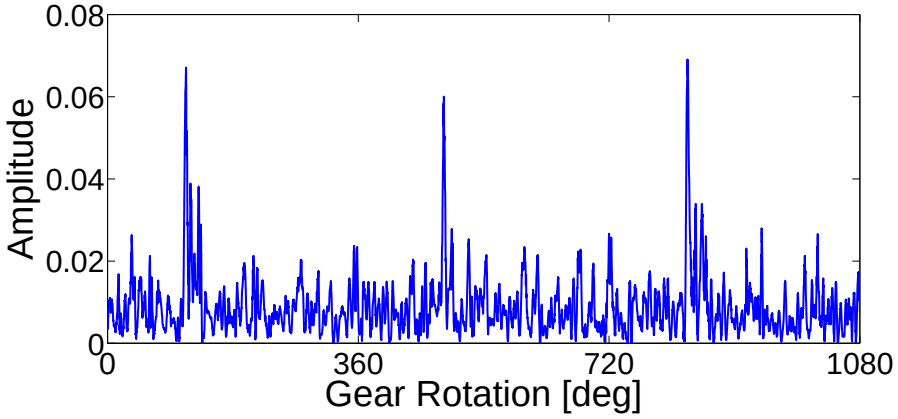


(a)

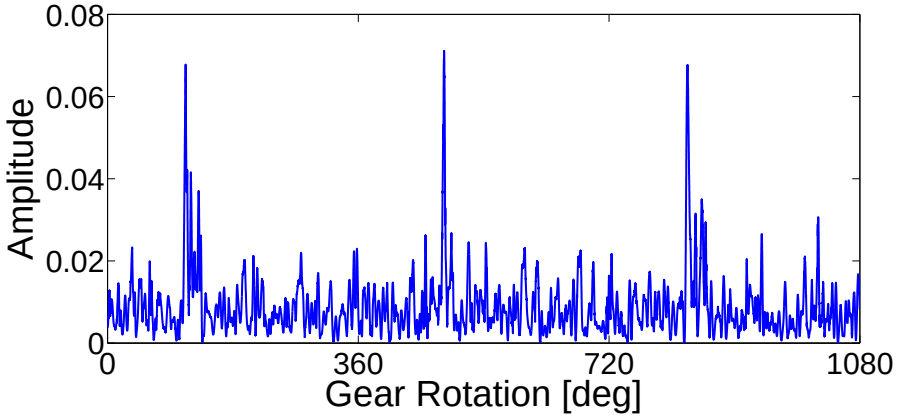


(b)

Figure 6.4: Instantaneous frequency of the 3rd IMF of the third simulated case: (a) Hilbert Transform (b) TKEO and Desa-1.

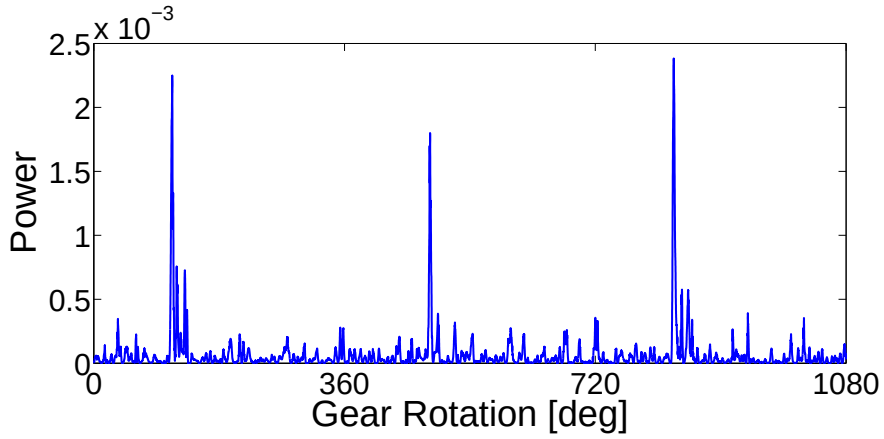


(a)

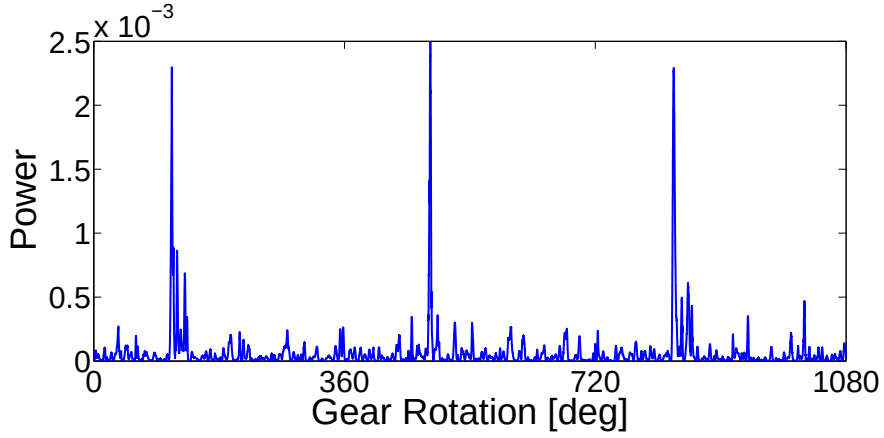


(b)

Figure 6.5: Instantaneous amplitude of the 3rd IMF of the third simulated case: (a) Hilbert Transform (b) TKEO and Desa-1.



(a)



(b)

Figure 6.6: Power diagram of the 3rd IMF of the third simulated case: (a) Hilbert Transform (b) TKEO and Desa-1.

6.2 Experimental data results

6.2.1 Bearing datasets

The reason that the bearing datasets are going to be decomposed using the EMD method in this section is mostly to show that the particular experimental results do agree with the idea that an increased number of the IMFs can show some indication of damage. These datasets will be much more appropriate to show this, in comparison with the gearbox datasets, whose results will be discussed later, due to the fact that the experiment was conducted in a lab environment and the exact condition of the bearing was known during the experiment. There will be no further analysis for damage detection on these measurements simply because the previously used methods performed quite well. Time-frequency analysis using the EMD method will be demonstrated in the next subsection (6.2.2), applied on the gearbox datasets.

Figure 6.7 shows the IMFs produced with application of the EMD algorithm on the bearing datasets. Eleven IMFs were produced for the first dataset (26/11/2012: undamaged case) and thirteen IMFs were produced for the second dataset (28/11/2012: damaged case). This result agrees with what was described in the simulations. In order to see the frequency content of each IMF the PSD diagrams of all the IMFs for each case are given in Figure 6.8. One can observe that while the IMFs produced in the undamaged case constitute an approximately dyadic filter bank, the frequency bands of the first and second IMFs in the second case (damaged bearing dataset) are overlapping on a large part of their frequency range, a fact that implies that there must have occurred some sort of mode mixing.

In Section 5.2.1, with the help of a simulation there was made an attempt to show what mode mixing is and when it is observed when using the EMD algorithm. The gearbox simulation results given in Section 6.1 of this chapter confirmed what was described. So generally, it was shown through simulations, but also through this specific experimental case that intermittencies in the analysed signals can increase the number of IMFs. So for faults that may appear as pulses or spikes in the vibration signals it is possible, if there is not a large amount of noise, then a higher number of IMFs will be decomposed due to the mode mixing problem. Generally, gear tooth cracks and bearing faults are this kind of fault.

An additional explanation to the assumption that an increased number of IMFs might indicate the existence of damage will be given in the following,

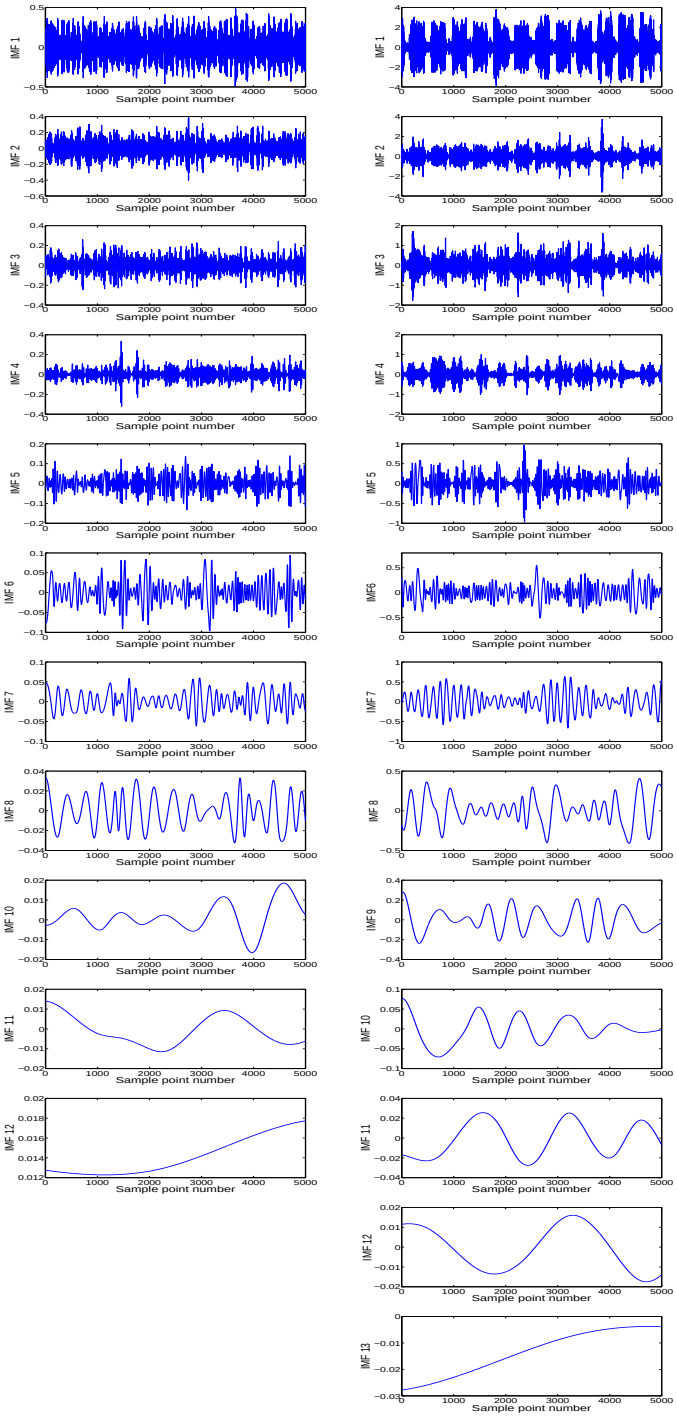
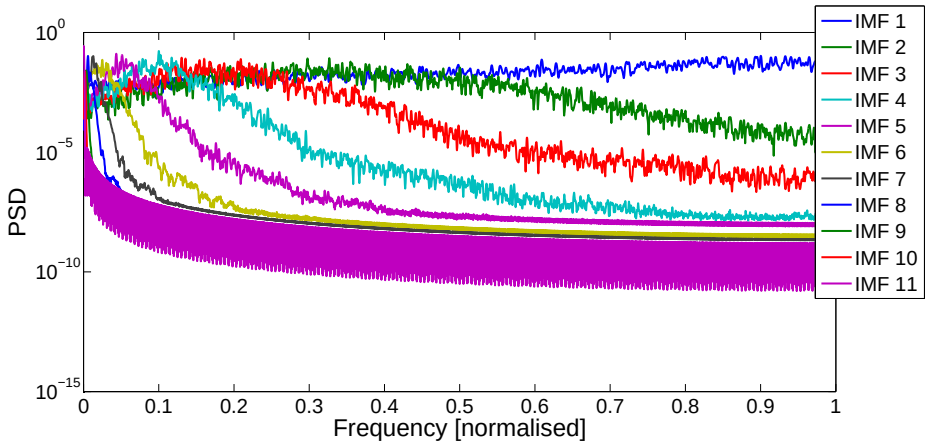
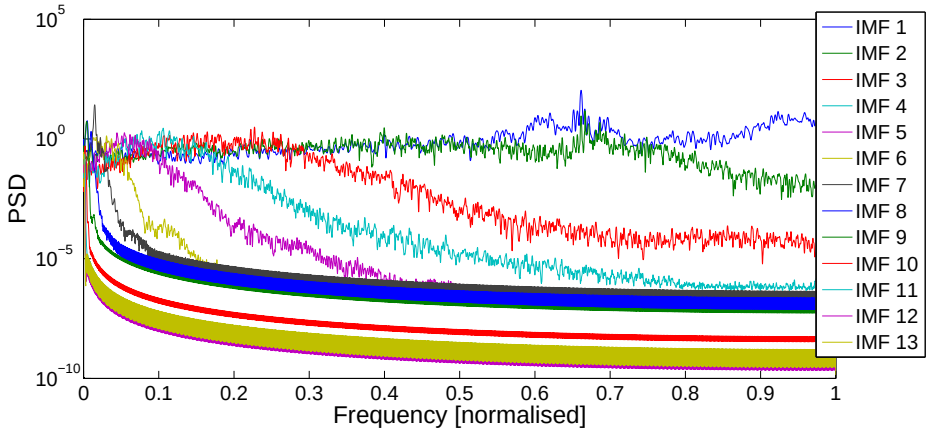


Figure 6.7: The IMFs produced from the analysis of the bearing data: (a) 26/11/2012, (b) 28/11/2012.



(a)



(b)

Figure 6.8: PSD of the decomposed IMFs for each dataset: (a) undamaged case (dataset 26/11/2012), (b) damaged case (dataset 28/11/2012).

and is not necessarily related to mode mixing. In reference [8] where the general principles of SHM are described, the last axiom (Axiom VIII) is the proposed principle that damage increases the complexity of the structure, signal or feature. Quantifiable measures of complexity can be defined by the application of statistics and signal processing, and generally in signal processing terms among such measures is the generation of new frequency components or harmonics in the power spectra of damaged signals. In this case, it is obvious from the theory related to what kind of spectrum one should expect from a damaged bearing, described in Section 4.1.2 (Figure 4.7) and also from the primary analysis that was performed in the same section (Figures 4.8 and 4.9), that damage indeed produced components that initially didn't exist in the healthy signal. From this point of view, and in the absence of a large amount of noise, the decomposition of the damaged vibration signal should be expected to generate additional IMFs, that would describe-contain in the frequency domain-these additional frequencies that exist in the vibration.

6.2.2 Gearbox datasets

The three different datasets that were obtained from the wind turbine gearbox data were also decomposed using the EMD method and the results are shown in Figure 6.9. The first dataset was decomposed into 13 IMFs, the second into 12 IMFs and the third into 13 IMFs. The diagrams show 4 rotations of the smaller wheel of the parallel gear stage 2 (damaged gear). Small speed fluctuations might exist so the rotating period might not be exactly constant for the whole duration of the datasets.

Only the first four IMFs are presented here. The frequency content of each IMF is shown in Figure 6.10, that shows the PSD of all the IMFs for each dataset. The first IMF is the highest frequency band of the signal (noise). The second IMF contains the fourth and part of the third harmonic of the meshing frequency (frequency band: 2500-5000 Hz). The third IMF contains mostly the second and the third harmonic of the meshing frequency (frequency band: 500-3000 Hz). The fourth IMF contains mostly the first harmonic of the meshing frequency (frequency band: 200-1500 Hz). The fifth IMF contains mostly the meshing frequency (frequency band: 150-600 Hz). The rest of the IMFs (not given here) are related to the meshing frequencies and harmonics of the other stages of the gearbox and to lower frequency components of the signals.

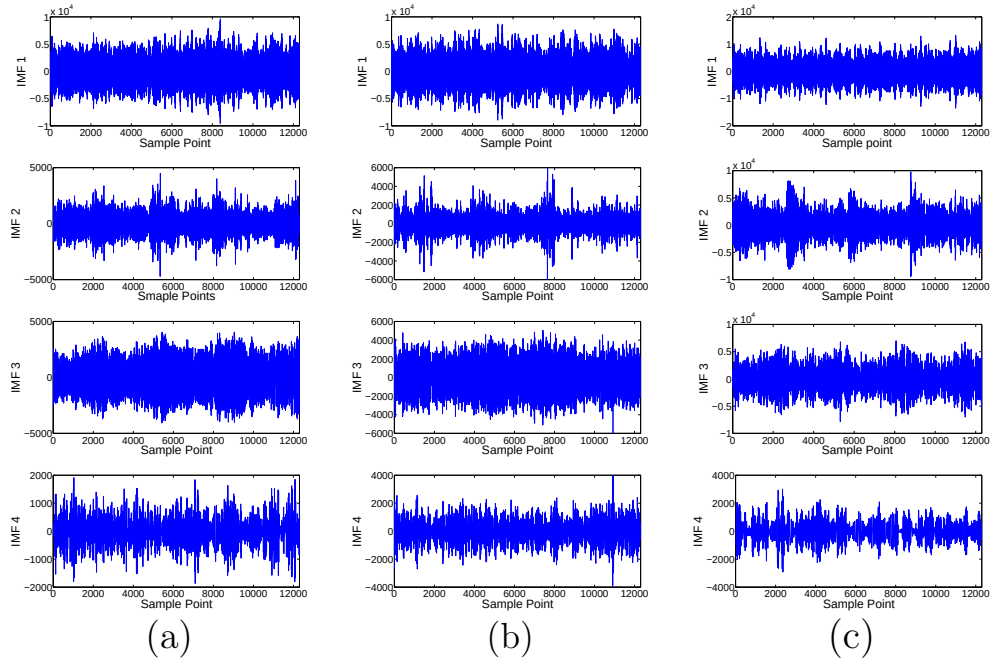
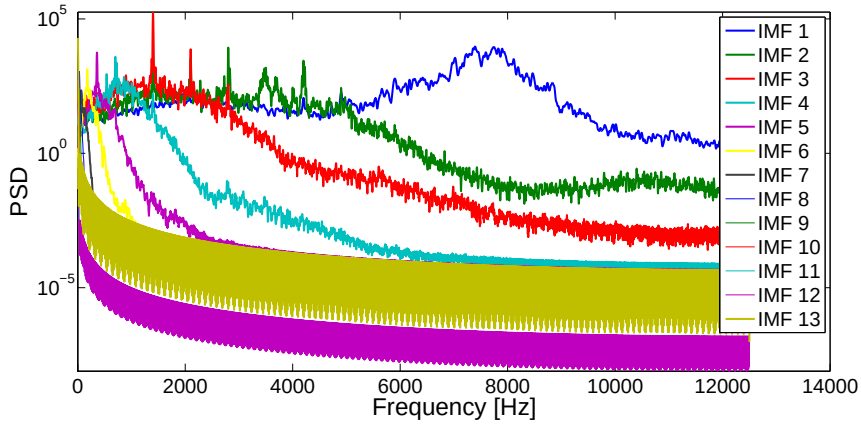


Figure 6.9: First four IMFs of the wind turbine gearbox data: (a) 31/10/2009, (b) 11/2/2010, and (c) 4/4/2010.

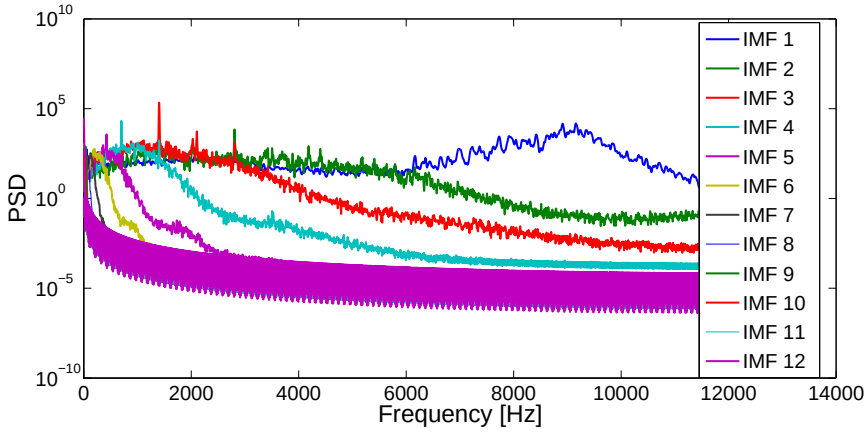
The tooth fault is most clearly shown in the second IMF as (almost) periodic pulses occurring at the time when the damaged gear (small wheel of the third gear stage) engages. The period of these pulses coincides with the rotating period of the damaged gear.

Periodic pulses can be seen, but less clearly, in the third IMF. Without proceeding to any further analysis, there is clearly some indication of damage in the first dataset which was initially considered to be describing the gearbox at an undamaged condition. This conclusion is drawn at this point by the fact that periodic pulses were produced in the second and third IMF of the first case as well. The EMD method produced 13 IMFs, which is the same number as the one produced in the third case, where damage was progressed, and one more in number than the number of IMFs of the second case.

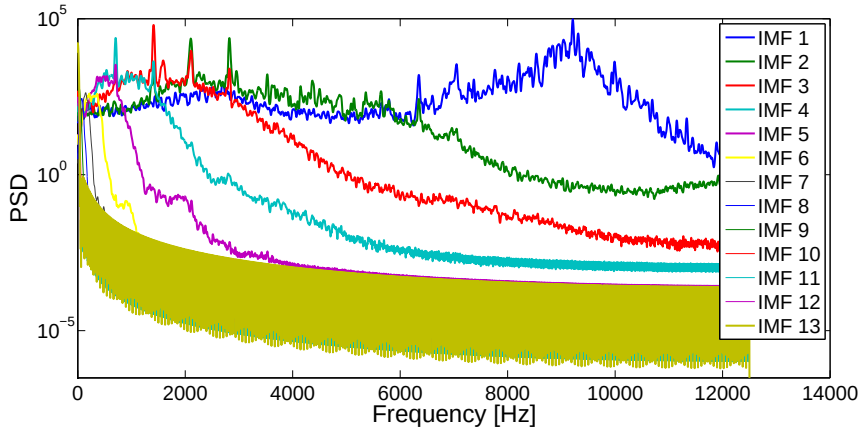
The authors were informed after having performed the analysis that the first dataset, described as a reference dataset, was actually taken when the gearbox had already shown some first indications of damage. The authors were also informed that it is actually not easily feasible to obtain datasets of a gearbox during both its healthy condition and damaged condition. Collecting all vibration data of a healthy gearbox with the expectation that it might fail in the near, or not so near future, would mean that huge data storage is



(a)



(b)



(c)

Figure 6.10: PSD of the decomposed IMFs for each dataset: (a) 31/10/2009, (b) 11/2/2010, and (c) 4/4/2010.

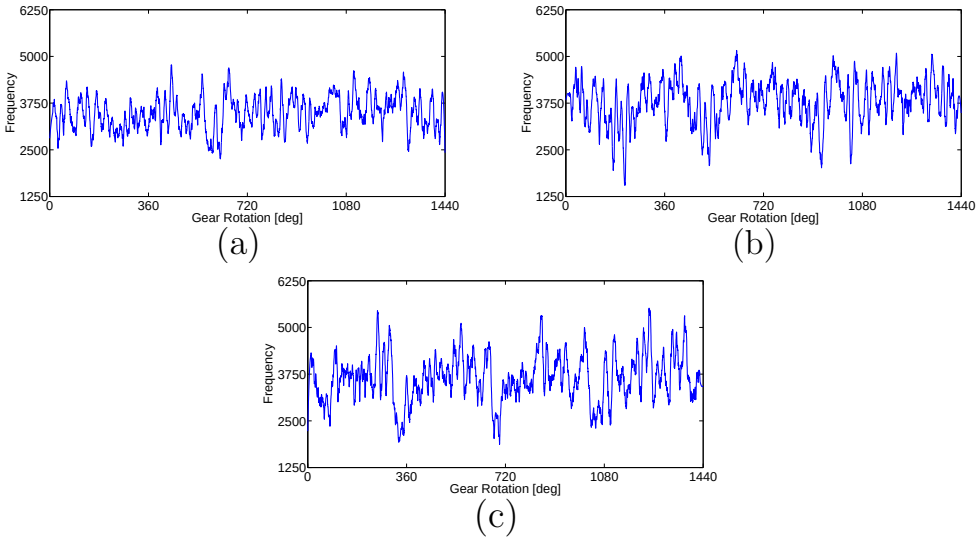


Figure 6.11: Instantaneous frequency diagrams (HT) of the 2nd IMF: (a) 31/10/2009 (b) 11/02/2010 and (c) 4/04/2010.

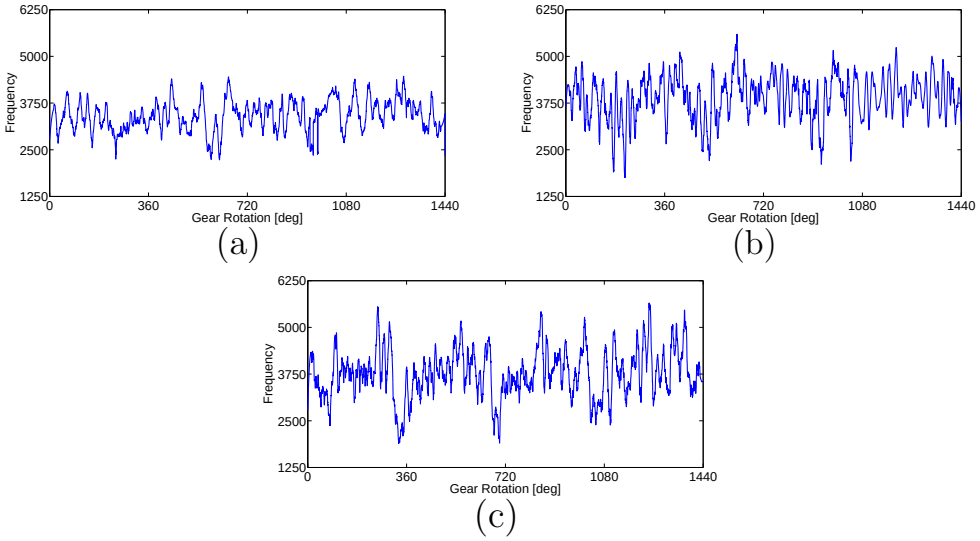


Figure 6.12: Instantaneous frequency diagrams (TKEO/ Desa-1) of the 2nd IMF: (a) 31/10/2009 (b) 11/02/2010 and (c) 4/04/2010.

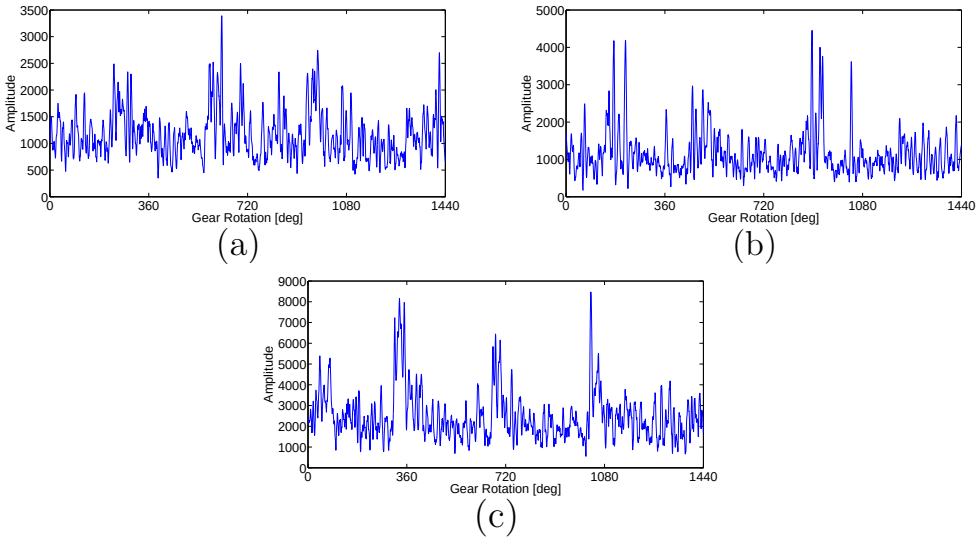


Figure 6.13: Instantaneous amplitude diagrams (HT) of the 2nd IMF: (a) 31/10/2009 (b) 11/02/2010 and (c) 4/04/2010.

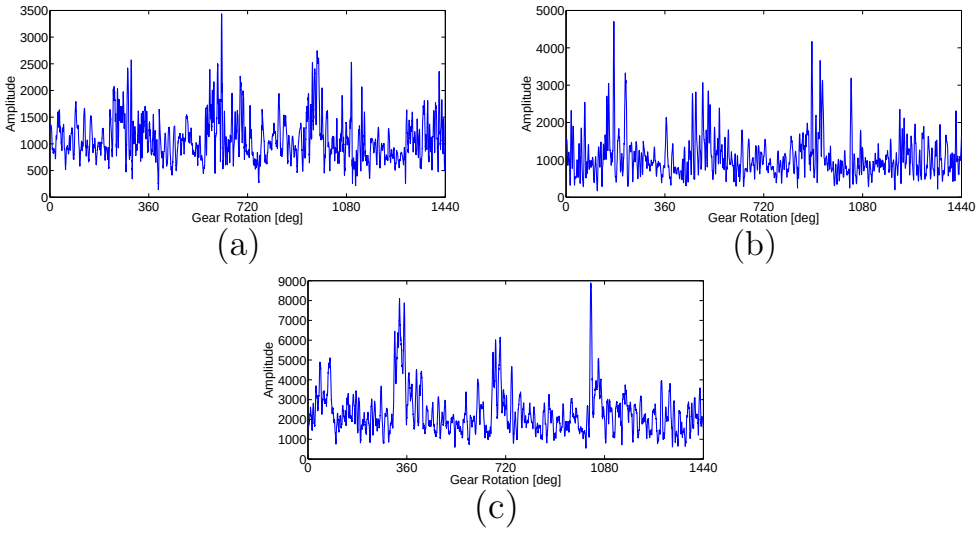
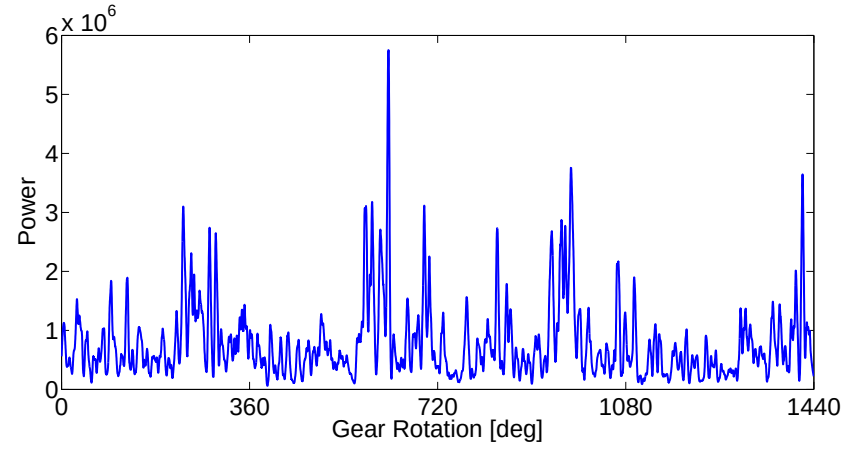
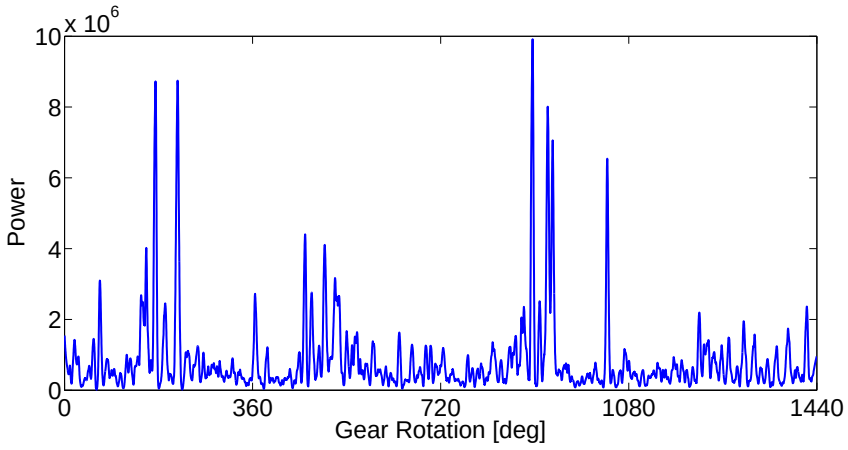


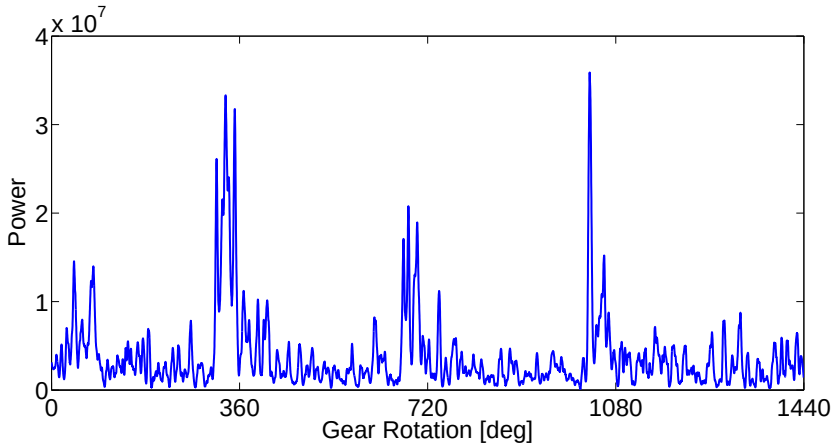
Figure 6.14: Instantaneous amplitude diagrams (TKEO/ Desa-1) of the 2nd IMF: (a) 31/10/2009 (b) 11/02/2010 and (c) 4/04/2010.



(a)

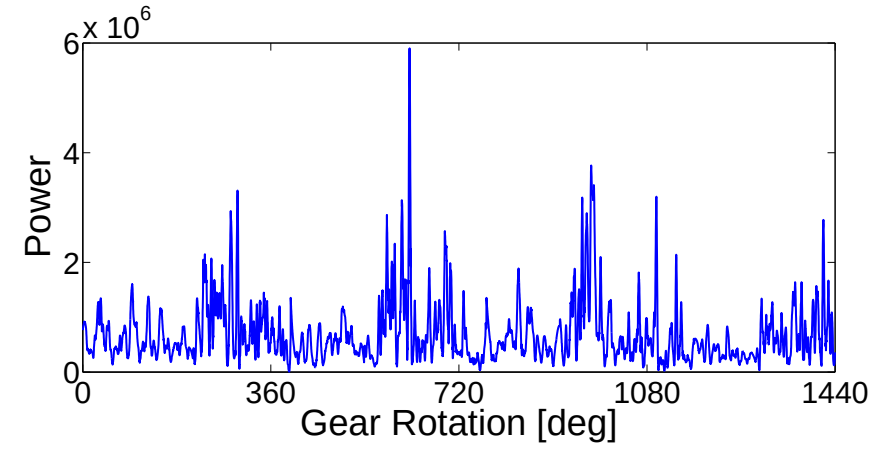


(b)

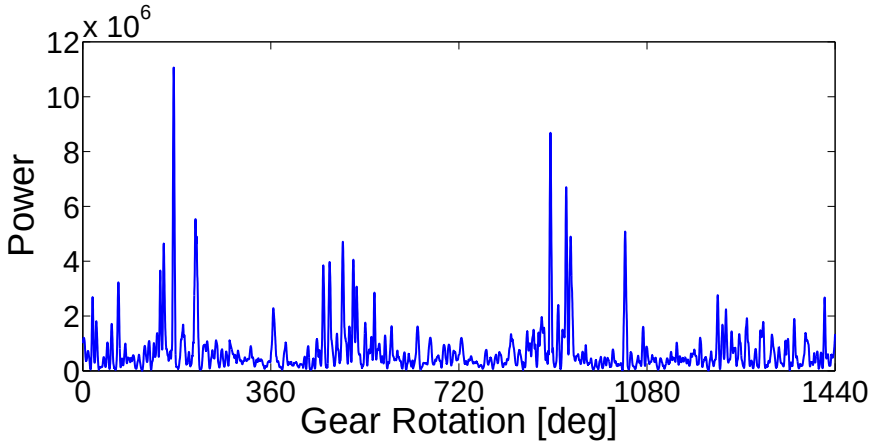


(c)

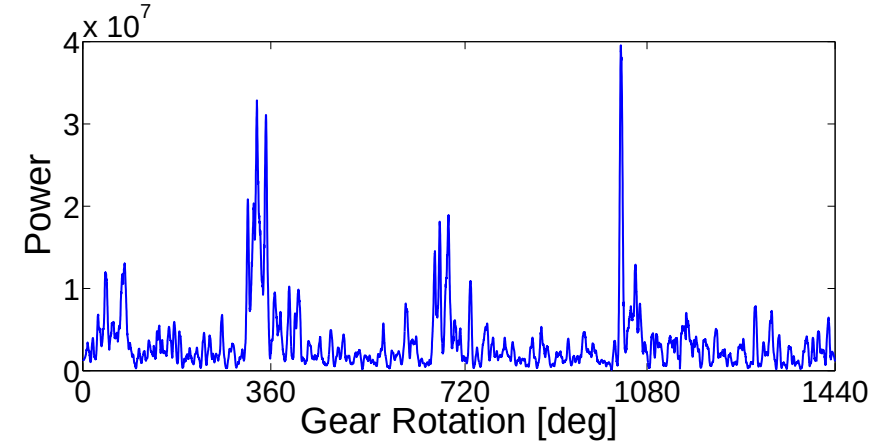
Figure 6.15: Power (HT) of the 2nd IMF: (a) 31/10/2009 (b) 11/02/2010 and (c) 4/04/2010.



(a)



(b)



(c)

Figure 6.16: Power (TKEO/ Desa-1) of the 2nd IMF: (a) 31/10/2009 (b) 11/02/2010 and (c) 4/04/2010.

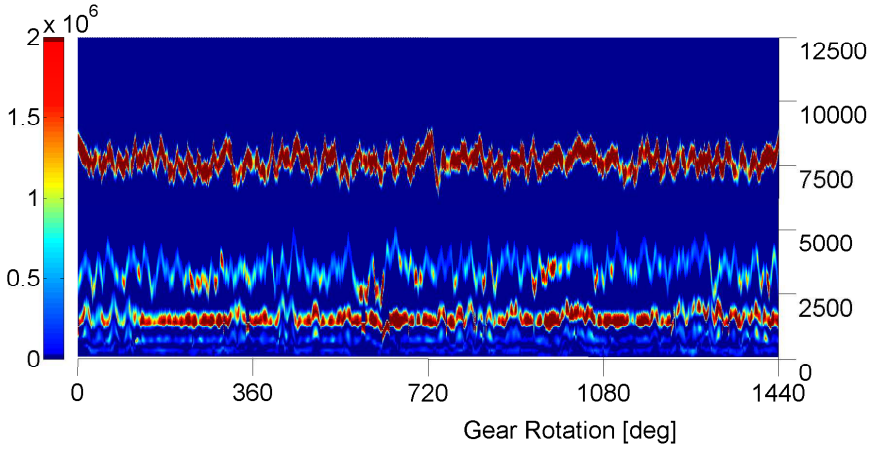
necessary.

A final remark concerning the influence of the varying load of the wind turbine gearbox will be given here. Despite the fact that the case examined here is concerning the third gear stage, which should not be as affected as the first gear stage, by the varying load caused by wind turbulences, the analysis of the experimental data shows that even this stage of the gearbox is influenced, underlying the importance of further examining the case of condition monitoring under varying load conditions. In this case a large part of the varying load influence on the vibration signals is carried in the first IMF, just as in the simulations, in addition to other types of noise that may exist in the wind turbine nacelle. Load variation due to wind turbulence influences mostly the highest frequency bands of the signal. Load variation may also be observed in other IMFs related to other frequency bands. The fact that the second case-dataset 11/2/2010-produced a lower number of IMFs than the other two, despite the fact that it was identified already by the condition monitoring systems of the wind turbine as describing a damaged gearbox, and therefore should have at least the same IMFs number as the first case, is probably explained by the different load conditions that may have existed at the specific measurements.

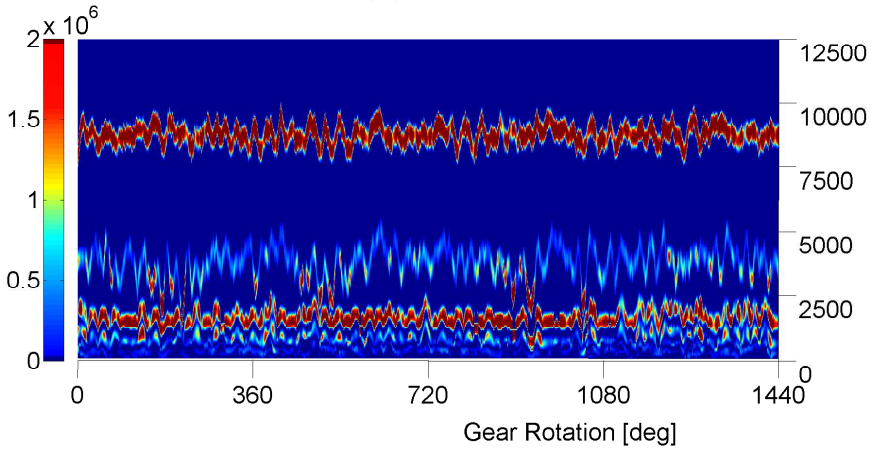
In fact, knowing now that the gearbox is damaged in all the datasets and since noise levels are high, means that this data cannot be used to support or contradict the hypothesis that damage could be reflected in the number of IMFs. Unlike the simulations, the experimental datasets all appear to have damage and have almost the same level of environmental loading. On the other hand, the data do not contradict the hypothesis. Further research is needed.

In order to obtain the amplitude envelope and the instantaneous frequency of the IMFs presented in the previous section, the IMFs were first filtered with a smoothing filter (*sgolay* smoothing filter in *Matlab*). Damage has again the same pattern as in the simulations, so basically the frequency of the second IMF, which occupies the frequency band 2500-5000 Hz (roughly) (Figures 6.17 and 6.18), drops each time that the damaged tooth engages. This is shown better in Figures 6.11 and 6.12 where the instantaneous frequency diagrams are plotted. This IMF represents the frequency band of the fourth harmonic of the *parallel gear stage II* (2820 Hz as mentioned earlier).

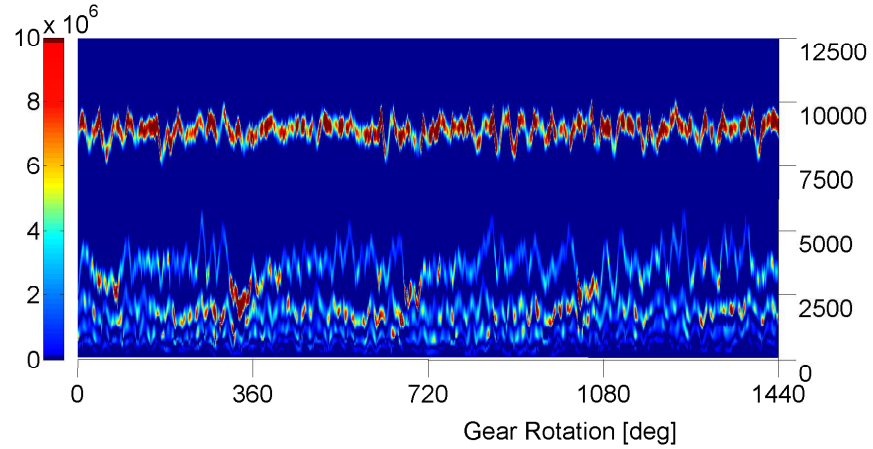
The power levels increase at that specific time as well, something that is more clearly shown in the amplitude and power diagrams of Figures 6.13, 6.14, 6.15 and 6.16. As damage increases, the power levels increase and so the scales in the 3D diagrams increase. In addition, it is clear in this



(a)

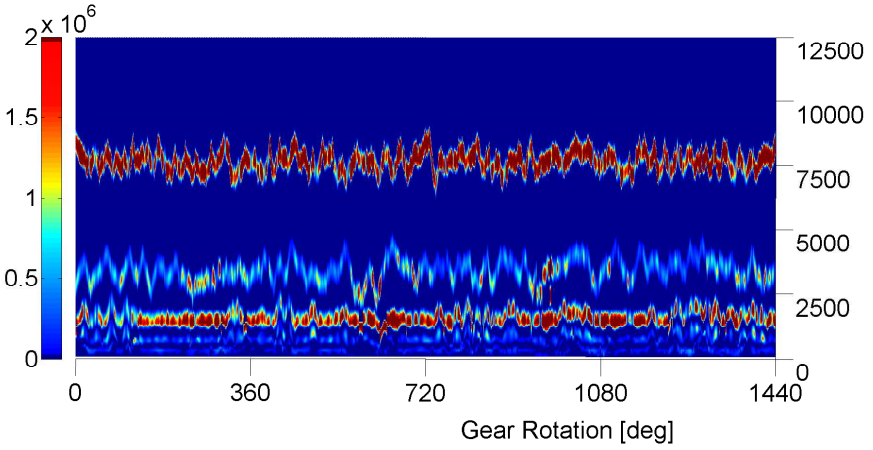


(b)

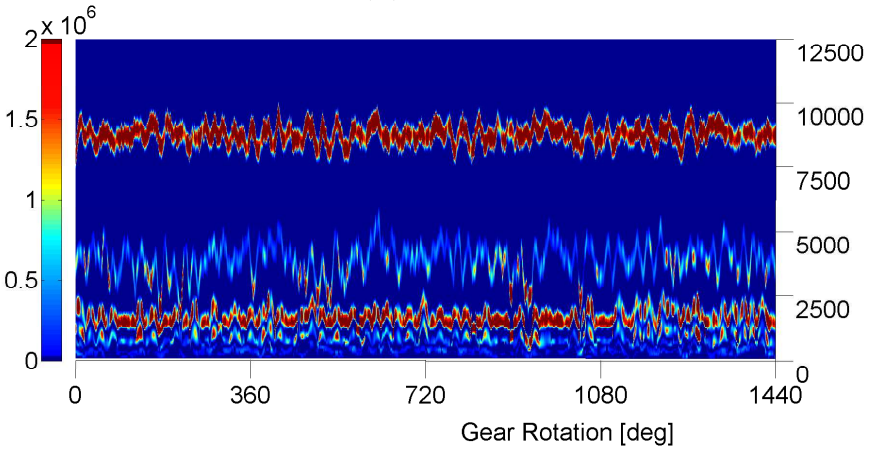


(c)

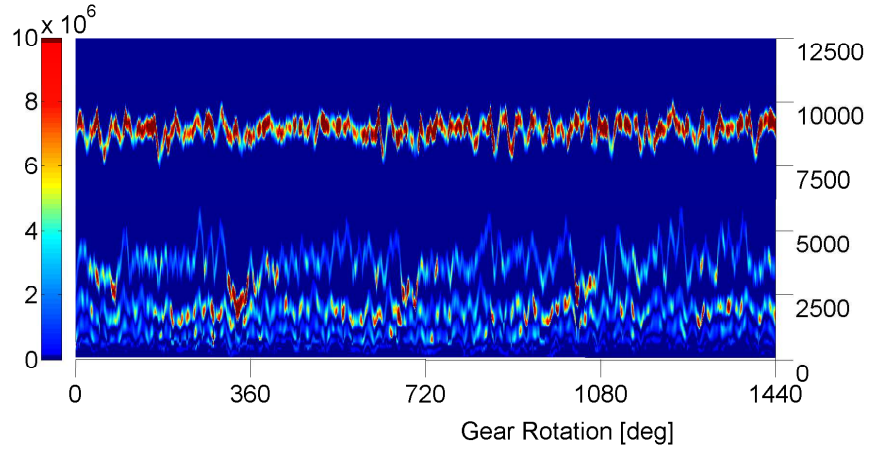
Figure 6.17: The Hilbert Spectra of the three datasets: (a) shows the spectra of 31/10/2009, (b) shows the spectra of 11/02/2010, and (c) of 4/04/2010. The horizontal axis shows the sample point and the vertical axis the instantaneous frequency.



(a)



(b)



(c)

Figure 6.18: The Teager Spectra of the three datasets: (a) shows the spectra of 31/10/2009, (b) shows the spectra of 11/02/2010, and (c) of 4/04/2010. The horizontal axis shows the sample point and the vertical axis the instantaneous frequency.

part of the analysis that there are some indications of damage in the first dataset too, and since it is at a relatively much lower power scale it must be at an early state. Finally Figures 6.11-6.13 show more clearly, that both of the amplitude-frequency separation techniques used perform equally well, at the same resolution quality, which means that TKEO/Desa-1 approach could be a good alternative to the HT approach, given its low computational requirements.

With the calculation of the instantaneous characteristics (instantaneous frequency and power) of the IMFs one can have the 3D diagrams of the signals, shown in Figures 6.17 and 6.18. Figures 6.17 (a), (b) and (c) are obtained with the HT approach and Figures 6.18 (a), (b) and (c) are obtain using the TKEO/Desa-1 algorithms. The diagrams show the power levels for each IMF and for almost four gear rotations. An interpolation was used in order to create clearer images. For significant frequency overlaps of the IMFs, there might be problems producing this diagram, and therefore only Figures 6.9-6.16 might be produced, although this issue did not occur in this particular case.

6.3 Conclusions

This chapter focuses mainly on two things:

- The first one is to test a recently emerged amplitude-frequency separation technique in order to perform time-frequency analysis-based condition monitoring with the use of the EMD method.
- The second one, is to examine the effect of time-varying load conditions present in wind turbine gearboxes when trying to perform condition monitoring of such systems.

Concerning the first aim of the work, it was shown that the TKEO, in combination with the Desa-1 energy separation algorithm, can be a good alternative approach to the HT, since it offers, under the condition of analysing monocomponent, smooth and clean signals, at least the same quality results as the HT. Concerning the second aim of the study, both the results of a simulated nonlinear gear model that included varying load conditions similar to those produced in wind turbines, as well as the results of real wind turbine gearbox data, confirmed that varying loads can create difficulties in condition monitoring.

The EMD method managed to separate the major part of the time-varying load influences from the vibration signals, leaving signal components that could be related to damage, available for further analysis either with the HT or the TKEO/Desa-1 techniques. In addition, the simulation part of the study suggested that the EMD method is able to separate the steady and time-varying load cases and the undamaged and damaged cases purely from the number of IMFs. This fact should not come as a surprise, it is well known that in the cases of both gearboxes and bearings, damage might create new frequency components, which would manifest as new IMFs if they are individually narrowband or monocomponent. The time varying loads modulate the signals and potentially broaden the signal components thus creating more IMFs by initiating mode mixing. This was not demonstrated on the gearbox experimental data, but the analysis of the bearing datasets with the EMD was encouraging.

Concerning the gearbox datasets, the fact that no additional IMFs were extracted for the two experimental vibration datasets describing the damaged cases, when compared with the first dataset (which was initially considered as describing the gearbox in an undamaged state), is explained by the fact that the load conditions that the gearbox was operating were similar for all the cases, and also the analysis presented here suggests that damage was in fact present even in the initial data set, which was later confirmed.

The analysis showed that the damage manifests as an intermittent effect in the response data and it is known that intermittency is one of the conditions that creates mode mixing. Because the intermittency due to damage appears to be present in even the first experimental data set, there is no increase in the number of IMFs as the damage progresses. Whether one is considering intermittency or new frequency components due to damage it appears that the vibration signals that contain damage information can create a (sometimes) significantly higher number of produced IMFs, something that has not been observed in other studies, or rather, not exploited, until now. Further research in order to validate these claims on experimental data would be useful.

Chapter 7

Feature discrimination using novelty detection

In the introduction it was mentioned that the feature discrimination level of damage detection can be achieved by two different approaches, the *supervised* and the *unsupervised learning* approaches.

Unsupervised learning methods for damage detection do not require all the class labels of the acquired data. This is an advantage over the *supervised learning* methods, since it is the most common case that data describing all damage classes of a structure, cannot be provided. From a machine learning perspective therefore, establishing a two-class classifier, which can use the data from a normal condition as baseline data and distinguish any data corresponding to a damage state, solves the problem of data unavailability.

The *novelty detection* or *anomaly detection* methods belong to the unsupervised learning approaches, and usually in statistics they are referred to as *outlier detection* methods. The idea of these methods is that one only needs measurements from the undamaged structure which can then be used to construct a statistical (or other) model of the data. Any subsequent data from the system can be tested for conformity in some strict sense with the model of the undamaged state [8]; nonconformity can then be said to infer damage.

In this chapter, the *outlier analysis* method will be applied to specific IMFs of the decomposed-using the EMD method-gearbox experimental datasets. This technique however, will be applied in a slightly different way than previous studies where features were obtained from e.g. the transmissibility functions [139, 140]. This different form of application is chosen in this case in

order to maintain the time-frequency analysis perspective. In addition, a second thresholding method will be proposed for this part of the analysis, known as *3D-phase space thresholding*, as novel approach. In the following sections, a brief theory of the methods used in this chapter will be given, for the better understanding of the proposed application of the methods.

7.1 Some basic theory on outlier analysis

Novelty detection can be achieved by assuming that selected features defining a normal condition follow a particular form of probability distribution. This distribution is usually the *Gaussian distribution*, since it is the simplest one. In this case, the probability density function of the data, in the normal condition, can be described by a small number of parameters.

For univariate data, these parameters are the mean and variance/standard deviation. The *outlier analysis* (OA) method computes *discordancy measures* for data and compares them with a threshold. A *discordant outlier* in a data set is an observation that appears to be inconsistent with the rest of the data and therefore has a large discordancy measure, exceeding the calculated threshold. The discordancy measure in this case is the deviation statistic:

$$z = \frac{x_\zeta - \overline{x}}{\sigma_x} \tag{7.1}$$

where x_ζ is the candidate outlier and $\overline{x}$ and σ_x the mean and standard deviation of the training (undamaged) data sample respectively [8].

In the case of multivariate data, damage detection is more difficult than the univariate situation. The discordancy measure, equivalent to equation (7.1), is the *Mahalanobis squared-distance*:

$$D_\zeta^2 = (\{x\}_\zeta - \{\overline{x}\})^T [\Sigma]^{-1} (\{x\}_\zeta - \{\overline{x}\}) \tag{7.2}$$

where $\{x\}_\zeta$ is the feature vector corresponding to the candidate outlier, $\{\overline{x}\}$ is the sample mean of the normal condition features and $[\Sigma]$ is the normal condition feature sample covariance matrix.

The threshold value that labels the inlier and outlier observations can be estimated through the employment of a *Monte Carlo* method. Briefly, a $p \times n$ matrix (p-observations, n-dimensions, in order that the matrix dimensions

match the dimensions of the extracted features) is generated and populated with elements randomly drawn from a zero-mean, unit standard deviation Gaussian distribution. Then the Mahalanobis squared distance is calculated for all elements and the largest value stored. This is repeated a large number of times, each time storing the largest Mahalanobis squared-distance, which are then sorted in order of magnitude. The threshold is assigned as a percentage of the resulting array of Mahalanobis squared-distances. A 99% confidence threshold for example, means that any values above that threshold have less than 1% probability of occurrence. The threshold is a essentially an extreme value measure.

It is important to mention here, that the application of novelty detection in SHM, and consequently CM, is complicated because of the influence of the changing environmental and operational conditions on the vibration signals. The limitation here is that the damage sensitive feature of the monitored vibration signal should remain, within some limits, stationary, in the normal condition. This means that the occurrence of damage is inferred by any significant change occurring in the feature [141]. If responses driven by, for example, a temperature variation, are not defined within the normal condition, then a novelty detection process will wrongly assign these responses as anomalies.

7.2 Experimental datasets results for Mahalanobis squared distance

OA was performed in the power of the 2^{nd} IMF for each gearbox dataset analysed in Chapter 6. This was the mode identified as the most sensitive to damage. In this part, the estimated power using the TKEO/Desa-1 method (Figure 6.16) will be used for outlier analysis and these results will be displayed here only, since they are very similar with those obtained with the HT method.

Concerning the novel approach chosen here for outlier analysis and the way features were selected, in the power diagrams a 10-dimensional feature was defined as a 10-point time window. One series of 200 features was defined as reference and used for the training data. The training data were chosen carefully in order that no peaks (of the power measure) would be included in them. In this way whenever a fault appears, the outlier statistics diagram should show a peak that is distinct from the normal condition data.

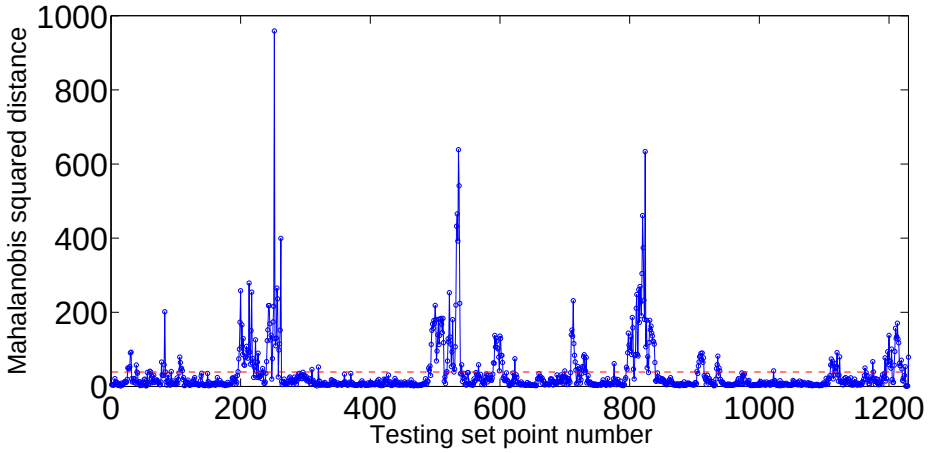


Figure 7.1: Outlier statistics for the power of the 2^{nd} IMF for the 31/10/2009 dataset (training data 280-480).

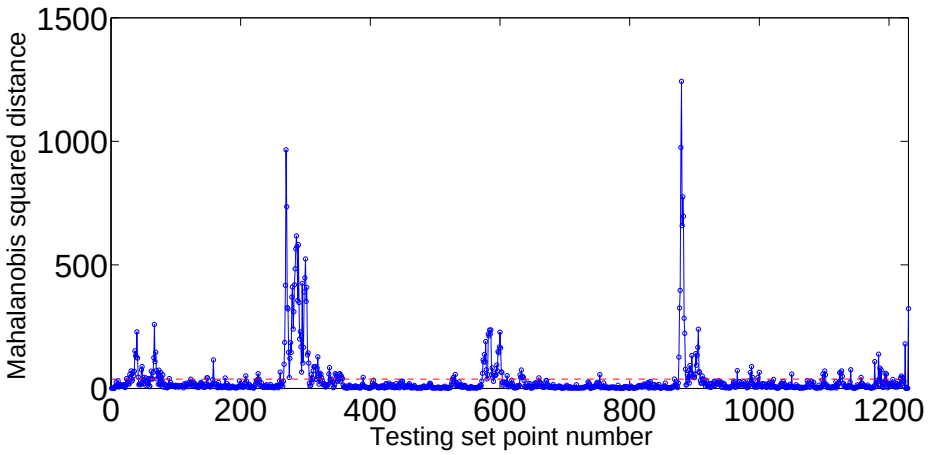


Figure 7.2: Outlier statistics for the power of the 2^{nd} IMF for the 11/2/2010 dataset (training data 200-400).

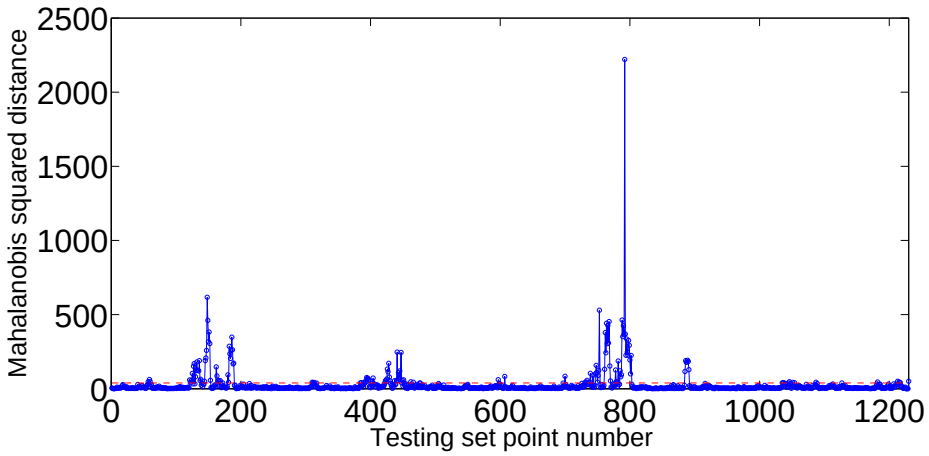


Figure 7.3: Outlier statistics for the power of the 2^{nd} IMF for the 4/4/2009 dataset (training data 350-550).

The results of the outlier analysis using the Mahalanobis squared distance measure, presented here in Figures 7.1, 7.2 and 7.3, show that the method manages to detect the peaks in the power. In order to eliminate the additional alarms detected, a different solution will be given in the next section however, where a spatially adaptive thresholding method will be proposed.

7.3 Phase-Space Thresholding Method

The reason for choosing this thresholding method was simply the observation that what appears as a feature in the power diagrams of the analysed data resemble spikes. This particular method was developed specifically for detecting spikes in *acoustic dopler velocimeter* (ADV) data in reference [142], and was used since then successfully in several studies in the same field [143].

As mentioned in [142] the method is based on the following three concepts:

- that differentiation enhances the high frequency portion of a signal,
- that the expected maximum of a random series is given by a *universal threshold*, and
- that good data cluster in a dense cloud in phase space or Poincaré maps.

Briefly, the method constructs an ellipsoid in three-dimensional phase space

and points lying outside the ellipsoid are designated as spikes. Other algorithms for spike detection, well known in electrical engineering, tested in this paper are: *RC filters* method, *Tukey 53H* method, *acceleration thresholding* method and *wavelet thresholding* method. All of them seem to perform worse according to [142].

A three-dimensional Poincaré map or a phase space map is a plot in which a variable and its derivatives are plotted against each other. The threshold used in this case is defined by the Universal criterion and an ellipsoid is formed in this way in the three-dimensional space that separates inliers from outliers. The universal threshold arises from a theoretical result from normal probability distribution theory. It was introduced in reference [144], as part of the *wavelet thresholding method*. For n independent, identically distributed, standard, normal random variables ξ_i the expected absolute maximum is:

$$E(|\xi_i|_{max}) = \sqrt{2\log n} = \lambda_U \tag{7.3}$$

where λ_U is termed the *universal threshold*. For a normal, random variable, that consists of n data points and whose standard deviation is estimated by σ and the mean is zero, the expected absolute maximum is:

$$\lambda_U \sigma = \sqrt{2\log n}\sigma \tag{7.4}$$

The algorithm consists of a number of sequential iterations that stop when the number of good data become constant (or, equivalently the number of new points identified as outliers, peaks in the power diagrams in this case, falls to zero). If u_i is the dataset being analysed, then each iteration has the following steps:

- Calculate *surrogates* Δu_i and $\Delta^2 u_i$ for the first and second derivatives from:

$$\Delta u_i = (u_{i+1} - u_{i-1})/2 \tag{7.5}$$

$$\Delta^2 u_i = (\Delta u_{i+1} - \Delta u_{i-1})/2 \tag{7.6}$$

- Calculate the standard deviations of all three variables σ_u, $\sigma_{\Delta u}$ and $\sigma_{\Delta^2 u}$, and then the expected maxima using the *universal criterion*.

- Calculate the rotation angle of the principal axis of the $\Delta^2 u_i$ versus u_i using the cross correlation:

$$\theta = tan^{-1}(\frac{\sum u_i \Delta^2 u_i}{\sum u_i^2}) \tag{7.7}$$

- Each set of variables $\{u_i, \Delta u_i, \Delta^2 u_i\}$, determines a point in spherical coordinates. For each pair of these variables, an ellipse can be calculated. Therefore, for Δu_i versus u_i the major axis is $\lambda_U \sigma_u$ and the minor axis is $\lambda_U \sigma_{\Delta u}$; for $\Delta^2 u_i$ versus Δu_i the major axis is $\lambda_U \sigma_{\Delta u}$ and the minor axis is $\lambda_U \sigma_{\Delta^2 u}$; and for $\Delta^2 u_i$ versus u_i the major and minor axes a and b respectively, can be shown by geometry to be the solution of:

$$(\lambda_U \sigma_u)^2 = a^2 cos^2\theta + b^2 sin^2\theta \tag{7.8}$$

$$(\lambda_U \sigma_{\Delta^2 u})^2 = a^2 sin^2\theta + b^2 cos^2\theta \tag{7.9}$$

- For each projection in space the points that lie outside of the ellipse are identified and replaced with a smoothed estimate in order to perform the next iteration.

At each iteration, replacement of the outliers reduces the standard deviations calculated in two and thus the size of the ellipsoid reduces until further outlier replacement has no effect.

Figures 7.4, 7.5 and 7.6 show the 3D phase space maps produced by the application of the method on the power diagrams shown in Figure 6.16 again. They show the ellipsoid produced by the sequential iterations described by the method's steps previously. The outliers here are shown as red points out of the ellipsoid. Finally Figures 7.7, 7.8 and 7.9 show the final estimation of outliers using the method. The diagrams are very satisfying, since all of the power peaks have been detected.

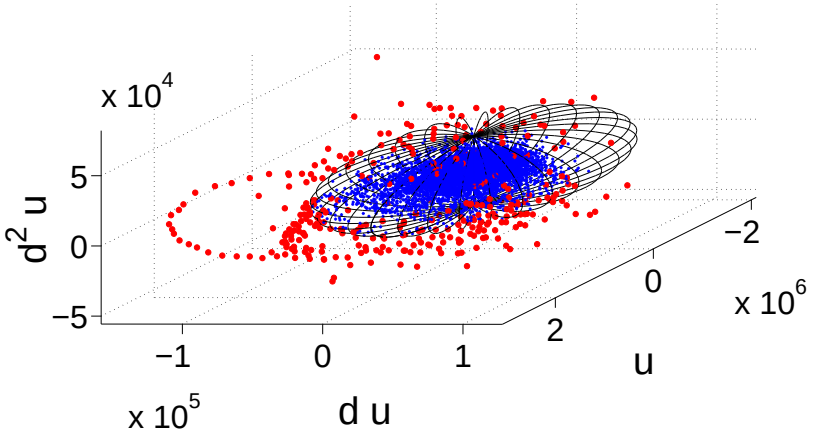


Figure 7.4: Constructed ellipsoid in 3D phase space for the power of the 2^{nd} IMF for the 31/10/2009 dataset.

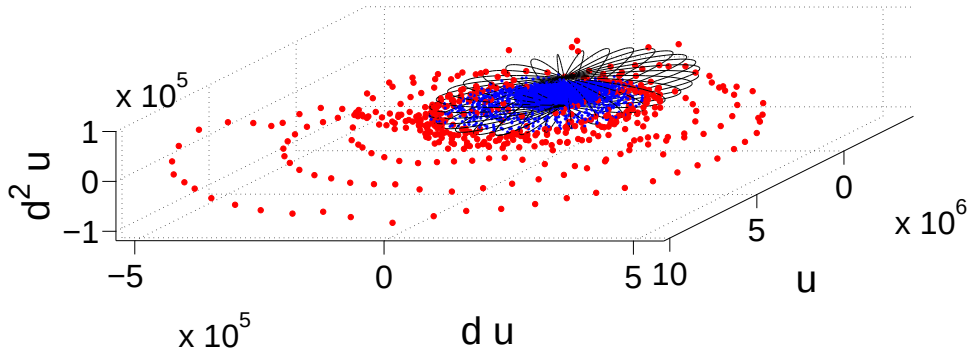


Figure 7.5: Constructed ellipsoid in 3D phase space for the power of the 2^{nd} IMF for the 11/2/2010 dataset.

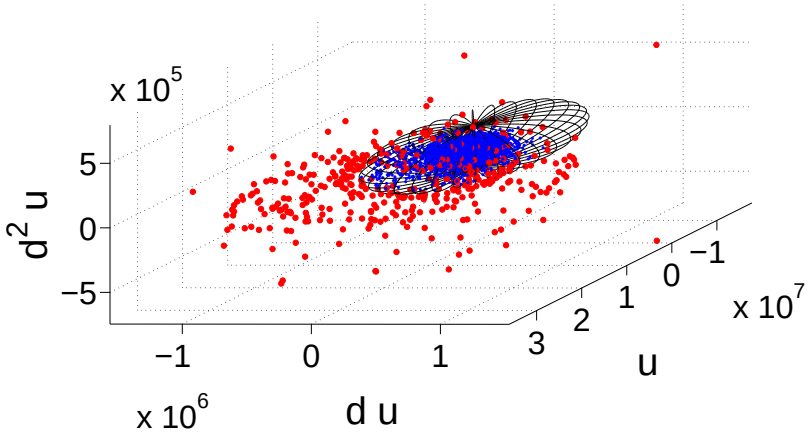


Figure 7.6: Constructed ellipsoid in 3D phase space for the power of the 2^{nd} IMF for the 4/4/2009 dataset.

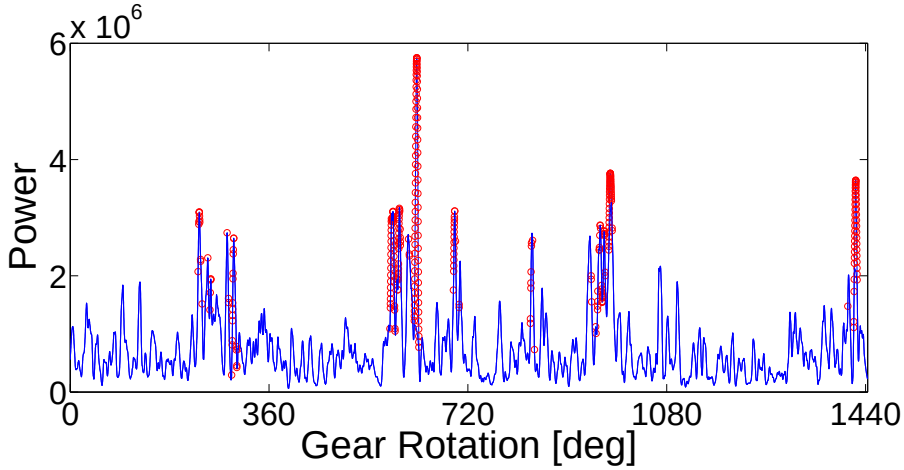


Figure 7.7: Estimated outliers using the spatially adaptive threshold (power of the 2^{nd} IMF for the 31/10/2009 dataset).

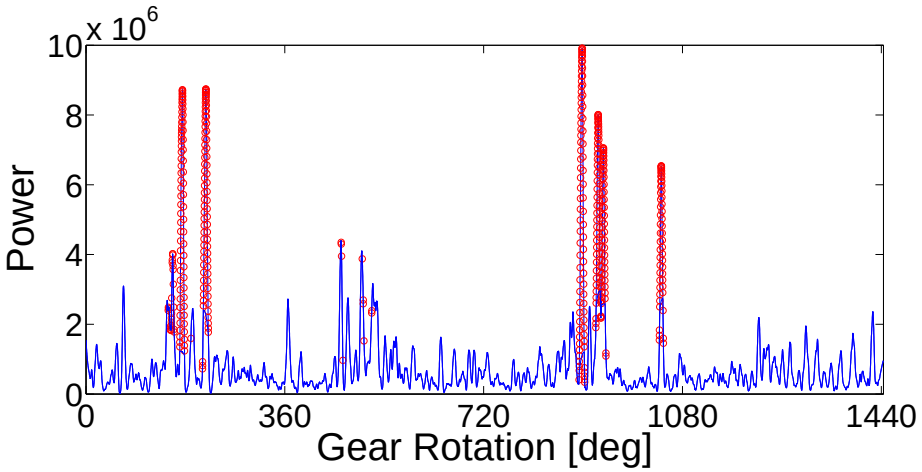


Figure 7.8: Estimated outliers using the spatially adaptive threshold (power of the 2^{nd} IMF for the 11/2/2010 dataset).

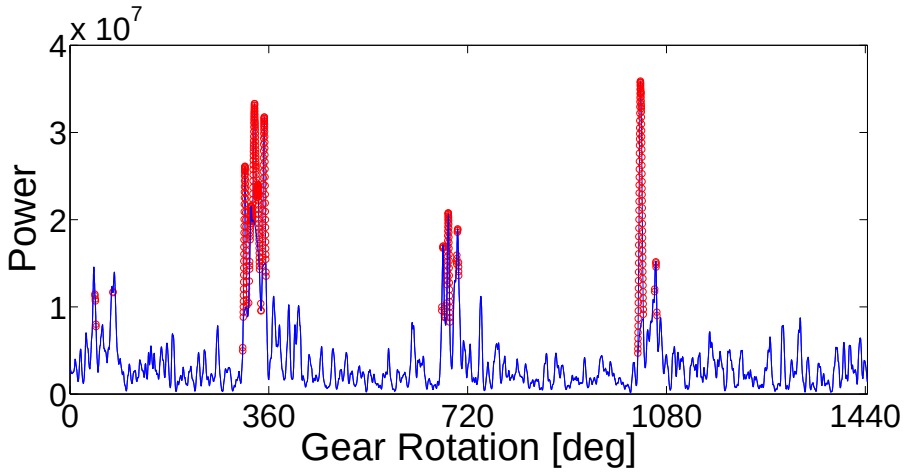


Figure 7.9: Estimated outliers using the spatially adaptive threshold (power of the 2^{nd} IMF for the 4/4/2009 dataset).

7.4 Conclusions

In this chapter two different novelty detection methods are applied to the features obtained from the analysis of the gearbox datasets in the previous chapter. The first one has been applied many times in the SHM field, but in a quite different way than what was proposed in this part of the study. The reason for that is mainly the fact that the work presented in the thesis is based on time-frequency analysis. This attempt to use the Mahalanobis Squared distance for OA in the proposed way, gave interesting results that suggest that the method could be a simple choice for the specific application.

The second method presented and used in this chapter is an adaptive thresholding method (*3D-phase space thresholding*) that has not been used before in the CM field. The fact that it was initially introduced and proposed for the purpose of spike removal is what drew the attention in this particular case. Since the kind of features examined here are peaks in the power diagrams of the Hilbert or Teager spectra (and generally any kind of time-frequency analysis spectra), the method proved quite effective, identifying all the points in the diagrams related to damage.

All the analysis presented upto this chapter could be considered to be a complete damage detection scheme for a wind turbine gearbox, with satisfying results. In the chapter that follows, a different kind of approach will be proposed, that originates from econometrics and has been recently used in the SHM field, for damage detection of bridges. It is completely different from a time-frequency or time-scale analysis method from the point of view that it does not decompose a signal into a set of time-scale signal components; what the method is able to do is to remove common trends by constructing a residual that is the linear combination of the several measurements taken from the examined structure.

Chapter 8

How cointegration helps in CM of the wind turbine gearbox

The use of cointegration has been proposed recently as a potentially powerful means of removing confounding influences from SHM data. The reason for this application is that in the real world, SHM is mainly inhibited by the fact that benign influences, most of the time related to environmental and operational changes, can reduce the effectiveness of damage detection methods. Sometimes, these effects are shown in the measured signals as long term trends. In addition, although it is the longer time scales of the data that are mostly influenced by the environmental/operational variations around the structure, it is sometimes also the case that nonstationarity from benign causes may appear on different time scales as well, and possibly the ones that damage is expected to manifest on. Removing such trends from nonstationary time-series data is therefore of great interest in the SHM field and unfortunately the application of just a bandpass filtering technique probably would not prove sufficient to remove nonstationary effects from the signals not related to damage at the scales that are important for damage detection. Cointegration, a technique taken from econometrics, on the other hand seems to solve this problem.

In the context of CM of wind turbine gearboxes, it has been discussed in the introduction and other chapters of this thesis that operational variations do exist. Varying loads and varying speeds of the drivetrain might influence

the time-scales on which damage features might occur. As a consequence, the application of cointegration for removing such influences not related to damage would be of great interest. In this chapter, the basic theoretical background of cointegration will be given, and in the next sections a multi-scale cointegration approach will be proposed that could be applied for CM purposes. In addition, since as it will be shown cointegration requires the existence of measurements taken from more than one sensor, a technique in order to make it possible to apply cointegration to just a single sensor measurement will be proposed. This idea was found useful for cases where there might be lack of data from different sensors.

In the next section the theory of cointegration follows.

8.1 Cointegration theoretical background

Cointegration is a property of some nonstationary multivariate time series; an n-dimensional time series is cointegrated if some linear combination of the component variables is stationary. The combination is called a *cointegrating relation*, and the coefficients form a *cointegrating vector*. In general, there may be multiple cointegrating relations among the variables in the time series. A brief description of the cointegration process for SHM in general, and as a consequence CM in particular, will follow. A major part of what is going to be presented is taken from reference [145]. Since the mathematics around the theory of cointegration is sometimes rather complex, here only the basic steps of the method will be described. References [141, 145] provide a tutorial treatment of cointegration in the context of SHM problems for those who would like to have a more detailed view of the process and study its application to the SHM of bridges.

The analysis approach is largely based on the *Johansen procedure* which is able to find the cointegrating vector that will result in the most stationary combination of the variables.

It is important to say here that the time series, apart from sharing common trends, should also have the same *degree of nonstationarity* if they are cointegrated. This degree of nonstationarity is related to what is known as the *order of integration* of the time-series. A nonstationary time series is integrated if it can become stationary through differencing. The number of differences needed in order to achieve stationarity is called the *order of integration* and can be estimated by a stationarity test, in this case the

augmented Dickey-Fuller test (ADF) [146] is the one that is used.

The summary of the cointegration process is given in the following steps. The Johansen procedure is used to find the most stationary linear combination of a set of monitored variables.

- The first step is to check the suitability of the monitored variables for cointegration. They should have the same order of integration. Because of an explanation given in [147], for most of the SHM cases the vibration signals are of integrated order one $I(1)$, i.e. a non-stationary variable with first difference stationary. So the cointegration procedure will be described this kind of variable. First, the ADF test is applied on each variable:
  - Fit each variable in question to the following time-series model using a least squares approach,

$$\Delta y_i = \rho y_{i-1} + \sum_{j=1}^{p-1} b_j \Delta y_{i-j} + \epsilon_i \tag{8.1}$$

    where Δ is a difference operator defined as $\Delta y_{i-j} = y_{i-j} - y_{i-j-1}$ and ϵ_i is Gaussian noise.
  - The least-squares estimate of the parameter ρ is used to infer the degree of nonstationarity of the variable. This is basically a unit root test. If ρ is statistically close to zero, the process will be non-stationary and integrated order one. The null hypothesis of $\rho = 0$ is tested by comparing the value of the test statistic.

$$t_\rho = \frac{\widehat{\rho}}{\sigma_\rho} \tag{8.2}$$

    where $\widehat{\rho}$ is the least squares estimate of ρ and σ_ρ is the variance of the parameter, with the critical values from the *Dickey-Fuller tables* (DF, see [148]). If the hypothesis is accepted, the time series has a unit root and is $I(1)$. If it is rejected, the test is repeated for Δy_i, if the hypothesis is then accepted, y_i is a $I(2)$ nonstationary sequence. This is continued until the integrated order of the time-series is ascertained.
- The Johansen procedure can now be applied to variables found to be order one. If they are cointegrated the Johansen procedure will find the most stationary linear combination of a set of monitored variables:

– Fit the variables in question with a *vector autoregressive* (VAR) model:

$$\{y_i\} = [A_1]\{y_{i-1}\} + [A_2]\{y_{i-2}\} + \cdots + [A_1]\{y_{i-1}\} \tag{8.3}$$

and use the *Akaike information criterion* (AIC) to determine the most suitable model order.

– The best *cointegrating vectors* are found as the columns of $[\beta]$ in the *vector error-correction model* (VECM) of the parameter set that takes the form:

$$\{z_{0i}\} = [a][\beta]^T\{z_{1i}\} + [\Psi]\{z_{2i}\} + \epsilon_i \tag{8.4}$$

where $\{z_{0i}\} = \{\Delta y_i\}$, $\{z_{1i}\} = \{y_{i-1}\}$, $\{z_{2i}\} = \{\{\Delta y_{i-1}^T\}, \{\Delta y_{i-2}^T\}, \ldots, \{\Delta y_{i-p}^T\}, \{D\}^T\}$, p is the model order ascertained previously in the first step of the Johansen procedure and $\{D\}^T$ is a deterministic trend. To find $[\beta]$, the cointegrating vectors, first establish the residuals $\{R_{0i}\}$ and $\{R_{1i}\}$ of the following regressions:

$$\begin{cases} \{z_{0i}\} = [C_1]\{z_{2i}\} + R_{0i} \\ \{z_{1i}\} = [C_2]\{z_{2i}\} + R_{1i} \end{cases} \tag{8.5}$$

– Define the product moment matrices:

$$[S_{mn}] = \frac{1}{N}\sum_{i=1}^{N}\{R_{mi}\}\{R_{ni}\}^T \tag{8.6}$$

The required cointegrating vectors are the eigenvectors of the generalised eigenvalue problem:

$$(\lambda_i'[S_{11}] - [S_{10}][S_{00}]) \tag{8.7}$$

The eigenvector with the corresponding largest eigenvalue is the a cointegrating vector that gives the most stationary combination of the original variables.

- Once a suitable cointegrating vector is found, new data from the monitored variables are projected on it. If data from the normal condition of the structure were used in order to create the cointegrating vector, then the residual from the linear combination will continue to be stationary all the time the structure continues to operate in its normal condition. If the residual deviates from stationarity, this is taken as an indication that the structure may be damaged.

- Finally a trace test statistic can also be carried out to indicate the number of cointegrating vectors possible for a set of variables.

8.2 Cointegration using a single sensor

It is implied from the above that in order to perform the cointegration algorithm, datasets from different sensors are needed. This requires additional instrumentation and memory storage, something not always feasible for the case of on-line monitoring. In this section a method is proposed that potentially allows cointegration to remove trends when a single sensor is present. While this idea has been used implicitly before where multivariate natural frequency data have been extracted from accelerometer data [149] and a single acceleration could potentially have sufficed for this purpose; the analysis in [149] was carried out in batch mode at intervals much longer than the sampling interval for the raw data. The requirement for off-line analysis reduced the potential response time of the diagnostic system. Here, a general method is proposed that is applicable to any measurement variable, operating in a truly on-line manner and does not require *a priori* specification of a physics-based model form. The idea is based on the application of recursive AR-modelling and is demonstrated on synthetic data, showing the method's potential to be applied on vibration data measured from a wind turbine transmission system, where cointegration can be adapted as a solution for extracting the load variation influences in the gearbox vibration signals.

8.2.1 Recursive Least Squares parameter estimation

The *recursive least squares* (RLS) algorithm is a parameter estimation method which updates the AR model coefficient estimates, each time a new sample data becomes available [150].

In the current application, this procedure is not associated with the steps followed in order to perform cointegration. It will be performed separately, before the application of cointegration and therefore has nothing to do with the model fitting steps of cointegration.

Suppose $\{\beta\}_i$ is the parameter vector at the i^{th} sampling instant and $[P]_i$ is the corresponding iterate of the covariance matrix. The RLS algorithm can be summarised by the following equations:

$$\{\theta\}_{(i+1)} = (\dot{y}_{(i-1)}\, \dot{y}_{(i+1)}\, y_{(i+1)})^T \tag{8.8}$$

$$\{K\}_{(i+1)} = \frac{[P]_i\{\theta\}_{(i+1)}}{1 + \{\theta\}^T_{(i+1)}[P]_i\{\theta\}_{(i+1)}} \tag{8.9}$$

$$[P]_{(i+1)} = (1 - \{K\}_{(i+1)}\{\theta\}^T_{(i+1)})[P]_i \tag{8.10}$$

$$\{\beta\}_{(i+1)} = \{\beta\}_i + \{K\}_{(i+1)}(x_{(i+1)} - \{\theta\}^T_{(i+1)}\{\beta\}_i) \tag{8.11}$$

The iteration is started with $\{\beta\}_0 = \{0\}$ and $[P]$ is initialised as a diagonal matrix with large entries, because large entries encode the little confidence in the initial estimate. With a simple modification to the above equations one can allow past data to be weighted with a forgetting factor λ, that essentially imposes an exponential window on the past data (allowing to the algorithm to forget samples faster during the parameter estimation updating). In this case the update formulae are:

$$\{K\}_{(i+1)} = \frac{[P]_i\{\theta\}_{(i+1)}}{\lambda + \{\theta\}^T_{(i+1)}[P]_i\{\theta\}_{(i+1)}} \tag{8.12}$$

$$[P]_{(i+1)} = \frac{1}{\lambda}(1 - \{K\}_{(i+1)}\{\theta\}^T_{(i+1)})[P]_i \tag{8.13}$$

Usually a value for λ in the range of $0.9 - 0.999$ is adopted. The choice of λ is a compromise between the ability of the algorithm to track changes in the parameters (small λ values needed), and the need to suppress unwanted variations due to measurement noise (when the λ value should be close to unity).

8.2.2 Simulation example

In this section a simulation example of the proposed approach for removing environmental or operational trends from one damage sensitive variable will be illustrated. The dynamic response of a *three-degree-of-freedom* (3 DOF) system, shown in Figure 8.1, to a random excitation is simulated. The random excitation is applied to the first mass as shown in the figure. The system masses are $M_1 = M_2 = M_3 = 1$ $[kg]$, and the damping values are $C_1 = C_2 = C_3 = 20$ $[Ns/m]$. The simulation for the first 3000 points runs with steady stiffnesses $K_1 = K_2 = K_3 = 1000000$ $[N/m]$. In order to introduce effects imitating environmental or operational conditions the next

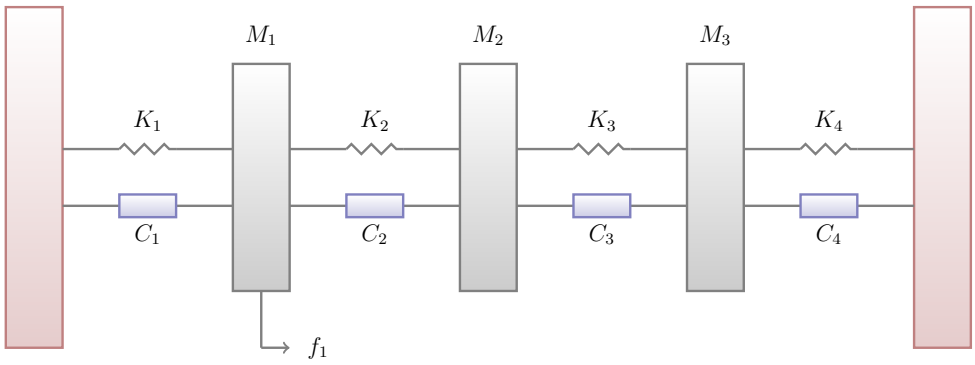


Figure 8.1: Simulated 3 degree of freedom lamped mass model.

3000 points, (sample points $3001 - 6000$), are obtained from a simulated system that this time includes a sinusoidal variation in its stiffnesses, more particularly: $K_1 = 1000000 + 500000sin(t)$, $K_2 = 1000000 + 400000sin(t)$, $K_3 = 1000000 + 300000sin(t)$ and $K_4 = 1000000 + 200000sin(t)$, where $\omega = 3\pi$. Finally in order to introduce damage, the simulation runs again for the next 3000 points, (sample points $6001 - 9000$) with reduced stiffness parameters that include the sinusoidal variations and this time correspond to the values: $K_1 = 750000 + 500000$, $K_2 = 750000 + 400000sin(t)$, $K_3 = 750000 + 300000sin(t)$ and $K_4 = 750000 + 200000sin(t)$.

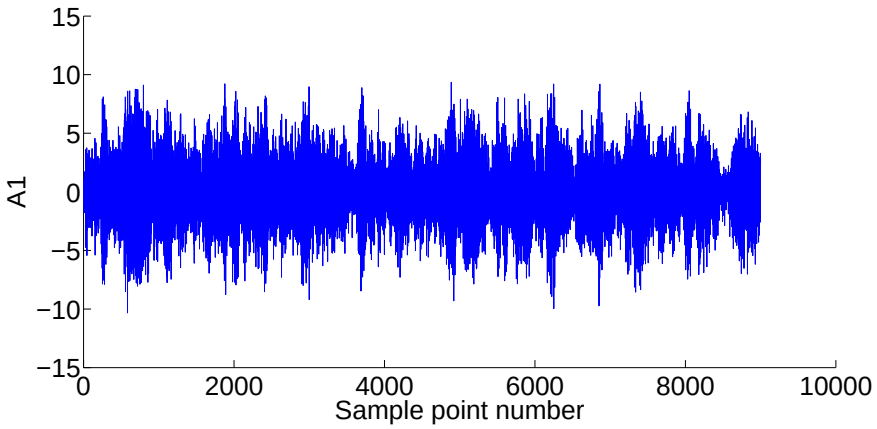
The RLS algorithm with a forgetting factor $= 0.9$, described in the previous section, was used to fit AR models to the simulated accelerations at the model's masses (A1, A2, A3). The order of the AR models was chosen to be 20, so 20 coefficients were estimated for each acceleration. The proposed method produces updated estimates in time, and an appropriate choice of the forgetting factor can help tracking the sinusoidal changes related to the stiffness variation, in the estimated coefficients, if the sampling time is chosen to be faster than the variations in the dynamics of the system. The first five coefficients produced were used to perform cointegration analysis. In this way, the ability of cointegration to remove the effects not related to the simulated damage will be tested, since these coefficients carry information about the system's parameters. The training datasets were chosen to be between the data points $2200 - 4000$ as depicted by the dashed lines in the Figures 8.3 (b), 8.4 (b) and 8.5 (b).

The simulations produced the acceleration datasets shown in Figure 8.2. The first step of the study is to produce the AR coefficients using the recursive least squares method with a forgetting factor 0.9. The choice of the forgetting factor in this case was made in such a perspective that the stiffness

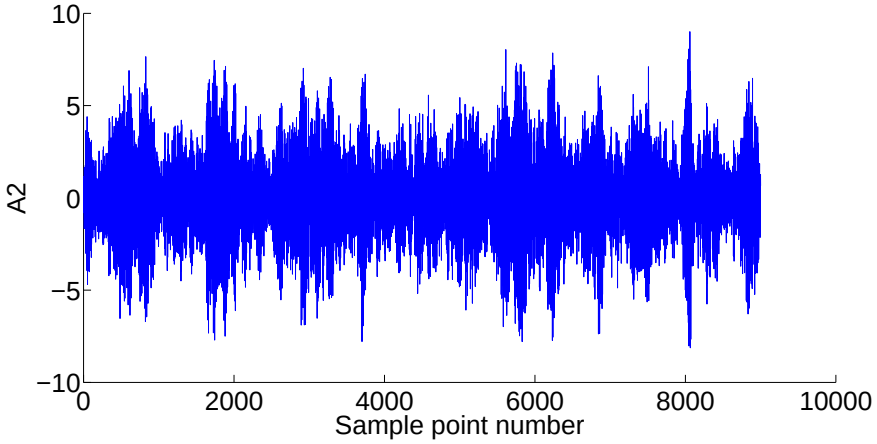
simulated sinusoidal variations, representing environmental or operational variations, can be tracked by the algorithm as mentioned earlier. The damage sensitive features examined in this case are the first five coefficients of the three simulated accelerations, but the coefficients of the velocities or displacements could be used as features instead as well. The aim then, is to use cointegration in order to perform damage detection by creating a feature insensitive to the simulated variations, that could represent either temperature or operational variations, but is still sensitive to damage.

Figures 8.3 (a), 8.4 (a) and 8.5 (a) show the first five, (out of twenty), coefficients produced, and the influence of the sinusoidal part of the stiffness is obvious for the last 6000 points of the datasets. The next step is to follow the procedure briefly described in the previous section for the cointegration of the coefficients. This procedure includes the application of several stationarity tests, that determine the order of integration, to all the datasets to be cointegrated (in this case the coefficients produced), and those suitable for cointegration are then combined using the Johansen procedure to create a stationary linear combination of them, (the residual). This stationary linear combination is established using the appropriate training set of data, that here includes points from the first part and the second part of the simulation. For a feature to be considered successful, it is necessary that it becomes non-stationary only on the occurence of damage regardless of other fluctuations. Figures 8.3 (b), 8.4 (b) and 8.5 (b) show the cointegrated residual produced for the training period chosen. The dashed horizontal lines indicate ± 3 standard deviations around the mean of the training data and act essentially as a statistical process control X-chart. Data points outside this threshold are considered abnormal.

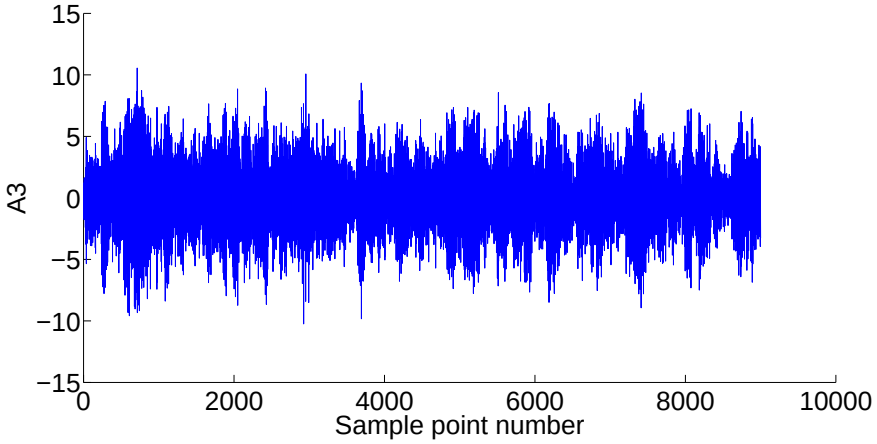
Figure 8.3 (b) showing the cointegration results of the coefficients of A1 are probably the least satisfying. The results get clearer while moving away from the force driving point (mass M1), with the results shown in Figure 8.4 (b), being slightly better. In Figure 8.5 (b), it is even more obvious that the approach proposed can be promising. The Johansen procedure created a stationary residual of the 5 coefficients over the training period, with one anomaly detected at the specific points where the second part of the simulation began (sample point number 3000). This is not a sustained excursion outside the confidence intervals however, so it rather shows the transition period of the dataset, after which the features return to an equilibrium and are still valid for anomaly detection. With the introduction of damage (sample point number 6000), a clear indication of damage is apparent. The residual becomes non-stationary, deviating periodically and significantly out of the control chart boundaries.



(a)

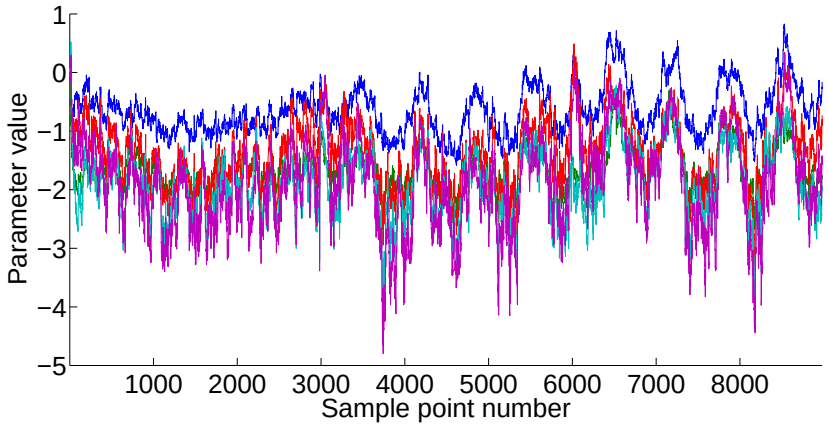


(b)

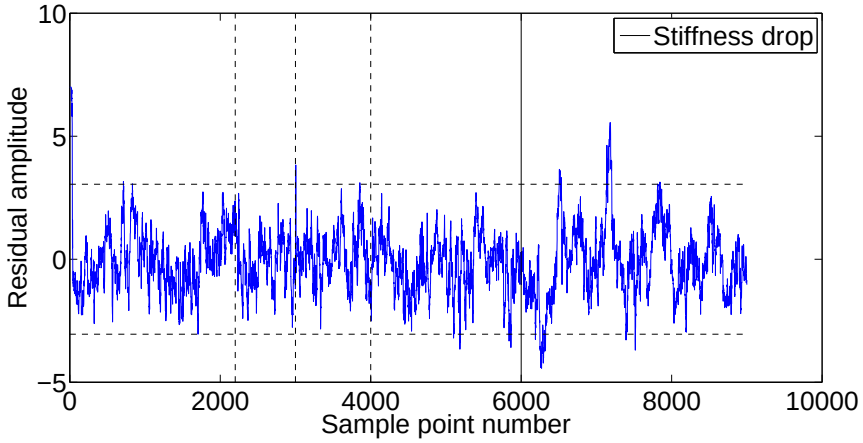


(c)

Figure 8.2: Simulated accelerations of the 3DOF system: (a) A1, (b) A2 and (c) A3.

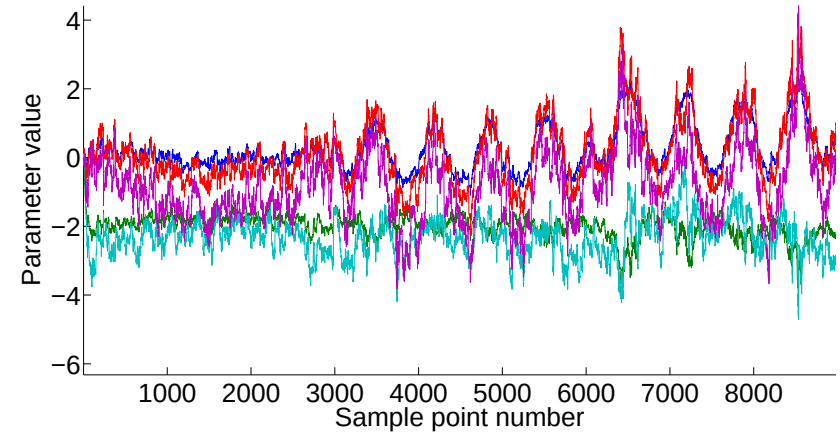


(a)

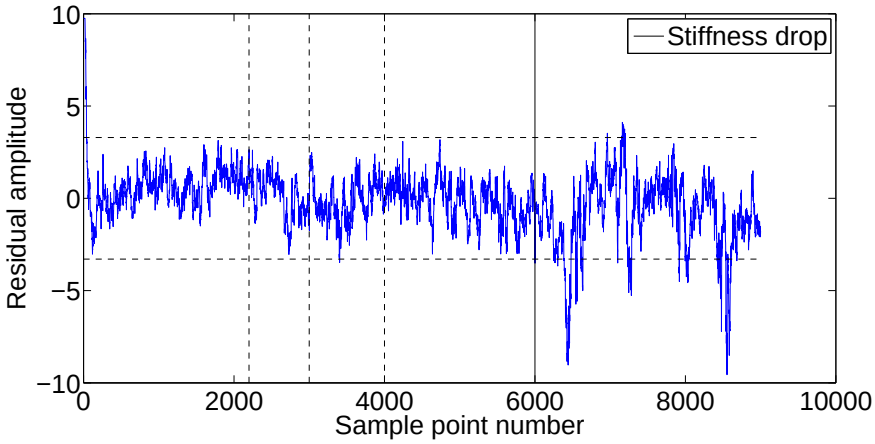


(b)

Figure 8.3: (a) First five estimated parameters of signal A1 and (b) the cointegrated residual produced



(a)



(b)

Figure 8.4: (a) First five estimated parameters of signal A2 and (b) the cointegrated residual produced.

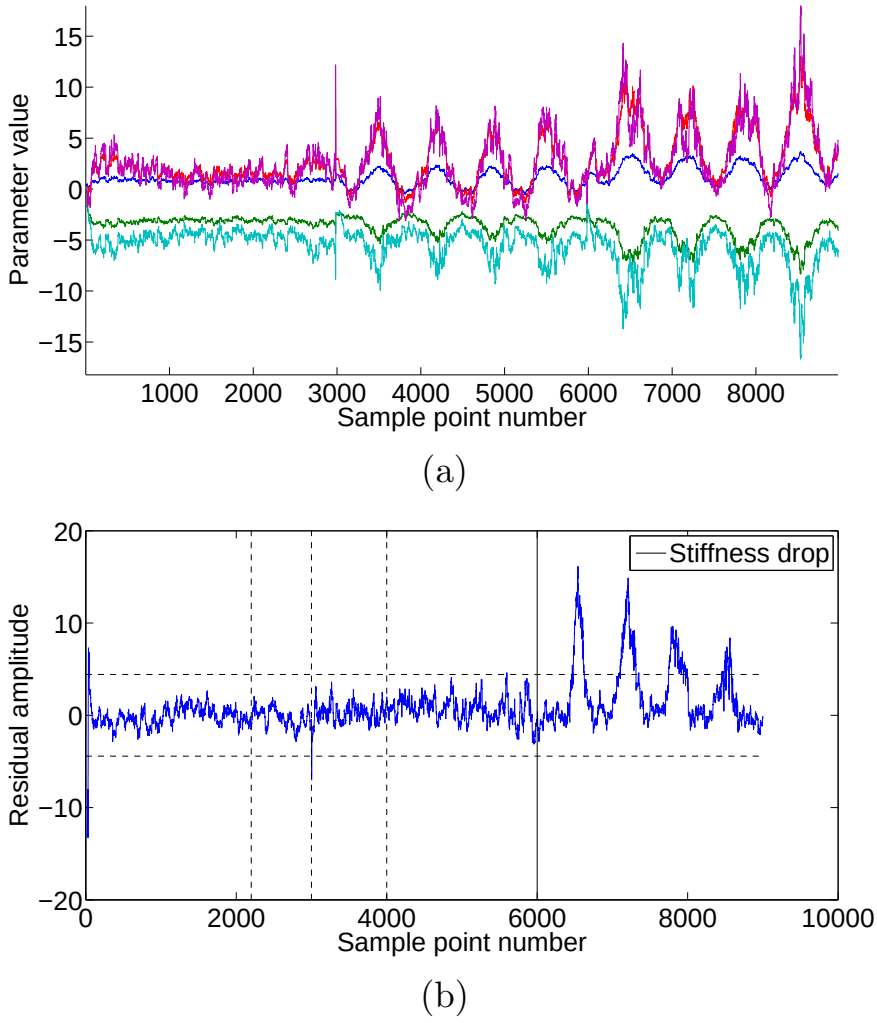


Figure 8.5: (a) First five estimated parameters of signal A3 and (b) the cointegrated residual produced.

8.3 Multiresolution cointegration explained in an experimental application

Previous work pointed out that despite the fact that it is the longer time scales of the data that are mostly influenced by the environmental variations around the structure, it is sometimes also the case that nonstationarity from benign causes may appear on different time scales as well, and possibly the ones that damage is expected to manifest on [151]. This means that the application of just a bandpass filtering technique probably would not prove sufficient to remove nonstationary effects from the signals not related to damage at the scales that are important for damage detection. In that paper [151], links between the cointegration algorithm and multi-scale decomposition using wavelets were highlighted and scale-specific cointegration was suggested as a possible option to the benefit of damage detection. The paper was deliberately quite speculative and exposed a number of issues that might prove problematic for the wavelet-based approach. One problem was related to linear dependence between variables as a result of the sparse wavelet basis at low frequencies which induced conditioning problems in the cointegration procedure. A second problem was that the wavelet levels selected for analysis had to be chosen on an *ad hoc* basis. In this part of the thesis the EMD is proposed as a potential solution to both of these problems.

8.3.1 Experimental datasets

The datasets used in this analysis were collected within the DAMASCOS consortium. DAMASCOS was a Brite-Euram project which attempted to use Lamb waves for damage detection in a composite plate subjected to cyclic temperature variations. The plate material was a *carbon fibre-reinforced plastic* (CFRP) laminate with a 0/90 degrees lay-up. Two identical piezoceramic discs bonded to the midpoint of the edges of the plate were used to transmit and receive fundamental symmetric and antisymmetric Lamb-waves, Figure 8.6. The instrumented composite plate was placed in an environmental chamber in the particular test used, and Lamb waves were recorded every minute.

Manson [152] used the DAMASCOS data to investigate the feasibility of discovering features for damage detection from the response spectrum that are insensitive to environmental variations. In that work the features chosen for analysis were 50 spectral lines from the frequency spectrum (numbers

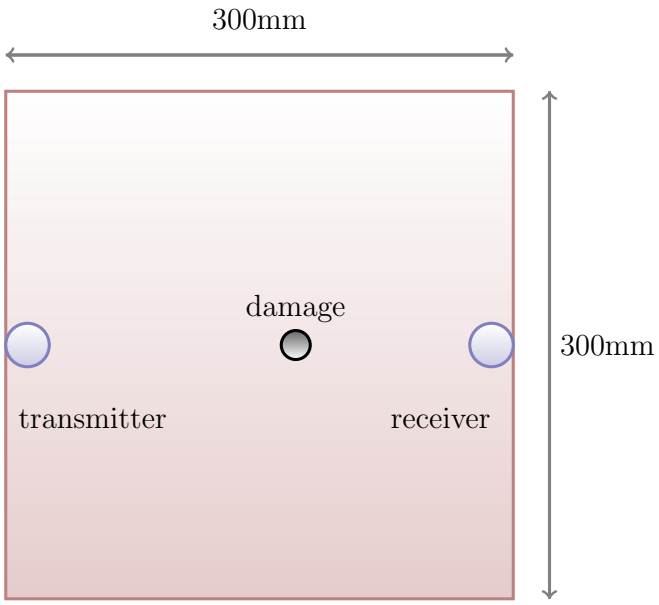


Figure 8.6: Thick composite plate (3mm) instrumented with piezoceramic transmitter.

$46-95$) which occurred around the peak. Figure 8.7 illustrates the variation of the amplitude of the 50 spectral lines under investigation for the duration of the testing. In the first part of the test, the temperature was constant at 25 degrees (sample points $0-1078$ in Figure 8.7). In the second part the temperature was cycled between 10 and 30 degrees every nine hours (sample points $1079-2159$). In the final part, apart from the temperature cyclic variation, there was an introduction of damage to the plate, by drilling a 10 mm hole between the two sensors (sample point 2207).

In a previous analysis [153], cointegration was used to create a feature that was insensitive to temperature-induced variation but still sensitive to damage. The details can be found in [153], but the basic principle was to use the Johansen procedure on training data encompassing the temperature variation but not damage, in order to create a cointegrated residual. This residual was then monitored for anomaly detection, with excellent results. In addition, a first attempt to combine multiresolution analysis and cointegration was made in the reference [151]. In that paper, discrete wavelet analysis was used as a means of characterising the time-scales on which nonstationarity manifested itself. The interesting results led to the current approach which aims to test whether the EMD method can improve the analysis. Here some of the wavelet results will be shown in parallel with the EMD results in order

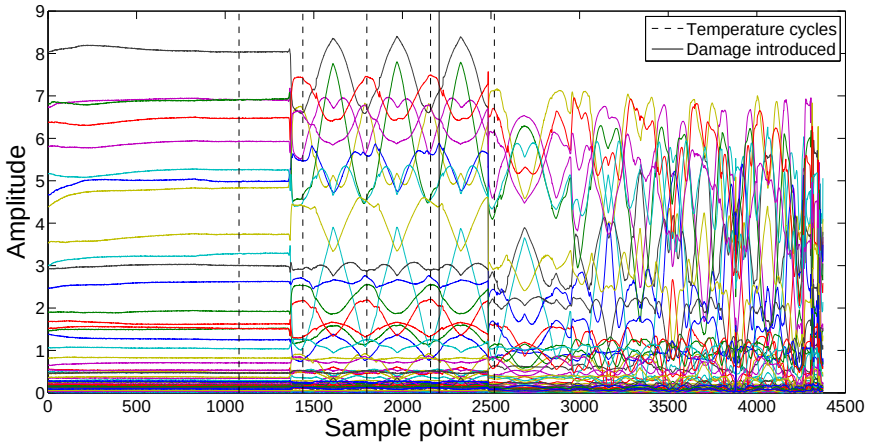


Figure 8.7: Features (spectral lines 46-95) over the whole period of the Lamb wave environmental test.

to have a clearer understanding of the advantages and disadvantages of both methods.

8.3.2 Multiresolution cointegration results: an EMD and Wavelets application

As explained in reference [151], the wavelet level decomposition is carried out by using the *orthogonal wavelet transform* (OWT, A.3). Because the OWT requires the number of samples to be a power of two, the feature data were truncated to 4096 (212) points by removing the first portion of the data where the temperature was held constant. This led to thirteen decomposed levels; however, the first two levels are essentially both copies of the mother wavelet and do not contain significant information from the signals. Unlike the EMD method the wavelet transform decomposes the signal to signal levels that begin from lower and end at higher frequencies. The Daubechies $N = 4$ wavelet was used for the analysis because, although it is less smooth than the higher-order wavelets, its shape is quite well matched to the *triangular wave* temperature ramping cycles. As shown in Figure 8.8 the level decomposition thus yielded a 13×20 matrix of level time series. (Levels are labelled here from 0 to 12. Level 0 is omitted from the figure as it contains the same wavelet as level 1 shifted by the signal mean).

The EMD method is also used to decompose the datasets examined, resulting in the IMFs shown in Figure 8.9.

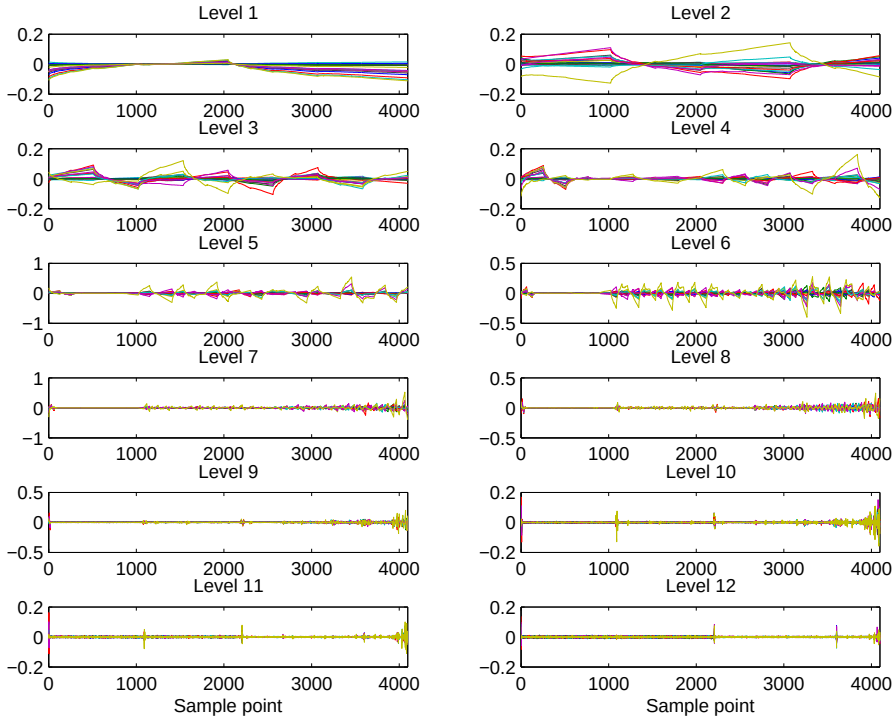


Figure 8.8: Wavelet levels for all Lamb wave spectral features [151].

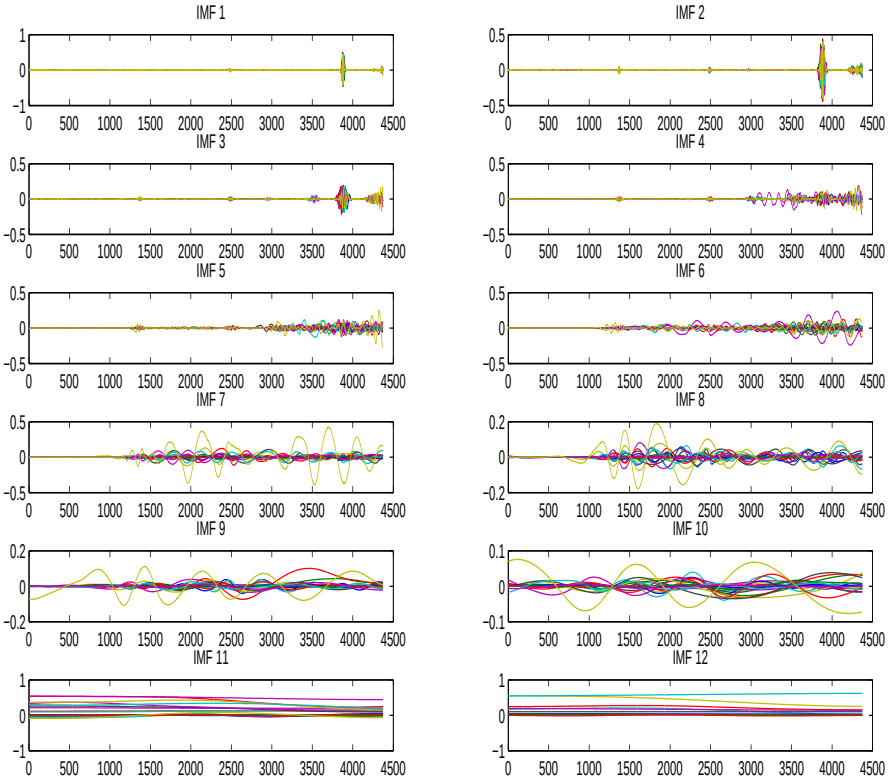


Figure 8.9: IMFs for all Lamb wave spectral features.

The stationarity of each wavelet level and each IMF can be then calculated using a statistical test, in this case the ADF statistic. In general, the ADF statistic becomes increasingly negative as signals become more stationary. The ADF statistic was computed for all the wavelet levels and all the IMFs (separately) and the results were averaged across the features to give a mean ADF statistic per level. These results are shown in Figures 8.10 and 8.11, where it can be seen that the wavelets decomposition process yields a sequence of signals steadily increasing in stationarity, while the IMFs are signal components with steadily increasing nonstationarity. Concerning Figure 8.10 the first 3 or 4 level values are probably not to be trusted as the signals actually reflect the nature of the mother wavelet rather than the the original signal. In a similar way (Figure 8.11) the last three IMFs mostly reflect the nature of the residual signal of the decomposition.

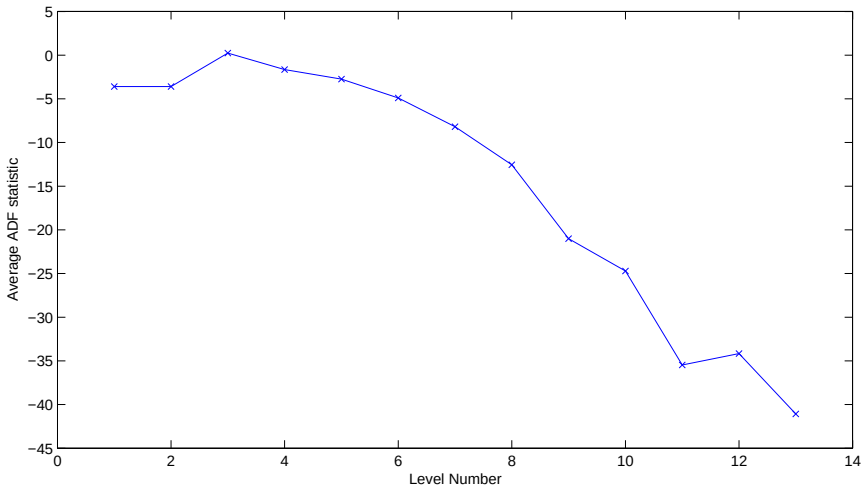


Figure 8.10: Mean ADF statistic across the feature set as a function of the wavelet level [151].

Multiresolution analysis offers the possibility to compute a cointegrated residual for damage detection on a level-by-level basis. This approach creates a large body of results, here only the results of the 12^{th} and 5^{th} wavelet level, as well as the 2^{nd} and 6^{th} IMFs are presented. The 12^{th} wavelet level is the 'most stationary' level of the wavelet decomposition analysis and in a similar manner the 2^{nd} IMF is the most stationary signal component (along with the 3^{rd} IMF) in the EMD analysis. Both the 12^{th} wavelet level and the 2^{nd} IMF represent the second highest frequency band of the signal, it is interesting to examine them, as an optimal case for SHM purposes. The cointegrated residuals for both cases are given in Figures 8.12 and 8.13. The dashed vertical

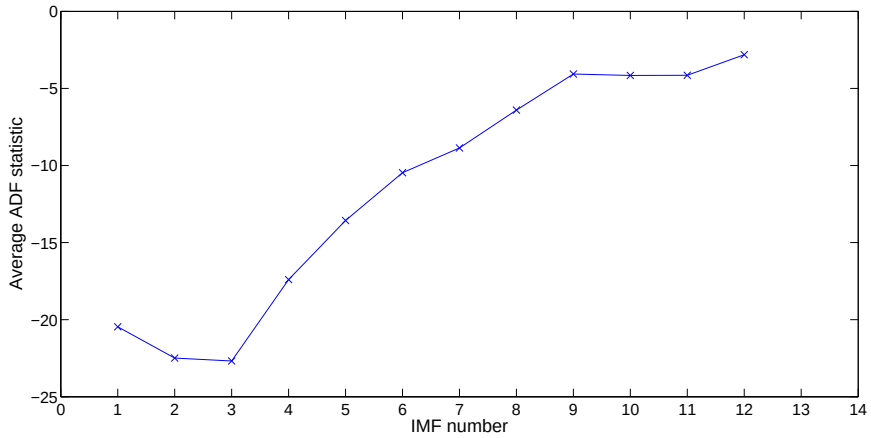


Figure 8.11: Mean ADF statistic across the feature set as a function of the IMF number.

lines in the figures indicate the region of training data used for the Johansen procedure (between samples 900 and 1500); they come some of the constant temperature regime and one full cycle of the temperature variation. The solid vertical line indicates the onset of damage. Finally, the horizontal dashed lines represent 3σ confidence levels as detection threshold computed from the statistics of the training data residual as is commonly used in statistical process control. The clearest indication of damage shown in both figures is the drop in the signal variance. This is not a good indicator of damage, as discussed in [151].

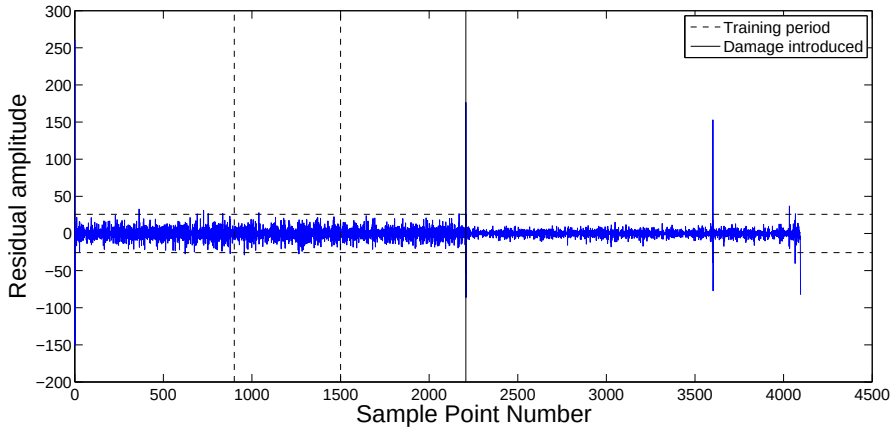


Figure 8.12: Cointegrated residual for 12^{th} wavelet level [151].

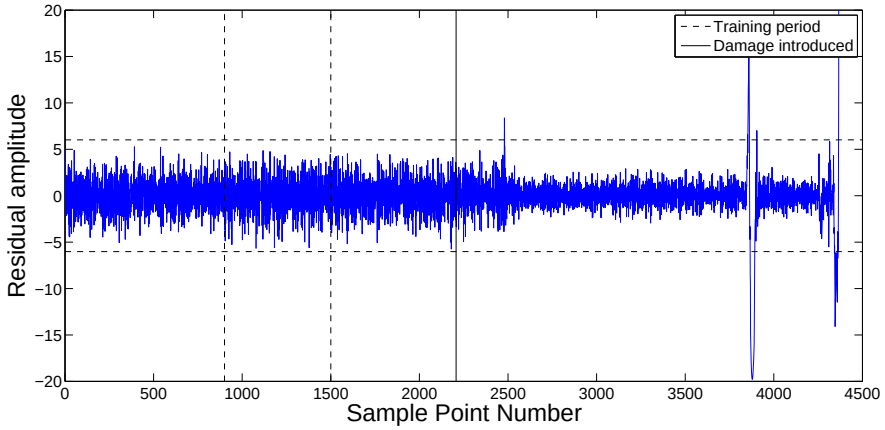
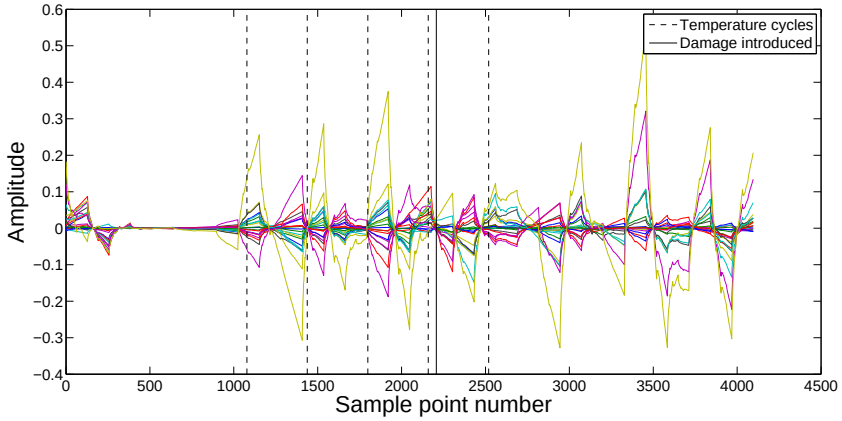


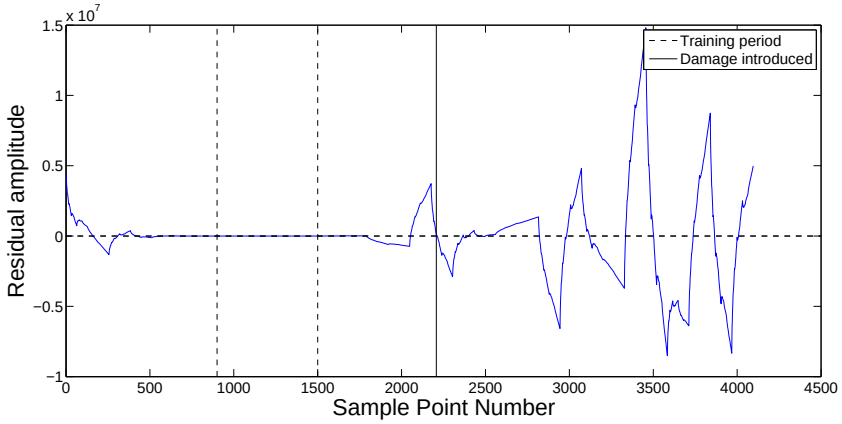
Figure 8.13: Cointegrated residual for IMF 2.

Having considered what happens at the most stationary signal components, it is interesting to see what happens at the 'least stationary' as well, and more particularly the ones corresponding to the time-scale of the temperature variations. For this purpose the cointegrated residual of the 5^{th} wavelet level and the cointegrated residual of the 6^{th} IMF are shown in Figures 8.14 and 8.16 (b) respectively. Apart from the fact that the residual causes an early alarm for the true damage, no other excursions are observed. The early alarm was discussed in [151]; it is not considered an immediate concern here for the following reason. The excursions before the actual onset of damage are akin to Gibbs phenomenon in Fourier analysis; the effects of the discontinuity (damage) are felt before and after it occurs at the scale of the individual wavelet level or individual IMF; because the two methods allow perfect reconstruction, the information to resolve the apparent non-causality will be present in other wavelet levels or IMFs, mainly corresponding to higher frequencies. The early alarms are simply an artefact of the batch processing applied here and will be removed when an on-line version of the methodology here is developed.

A major problem that arose in this case in the application of the Johansen procedure at level 5 (and some of the other levels) is an ill-conditioning issue; there exist more basis functions than pieces of data, since this level can only accommodate 16 basis functions, meaning that 20 features will be linearly dependent. This results in estimated residuals of magnitude that is 4 or 5 orders greater than the data themselves, Figure 8.14 (b). In order to solve this problem, regularisation is needed. The simplest way to do this is by adding noise to the feature data. A choice of a noise rms of approximately



(a)



(b)

Figure 8.14: (a) 20 features at level 5 of the decomposition and (b) Cointegrated residual for level 5: original data [151].

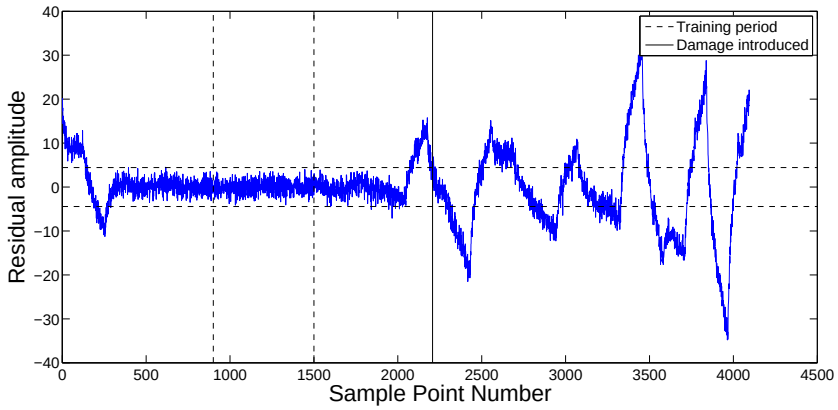
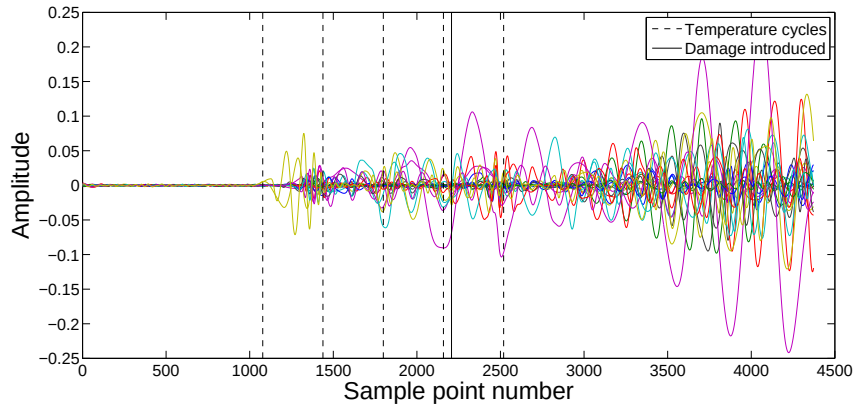


Figure 8.15: Cointegrated residual for level 5: 'regularised' data [151].

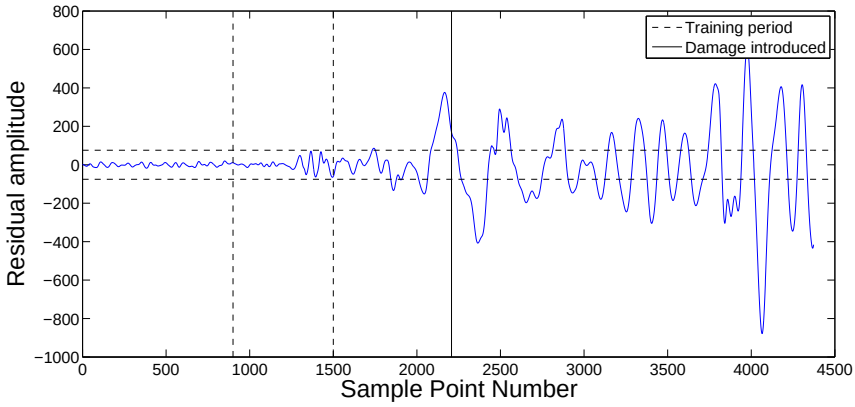
0.25% of the level variance, through a trial-and-error process, gave the results presented in Figure 8.15, that are considered to be good although the way of determining the noise levels was not performed in a principled manner. A more formal way of doing this should be through optimisation and cross-validation. This is probably the main disadvantage of the wavelet approach when compared to the EMD.

The EMD alternative to multiresolution cointegration might suffer from a different problem, though; the EMD method might not produce the same number of IMFs for the datasets analysed. This is a rather important issue since cointegration between components with slightly different time scales might occur. A solution to this, could probably be to use a noise assisted version of the EMD (e.g. the EEMD) that might act as a dyadic filter bank for all kinds of signals and would be able to produce possibly the same number of IMFs for every feature examined.

Despite the issues described above, it is clear from Figures 8.15 and 8.16 (b) that the cointegrated residual of the particular wavelet level and IMFs analysed is much more sensitive to damage than the previously presented results (Figures 8.12 and 8.13). This explains the fact that the most nonstationary signal component would be the most appropriate choice to perform cointegration and give excellent results that enhance the sensitivity to damage when compared to the application of a standard cointegration approach. According to [154] this because when damage breaks the relation imposed by cointegration the nonstationarity level in the signals is potentially the same as the nonstationarity level in the individual features.



(a)



(b)

Figure 8.16: (a) 20 features at IMF 6, and (b) cointegrated residual for IMF 6.

8.3.3 Conclusions

Previous work has shown that cointegration provides a powerful means of removing the effects of environmental and operational variations from data in order to enhance damage detection capability.

In this chapter a technique that can be used when one wishes to perform cointegration but lacks data from multiple sensors, was demonstrated. This is not uncommon, especially for real life on-line applications, where memory storage and sensor costs can be limitations in order to adopt any SHM method with high requirements of such kind. Applying cointegration to the condition monitoring of wind turbine gearboxes could be a solution to the main problem that inhibits signal processing analysis in on-line wind turbine gearbox condition monitoring, that is the time varying loads that these gearboxes undertake. Damage features in this case might be influenced in an unexpected way and false alarms could be produced because of these effects.

It was shown that the RLS parameter estimation method with a forgetting factor can produce coefficients, identifying the parameters of the system examined, that can track variations of the system parameters. These coefficients can be a suitable feature for damage detection and can be used successfully for cointegration analysis. The results presented in this chapter seem promising, but further research is important since the nature of the material of this study is rather exploratory.

Further work presented in this chapter, indicated that using cointegration in a multi-scale framework offered the possibility of enhancing the sensitivity of SHM algorithms; however, part of this later work based on wavelet analysis raised a number of issues that required more thought. One issue was that the sparse wavelet basis at low frequencies created conditioning problems with the Johansen procedure that required solution by regularisation; this problem has been solved by the adoption of the EMD as an alternative to the wavelet analysis. Another problem with the multi-scale approach was the fact that there are excursions above the alarm limit for the residuals before damage occurs; The problem is essentially a type of Gibbs phenomenon and is an artefact of the batch processing of the data; the solution to this problem will be to pass to a truly on-line algorithm for detection and this matter is currently under investigation.

A general comment about multiresolution cointegration, suggested by the results derived in reference [151], would be that there might be a connection between cointegration and multiscale analysis. This would be an interesting

concept for further investigation.

Chapter 9

Conclusions and further work

The main focus of this thesis was to develop and test advanced signal processing methods for wind turbine condition monitoring applications, that can overcome difficulties related to the varying operational conditions (load or speed) of these systems. For the feature extraction part of the damage detection process, time-frequency or time-scale analysis methods are the most appropriate in this case, in order for one to be able to represent the varying nature of the signals. For the pattern recognition part of the analysis, novelty detection methods should be chosen when only measurements from the undamaged structure can be used as a reference in the analysis. In addition what was attempted to be shown in this study is that there could be cases where just a filter bank method might not prove sufficient to detect damage when performing condition monitoring of wind turbine gearboxes or bearings or generally the drivetrain. In this case, a different kind of approach, such as cointegration, could probably provide a solution.

Summarising the work and the conclusions of this thesis:

- A description of the wind turbine gearbox experimental data that are mainly used in this study are given in Chapter 4. These measurements were taken by the EC Grupa, a Polish engineering company that maintains the wind turbine system. They come from a single accelerometer and were obtained at three different days representing different stages of damage. The damage was described as a tooth crack at the second gear stage of the gearbox.

Bearing datasets are also analysed at some points in this thesis. These come from a joint experiment with the Leonardo Centre for Tribology of the University of Sheffield. The measurements in this case were obtained during a fatigue test at two different dates, the first one when the bearing was healthy and the second one after introducing damage to it using the EDM method.

Finally, a simple gearbox model was simulated in this study, and is described in the same chapter. It allows for certain nonlinear and time-varying characteristics and takes into account varying loads similar to those found in wind turbines. These loads were simulated using FAST software and the applying such kind of loads to a gearbox model appears to be a novel idea.

- The time-frequency method proposed in this study is the EMD method. The main reasons for that are that it is an adaptive relatively recent method with some interesting features. Chapter 5 explains the basic theory of the EMD method. Various studies in the past have encountered a problem of the method, the mode mixing, that is why this problem is discussed. Through this discussion an interesting characteristic of this problem appears to have a potential trait to be used in the analysis that follows but with certain concerns and under specific cases. This trait is that mode mixing can create a significant higher number of IMFs when the signal analysed with the EMD method contains intermittencies. This was an interesting peculiarity of the algorithm in this particular case, since gear tooth damage and damage in bearings appears in the form of intermittencies.

  This chapter also introduces the TKEO and Desa-1 algorithms, as an alternative amplitude-frequency separation technique to HT. There are only very few applications of this operator in condition monitoring, and being a simple algorithm, its use in this study seemed of interest.

- In Chapter 6 the EMD method is applied first to the simulation datasets produced by the gearbox model of Chapter 4 and then to the bearing and gearbox experimental datasets. The simulation results showed that the EMD method manages to create signal components that could be related to damage, separating them from the varying load influences from the vibration signals for the specific loads examined in the simulations. These signal components are then available for further analysis with an amplitude-frequency separation technique, such as the HT or the TKEO/Desa-1. It was shown that the TKEO in combination with an energy separation algorithm can be a good alternative approach to

the HT under the condition of analysing monocomponent, clean and smooth signals. The simulation results also suggested that the number of IMFs could potentially separate the steady and time-carying load cases and the undamaged and damaged cases. This is not surprising, since it is generally accepted that damage increases the complexity of a structure (reference [155]), which in signal processing terms means that damage might create new frequency components. The EMD method, on the other hand for deterministic signals, or even signals that may contain a reasonable amount of noise, is known to decompose the signal analysed into meaningful signal components, the IMFs, so additional frequencies would mean additional IMFs. Apart from that, because gear tooth damage has an intermittency form in the signal, it initiates the mode mixing problem, therefore generating an even higher number of IMFs. In order to confirm this in experimental data, the experimental bearing datasets were used, giving results that agree with the previous assumptions. The noise levels in these datasets were considered rather low, since the tests were performed in a laboratory environment. This suggests that the EMD method did not act as a dyadic filter bank in this particular case, and therefore confirmed the initial assumption that had been made according to the simulation results. Unfortunately this could not be confirmed with the analysis of the gearbox experimental datasets, which were the only datasets coming from a wind turbine in operation, since all the datasets available were coming from damaged conditions of the gearbox. In addition, it is quite probable that due to the high levels of noise in this case, the EMD method would act as a dyadic filter bank and the mode mixing would not be initiated. Still, these datasets were used to perform time-frequency analysis and the $3D$ diagrams of their instantaneous characteristics were presented (both the Hilbert and the Teager spectra). Damage sensitive features were extracted from this analysis.

- In Chapter 7, feature discrimination using unsupervised learning is attempted at first by applying outlier analysis with Mahalanobis squared distance as a discordancy measure. Since, in this case time-frequency analysis is performed, this method should be applied in a different way than in previous studies that generally extract features from the FRFs or the transmissibility functions. So in this research work, outlier analysis is performed at the power of the 2^{nd} IMF for each gearbox dataset (that was obtained in the previous chapter), because this was the mode identified as the most sensitive to damage. In the power diagrams a 10-dimensional feature was defined as a 10-point window. One series

of 200 features was used for the training data, that were chosen carefully in order that no peaks (of the power) would be included in them. Whenever a fault appeared, the outlier statistic diagram should show a peak that is distinct from the normal condition data. The results of the method were satisfying and showed that the method could be used as a simple approach. A spatially adaptive thresholding method, known as $3D$ phase-space thresholding and appearing to be a novel method in the condition monitoring field was also tested on the same data. Because the peaks in the power signals, in this case, are the ones that show the existence of damage and since this thresholding technique was initiated particularly for spike detection, its application gave the desired results.

All the analysis, performed up to now can be considered to be a complete signal processing analysis for condition monitoring. A different approach to filter bank methods was also proposed in the next chapter.

- Cointegration is a technique which originates from econometrics, that linearly combines nonstationary response variables that are cointegrated in order to create a stationary residual whose stationarity represents the normal condition of the examined structure. For this reason, it has been recently applied in the structural health monitoring field as a means of removing environmental influences from data, and seems a promising approach in the condition monitoring field for the reasons described earlier. In Chapter 8 the possibility of cointegration using data from a single sensor is investigated through a simulated example. Apparently, in order to perform cointegration datasets from different sensors are needed, but this is not always feasible for the case of online monitoring for reasons of lack of storage and instrumentation. The idea proposed in order to achieve this, is based on the application of recursive AR-modelling (a recursive least squares algorithm with a forgetting factor) and is demonstrated on a 3DOF system synthetic data, that included sinusoidal stiffness variations. The correct choice of the forgetting factor helps the algorithm to track changes in the parameters. The results of the simulations produced AR coefficients that managed to track the stiffness sinusoidal variations simulated. In this way cointegration of the produced coefficients could be applied and tested. The results were promising, since the method managed to distinguish a simulated drop in the stiffness that represented damage.

  The second subject discussed in this chapter is the development of a multiresolution cointegration approach. The wavelets and the EMD methods are used for this purpose in combination with cointegration

and each approach offers different advantages and drawbacks. The datasets used in this analysis were collected by the DAMASCOS consortium, a Brite-Euram project which attempted to use Lamb waves for damage detection in a composite plate subjected to cyclic temperature variations. In a previous analysis [153], cointegration was used to create a feature that was insensitive to temperature-induced variation but still sensitive to damage.

The main objective of multiresolution cointegration has been to see if a multiresolution analysis can lead to enhancement of, the use of cointegration method for the removal of trends from structural health monitoring or condition monitoring data. Multiresolution cointegration was performed using the discrete wavelet analysis that was used as a means of characterising the time-scales on which nonstationarity manifested itself. The multiresolution results using the wavelet analysis were first presented in reference [151], and are given in order to have a better comparison of the two different approaches. One issue encountered in that paper was that the sparse wavelet basis at low frequencies created conditioning problems with the Johansen procedure that required solution by regularisation; the adoption of the EMD, as an alternative to the wavelets, solved this problem. Another problem with the multi-scale approach was the fact that there are excursions above the alarm limit for the residuals before damage occurs; this appears to manifest in the same manner with the EMD approach and is related to a type of Gibbs phenomenon. A truly on-line algorithm for detection will be able to solve this issue. A disadvantage of the EMD method on the other hand that is solved using the wavelet is that the number of IMFs produced might not be always the same.

In general, the results confirm that the multiresolution cointegration allows the removal of trends at independent time scales. Changes from environmental and operational variations will often show themselves on different time scales, so this possibility makes multiresolution cointegration interesting. In addition by cointegrating data at their most nonstationary timescale, one might potentially enhance the cointegration method, because when damage breaks the cointegration relation in this case, the nonstationarity effect reasserts itself most strongly.

At a deeper level, as discussed also in [151], the research has given support to the idea that cointegration and multiresolution analysis are fundamentally connected, although this assumption should be further explored.

Although the two methods proposed in the last chapter have been successful, the first one on synthetic data and the second one on structural health monitoring data, further work is needed to validate these results on experimental condition monitoring data.

9.1 Further work

Some aspects of the research presented in this study is of an exploratory character. Therefore, further investigations are required to fully establish certain parts of this thesis. These could include the following:

- This thesis focused on the condition monitoring of wind turbine gearboxes and to a lesser extend of wind turbine bearings. The kind of gearbox damage examined was limited to gear tooth damage, and as a result the kind of analysis that followed in order to extract appropriate damage features is defined by the way this kind of damage influences the gearbox vibration signals. Generally, the data coming from real wind turbine systems were limited and it would be interesting to try and test the methods proposed in a larger range of cases. Further work, that might be of interest would be for example to attempt applying similar signal processing methods to other components of the drivetrain, i.e. the shaft, that might need a completely different approach in the analysis in order to extract damage features.

- In the case of the EMD method, the lack of mathematical background is a restriction for the full understanding of the way the method works and also of the mode mixing problem of the EMD. Although, certain claims were made according to the simulation results that were confirmed by an experimental application, providing sufficient proof concerning the way gear tooth damage or bearing damage might influence the number of IMFs, under certain circumstances, should be performed in a much more theoretical background in order to make definite conclusions.

- In the case of the amplitude-frequency separation algorithms, the TKEO with the energy separation algorithm chosen, appeared to be a good alternative to the HT algorithm. Still, both algorithms have the disadvantage of being quite sensitive to not smooth data (or data containing noise). The solution examined here, was to use a smoothing filter (apart from the filter bank application that isolated noise only at the first IMFs). Testing other methods, such as the GZC method

for example, that might avoid a differentiation in order to estimate the frequency, would be interesting.

- For the feature discrimination part of the study the application of unsupervised learning methods seemed the most appropriate due to the lack of data from various damage cases and was generally considered successful. Still, this is just an initial step for damage detection, since novelty detection by itself lacks information concerning the kind and the exact level of damage. A supervised learning approach in this case would be able to distinguish for example whether a particular component in a drivetrain has reached a level of damage severity that suggests replacement. In order to achieve such a strategy though, either a huge range of data should be available or a high fidelity model should be created.

- Concerning the use of cointegration, most of the work presented is exploratory and has not been validated on real wind turbine gearbox data. In the first part of Chapter 8, a method was developed in order to perform cointegration with data obtained from just a single sensor. Although the simulation results in this case were successful, and the method is quite straightforward from a theoretical point of view, it has not been validated in real data and furthermore it has not been tested on either simulated or experimental gearbox data. For this purpose, an appropriate gearbox model should be simulated that can produce acceleration data from different placements of its cage and that allows load variations. In this way, cointegration using one sensor measurement can be compared with cointegration using various sensor measurements in a simulated environment. Finally, appropriate gearbox experimental data could validate the previous results.

- For the multiresolution cointegration approach using the wavelet transform an issue of sensitivity to the regularisation parameter (noise rms) adopted in the Johansen procedure should bear further investigation. In Chapter 8 a trial-and error approach was adopted with satisfying results but a more principled approach should be explored. This approach should be perhaps based on cross-validation on independent data. An analytical approach to regularisation might prove more powerful than adding noise. For the EMD alternative to multiresolution cointegration, this problem did not appear. Unfortunately, this case suffered from a different issue; the EMD method might not produce the same number of IMFs for the datasets analysed. In this case cointegration between components with slightly different time scales might

occur, possibly deteriorating the performance of the method. A solution to this, could probably be to use a noise assisted version of the EMD (e.g. the EEMD) that might act as a dyadic filter bank for all kinds of signals.

- One more improvement that could be made in the cointegration analysis would probably be to improve the detection thresholds for the novelty detection. The thresholds used in Chapter 8 are based on Gaussian statistics that might not be appropriate for every case examined. In fact they were almost certainly not for the lower wavelet levels were the statistics of the individual mother wavelet is likely to have greater influence.

- The multiresolution analysis presented insinuates that there might be a connection between cointegration and multiscale analysis, [151]. The idea behind cointegration is roughly that a number of signals sharing a common long-term trend can be combined, forming a signal from which this trend is removed. This combination of the signals which is constructed using the Johansen procedure, is a signal that has the stationarity characteristics of differenced versions of the original signals and differencing in this process makes the signals more stationary. As mentioned in [151], "the derivative operator - in spectral terms - is equivalent to a type of high-pass filter, so one can see how the sequence of detail/approximation divisions in the multi-resolution process will result in a sequence of signal components steadily changing their stationarity properties". It would be interesting to attempt to investigate this connection between cointegration and multi-resolution analysis at a deeper mathematical level in the future.

- Finally the algorithms used throughout the thesis are based on an offline approach to condition monitoring. On a practical level it would be useful to develop their on-line versions.

Publications

Author publications to date

Journal papers

1. **I. Antoniadou**, G. Manson, W. J. Staszewski, T. Barszcz and K. Worden, *A time-frequency analysis approach for condition monitoring of a wind turbine gearbox under varying load conditions*, Submitted to Mechanical Systems and Signal Processing.
2. K. Worden, E. J. Cross, **I. Antoniadou** and A. Kyprianou, *A multiresolution approach to cointegration for analysis of time-varying systems and structures*, Submitted to Mechanical Systems and Signal Processing.
3. A. Ziaja, **I. Antoniadou**, T. Barszcz, W. J. Staszewski and K. Worden, *Fault detection in rolling element bearings using wavelet-based variance analysis and novelty detection*, Submitted to Journal of Vibration and Control.

Reviewed conference papers

4. **I. Antoniadou**, T. P. Howard, R. S. Dwyer-Joyce, M. B. Marshall, J. Naumann, N. Dervilis and K. Worden, *Envelope analysis using the Teager-Kaiser energy operator for condition monitoring of a wind turbine bearing*, International Conference on Advances in Mechanical and Manufacturing Engineering, Kuala Lumpur, November 2013.

5. N. Dervilis, A. C. W. Creech, A. E. Maguire, **I. Antoniadou**, R. J. Barthorpe and K. Worden, *An SHM view of a CFD model of Lillgrund wind farm*, International Conference on Advances in Mechanical and Manufacturing Engineering, Kuala Lumpur, November 2013.

6. **I. Antoniadou**, E. J. Cross and K. Worden, *Cointegration and the Empirical Mode Decomposition for the analysis of diagnostic data*, 10^{th} International Conference on Damage Assessment of Structures, Dublin, July 2013.

7. **I. Antoniadou**, N. Dervilis, R. J. Barthorpe, G. Manson and K. Worden, *Advanced tools for damage detection in wind turbines*, 10^{th} International Conference on Damage Assessment of Structures, Dublin, July 2013.

8. **I. Antoniadou**, G. Manson, S. G. Taylor, N. Dervilis, K. Worden and C. R. Farrar, *Damage detection of RAPTOR Telescope systems using time-frequency analysis methods*, 5^{th} International Congress on Technical diagnostics, Krakow, September 2012.

9. N. Dervilis, M. Choi, **I. Antoniadou**, K. M. Farinholt, S. G. Taylor, R. Barthorpe, G. Park, C. R. Farrar and K. Worden, *Machine Learning applications to CX-100 wind turbine blade under continuous fatigue loading*, 5^{th} International Congress on Technical diagnostics, Krakow, September 2012.

10. N. Dervilis, M. Choi, **I. Antoniadou**, K. M. Farinholt, S. G. Taylor, R. Barthorpe, G. Park, K. Worden and C. R. Farrar, *Novelty detection applied to vibration data from CX-100 wind turbine blade under fatigue loading*, Modern Practice in Stress and Vibration Analysis, Glasgow, August 2012.

11. **I. Antoniadou**, G. Manson, W. J. Staszewski and K. Worden, *Damage Detection in Gearboxes considering intermittent faults and transient loads*, 2^{nd} International Conference on Smart Diagnostics of Structures, Krakow, November 2011.

Conference papers

12. **I. Antoniadou**, E. J. Cross and K. Worden, *Cointegration for the removal of environmental and operational effects using a single sensor*, 9^{th} International Workshop on Structural Health Monitoring, Stanford, September 2013.

13. **I. Antoniadou**, G. Manson, N. Dervilis, T. Barszcz, W. J. Staszewski and K. Worden, *Instantaneous characteristics estimation using the Teager-Kaiser operator for the condition monitoring of a wind turbine gearbox*, International Conference on Noise and Vibration Engineering, Leuven, September 2012.

14. **I. Antoniadou**, G. Manson, N. Dervilis, T. Barszcz, W. J. Staszewski and K. Worden: *Condition monitoring of a wind turbine gearbox using Empirical Mode Decomposition and outlier analysis*, 6^{th} European workshop on Structural Health Monitoring, Dresden, July 2012.

15. N. Dervilis, R. Barthorpe, **I. Antoniadou** and K. Worden, *Impact damage detection for composite material typical of wind turbine blades using novelty detection*, 6^{th} European Workshop on Structural Health Monitoring, Dresden, July 2012.

16. **I. Antoniadou**, G. Manson, N. Dervilis, W. J. Staszewski and K. Worden,*Damage detection in wind turbine gearboxes using outlier analysis*, SPIE Smart Structures/NDE, Industrial and Commercial Applications of Smart Structures Technologies, San Diego, March 2012.

17. N. Dervilis, R. Barthorpe, W. J. Staszewski, **I. Antoniadou** and K. Worden, *Damage detection in carbon composite material typical of wind turbine blades using auto-associative neural networks*, SPIE Smart Structures/NDE, Industrial and Commercial Applications of Smart Structures Technologies, San Diego, March 2012.

18. **I. Antoniadou**, W. J. Staszewski and K. Worden, *Intermittent fault detection in gearboxes - a simulated case study*, 18^{th} International Congress on Sound and Vibration, Rio de Janeiro, July 2011.

Nomenclature

ADF Augmented Dickey-Fuller test

ADV Acoustic Dopler Velocimeter

AIC Akaike Information Criterion

AM Amplitude Modulation

ANC Adaptive Noise Cancellation

$ANNs$ Aritificial Neural Networks

AR Auto Regressive

ARX Auto Regressive with eXogenous inputs

$BEMD$ Bivariate Empirical Mode Decomposition

$BPFI$ Ballpass Frequency Inner race

$BPFO$ Ballpass Frequency Outer race

BSF Ball Spin Frequency

BSS Blind Source Separation

CDA Canonical Discriminant Analysis

$CEEMD$ Complementary Ensemble Empirical Mode Decomposition

$CFRP$ Carbon Fibre-Reinforced Plastic

CM Condition Monitoring

CWD Choi-Williams Distribution

$DDMA$ Discriminant Diffusion Map Analysis

$Desa-1$ Discrete energy separation algorithm 1

DF Dickey-Fuller

DOF Degree of Freedom

DQ Direct Quadrature

EAs Evolutionary Algorithms

EDM Electrical Discharge Machining

$EEMD$ Ensemble Empirical Mode Decomposition

EMD Empirical mode Decomposition

FFT Fast Fourier Transform

FL Fuzzy Logic

FM Frequency Modulation

FNN Feedforward Neural Network

FRF Frequency Response Function

FT Fourier Transform

FTF Fundamental Train Frequency

GRP Glass Reinforced Plastic

$GWEC$ Global Wind Energy Council

GZC Generalised Zero Crossing

HMM Hidden Markov Models

HT Hilbert Transform

$i.i.d.$ independent and identically distributed signals

IA Instantaneous Amplitude

ICA Independent Component Analysis

IF Instantaneous Frequency

$IMFs$ Intrinsic Mode Functions

LDA Linear Discriminant Analysis

$MIMO$ Multiple input and multiple output system

MLP Multi-Layer Perceptron

non i.i.d. non independent and identically distributed signals

$NREL$ National Renewable Energy Laboratory

$NWIP$ Northern Wind Innovation Programme

OA Outlier Analysis

OWT Orthogonal Wavelet Transform

PCA Principal Component Analysis

PSD Power Spectral Density

RLS Recursive Least Squares

$SANC$ Self-Adaptive Noise Cancellation

SHM Structural Health Monitoring

SOM Self-Organising Maps

ST Synchrosqueezing Transform

STE Static Transmission Error

$STFT$ Short Time Fourier Transform

SVM Support Vector Machines

TDA Time Domain Averaging

$TKEO$ Teager-Kaiser Energy Operator

VAR Vector Autoregressive

$VECM$ Vector Error-Correction Model

WT Wavelet Transform

WVD Wigner-Ville Distribution

Appendix A

Signal processing

A.1 Fourier analysis

The FT rules over linear time-invariant signal processing, [156]. The theory presented here is largely based on the Fourier analysis chapter (Chapter 4) of the book given in reference [157].

If a signal $x(t)$ is a periodic function of time t, with period T,as shown in Figure A.1 then this signal can be expressed as an infinite trigonometric series, (*Fourier series*) of the form, [157]:

$$x(t) = a_0 + \sum_{k=1}^{\infty}(a_k cos\frac{2\pi kt}{T} + b_k sin\frac{2\pi kt}{T}) \tag{A.1}$$

where a_0, a_k, b_k are constant Fourier coefficients given by the equations:

$$a_0 = \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)dt \tag{A.2}$$

$$a_k = \frac{2}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)cos\frac{2\pi kt}{T}dt \tag{A.3}$$

$$b_k = \frac{2}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)sin\frac{2\pi kt}{T}dt \tag{A.4}$$

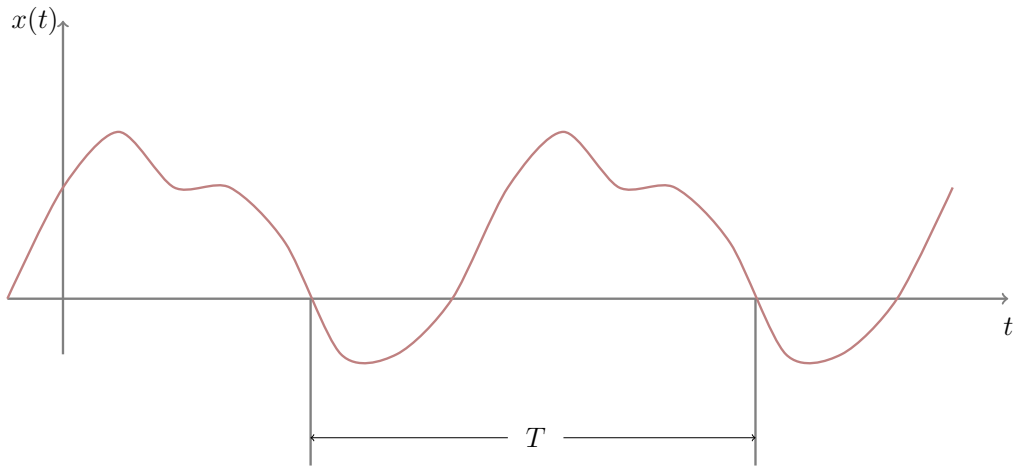


Figure A.1: A periodic signal [157].

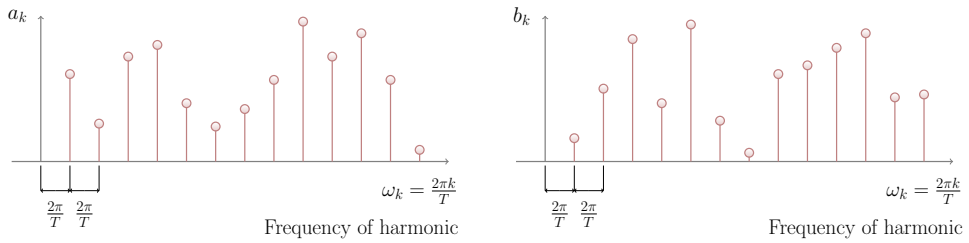


Figure A.2: Fourier coefficients [157].

The important restriction of the equation A.1 is that, when $x(t)$ is discontinuous, the series gives the average value of $x(t)$ at the discontinuity.

If the position of the t axis in Figure A.1 is adjusted so that the mean value of the signal is zero, then because of equation A.2, the coefficient a_0 will be zero and the coefficients a_k and b_k will generally be all different.

An illustration of their values is shown in Figure A.2.

The horizontal axis in this figure represents frequency. So the k_{th} coefficient is located at the frequency

$$\omega_k = \frac{2\pi k}{T} \tag{A.5}$$

which is the frequency of the k_{th} harmonic. When $T \to \infty$, the signal no longer represents a periodic phenomenon and can no longer be analysed into discrete Fourier components, subject to certain conditions though, the Fourier series of equation A.1 turns into a Fourier integral and the Fourier

coefficients of equation A.2 turn into continuous functions of frequency called Fourier transforms.

The Fourier integral is given in the following equation:

$$x(t) = 2\int_0^\infty A(\omega)\cos\omega t d\omega + 2\int_0^\infty B(\omega)\sin\omega t d\omega \tag{A.6}$$

with:

$$A(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(t)\cos\omega t dt \tag{A.7}$$

$$B(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(t)\sin\omega t dt \tag{A.8}$$

with the condition:

$$\int_{-\infty}^{\infty} |x(t)| dt < \infty \tag{A.9}$$

which means that the classical Fourier analysis theory can be applied only to functions that decay to zero when $|t| \to \infty$.

Making use of Euler's formula:

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{A.10}$$

$X(\omega)$ can be defined as:

$$X(\omega) = A(\omega) - iB(\omega) \tag{A.11}$$

one can derive the complex form of the continuous-time FT of the x(t):

$$X(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(t)(cos\omega t - isin\omega t)dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(t)e^{-i\omega t}dt \tag{A.12}$$

and the signal $x(t)$ can be regained from $X(\omega)$ by the inverse FT:

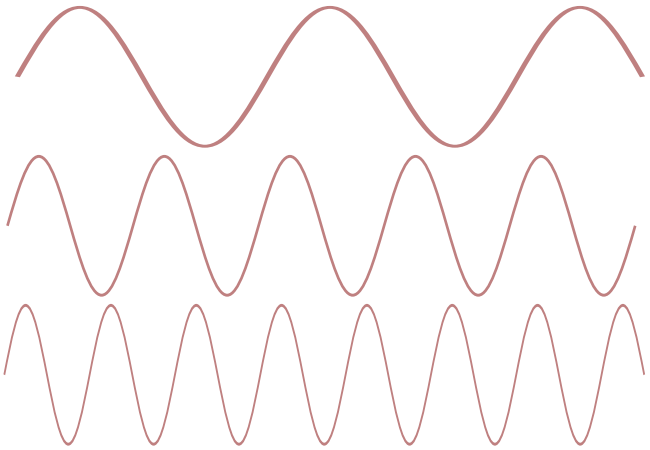


Figure A.3: Fourier basis functions (real parts of complex sinusoidal functions) [126].

$$x(t) = \int_{-\infty}^{\infty} X(\omega) e^{i\omega t} d\omega \tag{A.13}$$

$X(\omega)$ is the measure of the similarity between the signal $x(t)$ and complex sinusoidal functions [126] (Figure A.3). One important feature of the FT is that the basis functions in equation (A.13) and the dual functions in (A.12) have the same form. In addition, $e^{i\omega t}$ corresponds to an impulse at frequency ω. This means that the $X(\omega)$ is the signal's projections on the basis functions and precisely reflects the signal's behaviour at frequency ω. It is obvious from the above that the $X(\omega)$ represents the spectral content of the signal $x(t)$ in terms of frequency ω and magnitude $X(\omega)$. The square of the FT $|X(\omega)|^2$ is called *power spectrum*, and is a quantity that indicates how the signal energy is distributed in the frequency domain.

A.1.1 Frequency Response Functions

The equation of motion of a single-degree-of-freedom system [8] (Figure A.4) is:

$$m\ddot{y} + c\dot{y} + ky = x(t) \tag{A.14}$$

The above equation is written in the frequency domain (FT):

Figure A.4: Single-input-single-output SDOF system [8].

$$m(i\omega)Y(\omega) + c(i\omega)Y(\omega) + kY(\omega) = X(\omega) \tag{A.15}$$

or

$$(-m\omega^2 + ic\omega + k)Y(\omega) = X(\omega) \tag{A.16}$$

So if one defines

$$H(\omega) = \frac{1}{-m\omega^2 + ic\omega + k} \tag{A.17}$$

one finally obtains

$$Y(\omega) = H(\omega)X(\omega) \tag{A.18}$$

It is very easy to define the *frequency response function* (FRF) from the equation (A.18).

It is worth to point out that if:

$$h(t) = F^{-1}[H(\omega)] \tag{A.19}$$

then $h(t)$ is the solution of

$$m\ddot{h} + c\dot{h} + kh = \delta(t) \tag{A.20}$$

and for this reason it is called the *impulse response*.

A.2 The Short-Time Fourier Transform

Because the basis functions used in the classical Fourier analysis do not associate with any particular time instant, the resulting measurements (FT), do not reflect a signal's time-varying nature [126].

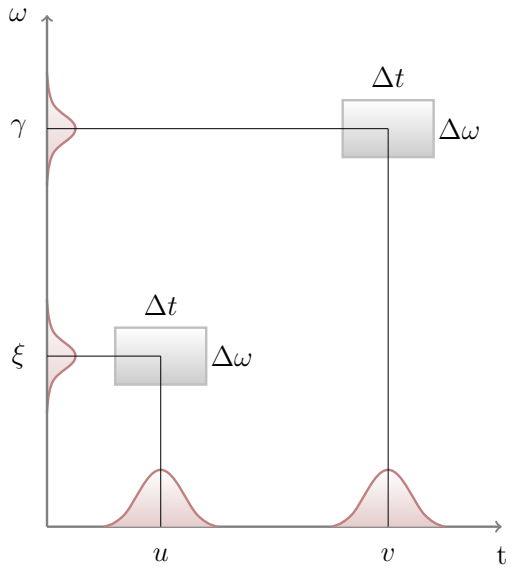


Figure A.5: Time-frequency boxes ("Heisenberg rectangles") representing the energy spread of two Gabor atoms [156].

The simplest way to overcome this is to compare the signal with elementary functions that are localised in the time and frequency domains simultaneously:

$$\mathbf{STFT}\{x(t)\}(t,\omega) = \int_{-\infty}^{\infty} x(\tau)\gamma^*(\tau - t)e^{-i\omega\tau}d\tau \tag{A.21}$$

which is a regular inner product and reflects the similarity between a signal $x(t)$ and the elementary function $\gamma(\tau - t)e^{-i\omega t}$. This function has a short time duration and is named the *window function*. Equation A.21 is the well known STFT, or windowed transform, defined by Gabor, [131]. One can write equation (A.21) on a frequency interval as well by applying the Fourier Parseval formula.

Gabor atoms are constructed by translating in time and frequency the time window γ:

$$\gamma_{u,\xi}(T) = \gamma(t - u)e^{i\xi t} \tag{A.22}$$

Figure A.5 shows the energy spread of two Gabor atoms [156].

One can explain the STFT by the following: the short time duration of the window function reflects the signal's local frequency properties. This is because, the window function is first multiplied with the signal and the FT of the product $x(\tau)\gamma^*(\tau - t)$ is computed.

One could also understand the STFT with the following concept: in STFT the signal is compared with a set of elementary functions $\gamma(\tau - t)e^{-i\omega t}$, concentrated in both time and frequency domains. If the time duration and frequency bandwidth of the window function are Δt and $\Delta\omega$ the $\mathbf{STFT}(t, \omega)$ indicates the signal's behaviour in the vicinity of $[t-\Delta t, t+\Delta t] \times [\omega-\Delta\omega, \omega+\Delta\omega]$. The values of Δt and $\Delta\omega$ should be as small as possible in order to have a better measurement of the signal's properties at a particular time and frequency (t, ω). Unfortunately, Heisenberg's uncertainty principle creates a limitation to that since Δt and $\Delta\omega$ are not independent:

$$\Delta t \Delta\omega \geq \frac{1}{2} \tag{A.23}$$

The above equation means that if $\gamma(t)$ is chosen to have good time resolution (Δt small), then its frequency resolution must be deteriorated and *vice versa*. The equality of the equation A.23 only holds when $\gamma(t)$ is a Gaussian function.

Fig. A.6 shows the STFT elementary functions. Unlike the FT, for which the elementary functions extend into the entire time domain, the Gabor elementary functions are centred in a particular time, and therefore they contain both time and frequency information. Both the FT and the STFT elementary functions though are harmonically-related complex sinusoidal functions.

The square of the STFT is called the STFT spectrogram and is the simplest time-dependent spectrum depicting roughly a signal's energy distribution in the joint time-frequency domain.

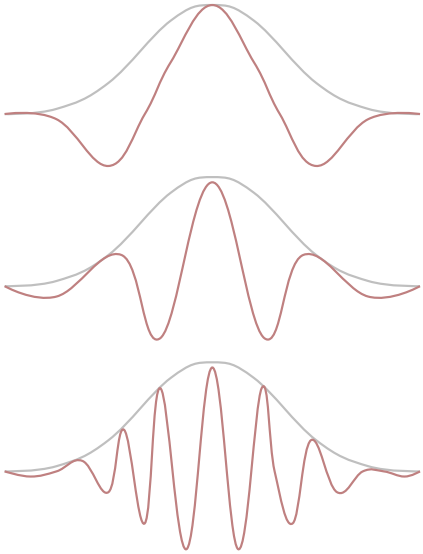


Figure A.6: Gabor elementary functions. In the Gabor transform, the change of the frequencies is obtained by frequency modulation [126].

A.3 Wavelets

The *continuous wavelet transform* (CWT), uses dilated (scaled) and translated (time-shifted) elementary functions, $\psi(a^{-1}(t-b))$ in order to produce the *time-scale representation* of the signal analysed. The elementary function $\psi(t)$ is known as *mother wavelet*. If $\psi(t)$ is centred at time zero and frequency ω_0, then the function $\psi(a^{-1}(t-b))$ is centred at time b and frequency ω_0/a. So b is a translation parameter indicating the time locality and $a(a>0)$ is a dilation or scale parameter and what the CWT does is to scale the time variable t of a given elementary function [126]. The equation describing this transform is the following:

$$\mathbf{CWT}(a,b) = \frac{1}{\sqrt{|a|}} \int x(t)\psi^*(\frac{t-b}{a})dt \tag{A.24}$$

for $a \neq 0$.

An important characteristic of the WT when compared to STFT is that although the time and frequency resolutions of STFT are uniform in the entire time-frequency domain, they vary in WT. Figures A.7 and A.8 illustrate this difference.

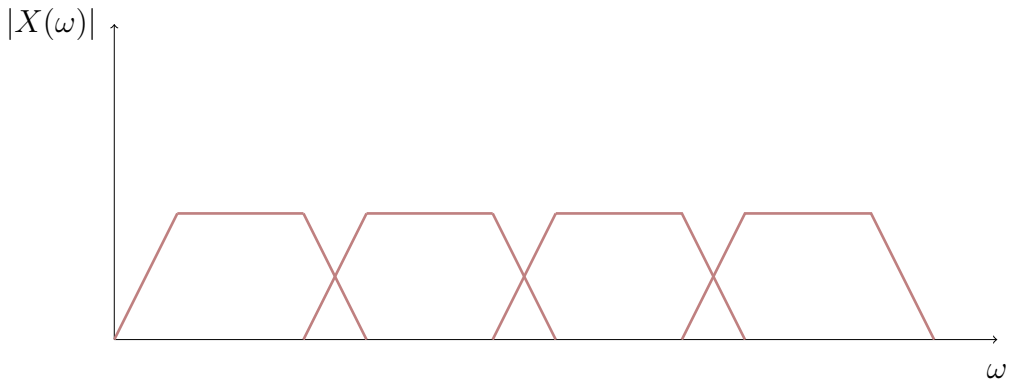


Figure A.7: Bandwidth of STFT is uniform in frequency domain [126].

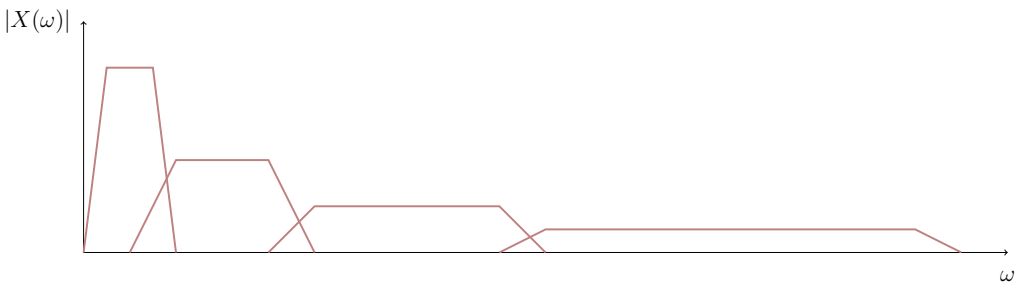


Figure A.8: Frequency response of WT [126].

If the data of interest are sampled, the *discrete wavelet transform* (DWT) can be used, that results from setting the dilations and translations to be $a_j = 2^j$ and $b_{j,k} = \frac{k}{2^j}$.

In this case the OWT can be defined by [151]:

$$x_k^m = \int_{-\infty}^{\infty} x(t)\psi_{m,k}(t)dt \tag{A.25}$$

where the function $\psi(t)$ is from a so-called orthogonal wavelet family,

$$\psi_{m,k}(t) = 2^{\frac{m}{2}}\psi(2^m t - k), (m, k\epsilon Z) \tag{A.26}$$

The transform is simply a linear combination of the basis functions and the decomposed function can be represented as a sum of m wavelet levels:

$$x_m(t) = \sum_k x_k^m \psi_{m,k}(t) \tag{A.27}$$

These levels represent the time behaviour of the signal within different scale bands and give their contribution to the total signal energy. Each level is determined by a number of wavelets which can be fitted into the interval of interest. If for example the Daubechies‘ wavelets are used, the level m is composed of $2^{\frac{m}{2}}$ basis functions. This decomposition corresponds to the division into levels of detail and approximation (Figure A.9), so it can be seen as the application of a sequence of filtering operations.

Starting from the original signal, one applies both high and low pass filtering [151]. The outcome of high pass filtering is designated as detail and that of low pass as approximation. This is repeated on the approximation signal to yield a second level of detail and new approximation. The process can be repeated until the low-frequency approximation is reduced to monotonic or even constant behaviour.

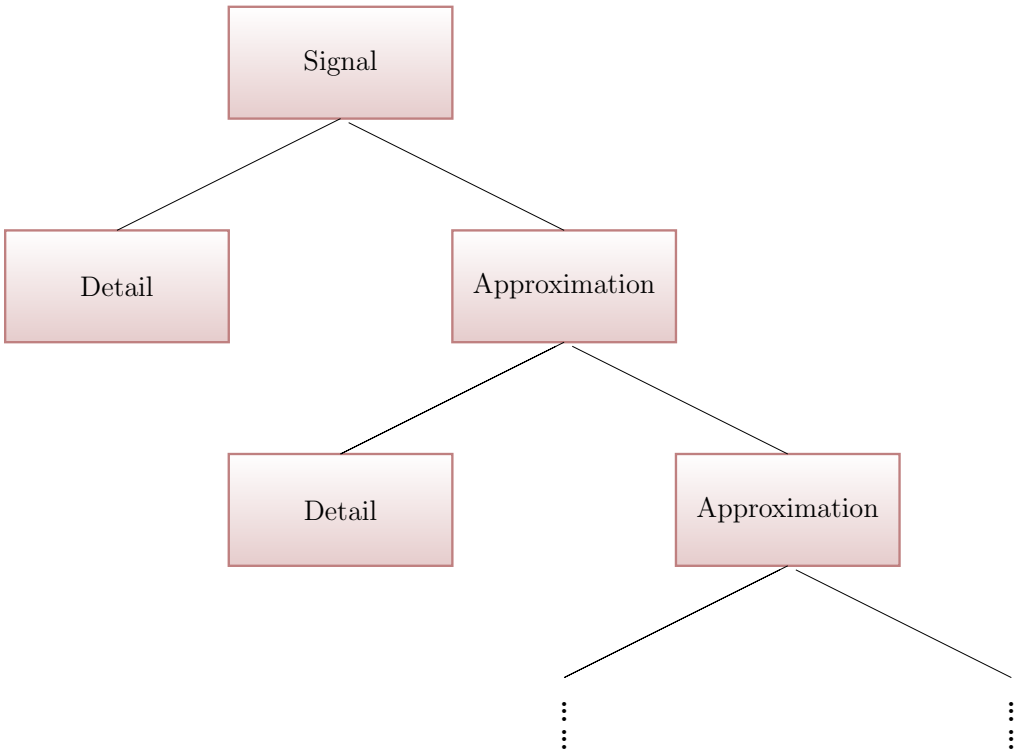


Figure A.9: A schematic for multi-resolution analysis [151].

Bibliography

[1] Global wind statistics. *Global Wind Energy Council*, 2012.

[2] Renewable UK, the voice of wind and marine energy. 2012.

[3] W.Q. Meeker and Y. Hong. Reliability meets big data: Opportunities and challengies. *Department of Statistics, Iwoa State University*, 2013.

[4] W. Musial, S. Butterfield, and B. McNiff. Improving wind turbine gearbox reliability. *European Wind Energy Conference*, 2007.

[5] K. Worden and J.M. Dulieu-Barton. An overview of intelligent fault detection in systems and structures. *Structural Health Monitoring*, 3: 85–98, 2004.

[6] R.B. Randall. Vibration-based condition monitoring: Industrial, Aerospace and Automotive applications. *Wiley*, 1993.

[7] A. Rytter. Vibration based inspection of civil engineering structures, Ph.D. thesis. *Department of Building Technology and Structural Engineering, University of Aalborg, Denmark*, 1993.

[8] C.R. Farrar and K. Worden. Structural health monitoring: A machine learning perspective. *John Wiley and Sons Ltd*, 2013.

[9] C.R. Farrar and K. Worden. Vibration-based structural identification. *Philosophical Transactions of the Royal Society London: A*, 359(1778): 131 – 149, 2013.

[10] W.J. Staszewski. The application of time-variant analysis to gearbox fault detection. *Ph.D. Thesis, Department of Engineering, University of Manchester, United Kingdom*, 1994.

[11] Chang F.K. Boller, C. and Y. Fujino. Encyclopedia of structural health monitoring. *Wiley-Blackwell*, 2009.

[12] A.K.S. Jardine, D. Lin, and D. Banjevic. A review of machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20:1483 – 1510, 2006.

[13] M. Nie and L. Wang. Review of condition monitoring and fault diagnosis technologies for wind turbine gearbox. *Procedia CIRP, Elsevier*, 11:287–290, 2013.

[14] C. Pachaud, R. Salvetat, and C. Fray. Improvement of the sensitivity of the scalar indicators (crest factor, kurtosis) using a denoising method by spectral subtraction: application to the detectionof defects in ball bearings. *Journal of Sound and Vibration*, 270:61 – 73, 2004.

[15] C. Pachaud, R. Salvetat, and C. Fray. Crest factor and kurtosis contributions to identify defects inducing periodical impulsive forces. *Mechanical Systems and Signal Processing*, 11:903 – 916, 1997.

[16] R.A. Collacott. Mechanical fault diagnosis and condition monitoring. *Chapman and Hall*, 1977.

[17] T. Barszcz and A. Jablonski. A novel method for the optimal band selection for vibration signal demodulation and comparison with the kurtogram. *Mechanical Systems and Signal Processing*, 25:431 – 451, 2011.

[18] P.D. McFadden. Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration. *Transactions of the ASME, Journal of Vib. Acoustics*, 108:165 – 170, 1986.

[19] D. Brie, M. Tomczak, H. Oehlmann, and A. Richard. Gear crack detection by adaptive amplitude and phase demodulation. *Mechanical Systems and Signal Processing*, 11:149 – 167, 1997.

[20] P.D. McFadden and J.D. Smith. Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of Sound and Vibration*, 96:69 – 82, 1984.

[21] H. Sohn and C.R. Farrar. Damage diagnosis using time series analysis of vibration signals. *Smart Materials and Structures*, 10:446 – 451, 2001.

[22] P.D. McFadden. Examination of a technique for the early detection of failure in gears by signal processing of the time domain average of the

meshing vibration. *Mechanical systems and signal processing*, 1:173 – 183, 1987.

[23] P.D. McFadden. Interpolation techniques for time domain averaging of gear vibration. *Mechanical systems and signal processing*, 3:87 – 97, 1989.

[24] R.B. Randall. A new method of modelling gear faults. *Trans. ASME, Journal of Mechanical Design*, 104:259 – 267, 1982.

[25] N. Sawalhi and R.B. Randall. Simulating gear and bearing interactions in the presence of faults: Part I. The combined gear bearing dynamic model and the simulation of localised bearing faults. *Mechanical Systems and Signal Processing*, 22:1924 – 1951, 2008.

[26] B. Bogert, M. Healy, and J. Tukey. The quefrency alanysis of time series for echoes: Cepstrum, Pseudo-Autocovariance, Cross-Cepstrum and Saphe Cracking. In *Proc. Symp. on Time Series Analysis*, pages 209 – 243, 1963.

[27] M. Inalpolat and A. Kahraman. A theoretical and experimental investigation of modulation sidebands of planetary gear sets. *Journal of Sound and Vibration*, 323(3-5):677–696, 2009.

[28] M. Inalpolat and A. Kahraman. A dynamic model to predict modulation sidebands of a planetary gear set having manufacturing errors. *Journal of Sound and Vibration*, 329(4):371–393, 2010.

[29] M. El Badaoui, F. Guillet, and J. Danière. New applications of the real cepstrum to gear signals, including definition of a robust fault indicator. *Mechanical Systems and Signal Processing*, 18(5):1031–1046, 2004.

[30] K. Worden and G.R. Tomlinson. Nonlinearity in structural dynamics: Detection, Identification and Modelling. *Taylor and Francis*, 2000.

[31] S. Braun and J.K. Hammond. Parametric methods. *Mechanical signature analysis. Academic Press London*, 1986.

[32] J. Antoni and R.B. Randall. Unsupervised noise cancellation for vibration signals: Part i. Evaluation of adaptive algorithms. *Mechanical Systems and Signal Processing*, 18:89 – 101, 2004.

[33] J. Antoni and R.B. Randall. Unsupervised noise cancellation for vibration signals: Part ii. A novel frequency-domain algorithm. *Mechanical Systems and Signal Processing*, 18:103 – 117, 2004.

[34] Z.K. Peng and F.L. Chu. Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography. *Mechanical Systems and Signal Processing*, 18(2):199–221, 2004.

[35] W.J. Wang and P.D. McFadden. Early detection of gear failure by vibration analysis: Part I. Calculation of the time- frequency distribution. *Mechanical Systems and Signal Processing*, 7:193 – 203, 1993.

[36] W.J. Staszewski and G.R. Tomlinson. Application of the wavelet transform to fault detection in a spur gear. *Mechanical Systems and Signal Processing*, 8(3):289–307, 1994.

[37] N.G. Nikolaou and I.A. Antoniadis. Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted morlet wavelets. *Mechanical Systems and Signal Processing*, 16:677 – 694, 2002.

[38] W.J Staszewski and K. Worden. Wavelet analysis of time-series: coherent structures, chaos and noise. *International Journal of Bifurcation and Chaos*, 9:455 – 471, 1999.

[39] A. Ziaja, T. Barszcz, and W.J. Staszewski. Fractal based signal processing for fault detection in ball-bearings. In 2^{nd} *International Conference on Condition Monitoring of Machinery in Non-Stationary Operations, Tunisia*, pages 385–392, 2012.

[40] S. Qian and D. Chen. In *Joint time-frequency analysis: Methods and Applications*. NJ: Prentice-Hall, 1996.

[41] W.J. Staszewski, K. Worden, and G.R. Tomlinson. Time-frequency analysis in gearbox fault detection using the Wigner-Ville distribution and pattern recognition. *Mechanical Systems and Signal Processing*, 11(5):673–692, 1997.

[42] N. Baydar and A. Ball. A comparative study of acoustic and vibration signals in detection of gear failures using Wigner-Ville distribution. *Mechanical Systems and Signal Processing*, 15(6):1091 – 1107, 2001.

[43] B. Tang, W. Liu, and T. Song. Wind turbine fault diagnosis based on morlet wavelet transformation and Wigner-Ville distribution. *Renewable Energy*, 35(12):2862 – 2866, 2010.

[44] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu. The empirical mode decomposition

and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, 1998.

[45] R. Ricci and P. Pennacchi. Diagnostics of gear faults based on EMD and automatic selection of intrinsic mode functions. *Mechanical Systems and Signal Processing*, 25(3):821 – 838, 2011.

[46] A. Parey, M. El Badaoui, F. Guillet, and N. Tandon. Dynamic modelling of spur gear pair and application of empirical mode decomposition-based statistical analysis for early detection of localized tooth defect. *Journal of Sound and Vibration*, 294(3):547 – 561, 2006.

[47] Y. Lei, J. Lin, Z. He, and M.J. Zuo. A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, 35(12):108 – 126, 2013.

[48] P. Flandrin, G. Rilling, and P. Goncalves. Empirical mode decomposition as a filter bank. *Signal Processing Letters, IEEE*, 11(2):112 – 114, feb. 2004.

[49] Z. Wu and N.E. Huang. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1:1 – 49, 2009.

[50] Z. Wu and N.E. Huang. A study of the characteristics of white noise using the empirical mode decomposition method. *Proceedings of the Royal Society A, London*, 460:1597 1611, 2004.

[51] J.R. Yeh, J.S. Shieh, and N.E. Huang. On instantaneous frequency. *Advances in Adaptive Data Analysis*, 2(2):135 – 136, 2010.

[52] G. Rilling, P. Flandrin, P. Goncalves, and J.M. Lilly. Bivariate empirical mode decomposition. *Signal Processing Letters, IEEE*, 14(12):936 – 939, 2007.

[53] Q. Du and S. Yang. Application of the EMD method in the vibration analysis of ball bearings. *Mechanical Systems and Signal Processing*, 21(6):2634 – 2644, 2007.

[54] Y. Yang, D.J. Yu, and J.S. Cheng. A roller bearing fault diagnosis method based on EMD energy entropy and ANN. *Journal of Sound and Vibration*, 294(1-2):269 – 277, 2006.

[55] W. Yang and P.J Tavner. Empirical mode decomposition, an adaptive approach for interpreting shaft vibratory signals of large rotating machinery. *Journal of Sound and Vibration*, 321(35):1144 – 1170, 2009.

[56] G. Gai. The processing of rotor startup signals based on empirical mode decomposition. *Mechanical Systems and Signal Processing*, 20 (1):222 – 235, 2006.

[57] T.H. Patel and A.K. Darpe. Coupled bending-torsional vibration analysis of rotor with rub and crack. *Journal of Sound and Vibration*, 326 (35):740 – 752, 2009.

[58] S.J. Loutridis. Instantaneous energy density as a feature for gear fault detection. *Mechanical Systems and Signal Processing*, 20:1239 1253, 2006.

[59] S.J. Loutridis. Damage detection in gear systems using empirical mode decomposition. *Engineering Structures*, 26:1833 1841, 2004.

[60] C. Li and M. Liang. Time-frequency signal analysis for gearbox fault diagnosis using a generalized synchrosqueezing transform. *Mechanical Systems and Signal Processing*, 26:205 – 217, 2012.

[61] J. Antoni. Cyclic spectral analysis in practice. *Mechanical Systems and Signal Processing*, 21(2):597 – 630, 2007.

[62] J. Antoni. Cyclostationary by examples. *Mechanical Systems and Signal Processing*, 23(4):987 – 1036, 2009.

[63] J. Antoni and R.B. Randall. Differential diagnosis of gear and bearing faults. *ASME Journal of Vibration and Acoustics*, 124:165 – 171, 2002.

[64] K. Worden, W.J. Staszewski, and J.J. Hensman. Natural computing for mechanical systems research: A tutorial overview. *Mechanical Systems and Signal Processing*, 25:4 – 111, 2011.

[65] Y. Huang, X.F. Zha, J. Lee, and C. Liu. Discriminant diffusion maps analysis: A robust manifold learner for dimensionality reduction and its applications in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*, 34(12):277 – 297, 2013.

[66] W.J. Staszewski and K. Worden. Classification of faults in gearboxes-pre-processing algorithms and neural networks. *Neural Computing and Applications*, 5(3):160 – 183, 1997.

[67] J. Rafiee, F. Arvani, A. Harifi, and M.H. Sadeghi. Intelligent condition

monitoring of a gearbox using artificial neural network. *Mechanical Systems and Signal Processing*, 21(4):1746 – 1754, 2007.

[68] A. Witodo and B.S. Yang. Support Vector Machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*, 21(6):2560 – 2574, 2007.

[69] P. Hayton, S. Utete, D. King, S. King, P. Anuzis, and L. Tarassenko. Static and dynamic novelty detection methods for jet engine health monitoring. *Philosophical Transactions of the Royal Society: Series A*, 365:493 – 514, 2007.

[70] D.M Yang, A.F. Stronach, P. Macconnell, and J. Penman. Third-order spectral techniques for the diagnosis of motor bearing condition using artificial neural networks. *Mechanical Systems and Signal Processing*, 16(23):391 – 411, 2002.

[71] B.A. Paya, I.I. Esat, and M.N.M. Badi. Aritificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor. *Mechanical Systems and Signal Processing*, 11(5):751 – 765, 1997.

[72] B. Sreejith, A.K. Verma, and A. Srividya. Fault diagnosis of rolling element bearing using time-domain features and neural networks. In *IEEE Region 10 and the Third international Conference on Industrial and Information Systems*, pages 1–6, 2008.

[73] C.K. Mechefske. Objective machinery fault diagnostics using fuzzy logic. *Mechanical Systems and Signal Processing*, 12(6):855 – 862, 1998.

[74] S. Sampath, S. Ogaji, R. Singh, and D. Probert. Engine-fault diagnostics: An optimisation procedure. *Applied Energy*, 73(1):47 – 70, 2002.

[75] Z. Chen, Y. He, F. Chu, and J. Huang. Evolutionary strategy for classification problems and its application in fault diagnostics. *Engineering Applications of Artificial Intelligence*, 16(1):31 – 38, 2003.

[76] Y.C. Huang and C.M. Huang. Evolving wavelet networks for power transformer condition monitoring. *IEEE Transactions on Power Delivery*, 17(2):412 – 416, 2002.

[77] J. Ma and J.C. Li. Detection of localised defects in rolling element bearings via composite hypothesis test. *Mechanical Systems and Signal Processing*, 9(1):63 – 75, 1995.

[78] H. Sohn, K. Worden, and C.R. Farrar. Statistical damage classification under changing environmental and operational conditions. *Journal of Intelligent Material Systems and Structures*, 13:561 574, 2002.

[79] M.L. Fugate, H. Sohn, and C.R. Farrar. Vibration based damage detection using statistical process control. *Mechanical Systems and Signal Processing*, 15(4):707 – 721, 2001.

[80] W.J. Staszewski, K. Worden, and G.R. Tomlinson. Time-frequency analysis in gearbox fault detection using the Wigner-Ville distribution and pattern recognition. *Mechanical Systems and Signal Processing*, 11(5):673 – 692, 1997.

[81] J.L. Goldwin and P. Matthews. Prognosis of wind turbine gearbox failures by utilising robust multivariate statistical techniques. In *IEEE conference on Prognostics and Health management*, pages 1–8, 2013.

[82] A.S. Pedersen and C.S. Steiniche. Safe operation and emergency shutdown of wind turbines. *Master's Thesis, Department of Electronic Systems, Aalborg University*, 2012.

[83] P. Gipe. Wind power: Renewable energy for home, farm, and business. *Chelsea Green Publishing, USA*, 2004.

[84] EMD International A/S. Turbulence impact assessment: Hiiumaa offshore wind farm, Estonia.

[85] E. Hau. Wind turbines. *Germany, Springer*, 2006.

[86] E.C. Morgan, M. Lackner, R.M. Vogel, and L.G. Baise. Probability distributions for offshore wind speeds. *Energy conversion and management*, 52:15 – 26, 2011.

[87] K. Worden. Signal processing and instrumentation: Lecture notes. *Department of Mechanical Engineering, The University of Sheffield*, 2009.

[88] K.H. Ruhm. Deterministic, nondeterministic signals. *Metrology: Measurement Science and Technology*, 2008.

[89] H. Long, J. Wu, and P.J. Tavner. Analysis of statistical loading conditions of wind turbine gearboxes based on scada data. *EWEA conference proceedings, Copenhagen, Denmark*, 2012.

[90] N. Baydar and A. Ball. Detection of gear deterioration under varying load conditions by using the instataneous power spectrum. *Mechanical Systems and Signal Processing*, 14(6):907 – 921, 2000.

[91] W. Bartelmus and R. Zimroz. Vibration condition monitoring of planetary gearbox under varying external load. *Mechanical Systems and Signal Processing*, 23(1):246 – 257, 2009.

[92] R. Zimroz and A. Bartkowiak. Two simple multivariate procedures for monitoring planetary gearboxes in non-stationary operating conditions. *Mechanical Systems and Signal Processing*, 38(4):701 – 712, 2011.

[93] W. Bartelmus and R. Zimroz. A new feature for monitoring the condition of gearboxes in nonstationary operating conditions. *Mechanical Systems and Signal Processing*, 23(1):1528 – 1534, 2009.

[94] X. Wang, V. Makis, and M. Yang. A wavelet approach to fault diagnosis of a gearbox under varying load conditions. *Journal of Sound and Vibration*, 329:1570 – 1585, 2010.

[95] Y. Zhan, V. Makis, and A.K.S. Jardine. Adaptive state detection of gearboxes under varying load conditions based on parametric modelling. *Mechanical Systems and Signal Processing*, 20:188 – 221, 2006.

[96] C.J. Stander and P.S. Heyns. Instantaneous angular speed monitoring of gearboxes under non-cyclic stationary load conditions. *Mechanical Systems and Signal Processing*, 19(4):817 – 835, 2005.

[97] C.J. Stander, P.S. Heyns, and W. Schoombie. Using vibration monitoring for local fault detection on gears operating fluctuating load conditions. *Mechanical Systems and Signal Processing*, 16(6):1005 – 1024, 2002.

[98] R. Zimroz, J. Urbanek, Bartelmus W. Barszcz, T., F. Millioz, and N. Martin. Measurement of instantaneous shaft speed by advanced vibration signal processing - Application to wind turbine gearbox. *Metrology and measurement systems*, 18:237 – 247, 2013.

[99] A.C.W. Creech, W.G. Fruh, and P. Clive. Actuator volumes and hr-adaptive methods for three-dimensional simulation of wind turbine wakes and performance. *Wind Energy*, 15:847 – 863, 2012.

[100] AGMA. Designing and specification of gearboxes for wind turbines. *ANSI/AGMA/AWEA 6006-A03*, 2003.

[101] R.B. Randall, J. Antoni, and S. Chobsaard. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, 15(5):945 – 962, 2001.

[102] A. Kahraman and R. Singh. Non-linear dynamics of a spur gear pair. *Journal of Sound and Vibration*, 142(1):49–75, 1990.

[103] A. Kahraman and R. Singh. Non-linear dynamics of a geared rotor-bearing system with multiple clearances. *Journal of Sound and Vibration*, 144(3):469–506, 1991.

[104] A. Kahraman and R. Singh. Interactions between time-varying mesh stiffness and clearance non-linearities in a geared system. *Journal of Sound and Vibration*, 146(1):135–156, 1991.

[105] R.G. Parker, S.M. Vijayakar, and T. Imajo. Non-linear dynamic response of a spur gear pair: Modelling and experimental comparisons. *Journal of Sound and Vibration*, 237(3):435–455, 2000.

[106] S. Theodossiades and S. Natsiavas. Non-linear dynamics of gear-pair systems with periodic stiffness and backlash. *Journal of Sound and Vibration*, 229(2):287–310, 2000.

[107] S. Theodossiades and S. Natsiavas. Periodic and chaotic dynamics of motor-driven gear-pair systems with backlash. *Chaos, Solitons & Fractals*, 12(13):2427–2440, 2001.

[108] M. Vaishya and R. Singh. Analysis of periodically varying gear mesh systems with coulomb friction using floquet theory. *Journal of Sound and Vibration*, 243(3):525–545, 2001.

[109] T.C. Lim and R. Singh. Vibration transmission through rolling element bearings. Part III: Geared rotor system studies. *Journal of Sound and Vibration*, 151(1):31–54, 1991.

[110] H.N. Ozguven and D.R. Houser. Mathematical models used in gear dynamics – A review. *Journal of Sound and Vibration*, 121(3):383–411, 1988.

[111] J. Wang, R. Li, and X. Peng. Survey of nonlinear vibration of gear transmission systems. *Applied Mechanics Reviews*, 56(3):309–329, 2003.

[112] A. Parey and N. Tandon. Spur gear dynamic models including defects: A review. *The Shock and Vibration Digest*, 35:465–478, 2003.

[113] J. Jonkman. AERODYN, FAST and ADAMS design codes. 2012. URL `http://wind.nrel.gov/designcodes/simulators`.

[114] J.D. Smith. Gear noise and vibration. 1999.

[115] H. Endo, R.B. Randall, and C. Gosselin. Differential diagnosis of spall vs. cracks in the gear tooth fillet region: Experimental validation. *Mechanical Systems and Signal Processing*, 23(3):636–651, 2009.

[116] I. Howard, S. Jia, and J. Wang. The dynamic modelling of a spur gear in mesh including friction and a crack. *Mechanical Systems and Signal Processing*, 15(5):831 – 853, 2001.

[117] F. Chaari, W. Baccar, M.S. Abbes, and H. Mohamed. Effect of spalling or tooth breakage on gearmesh stiffness and dynamic response of a one-stage spur gear transmission. *European Journal of Mechanics - A/Solids*, 27(4):691–705, 2008.

[118] J. Antoni. The spectral kurtosis: a useful tool for characterising non-stationary signals. *Mechanical Systems and Signal Processing*, 20:282 – 307, 2006.

[119] B. Boashash. Estimating and interpreting the instantaneous frequency of a signal: Part II. Algorithms and applications. *Proceedings of the IEEE*, 80(4):540 – 568, apr 1992.

[120] Z. Yang, Z. Yu, C. Xie, and Y. Huang. Application of Hilbert-Huang transform to acoustic emission signal for burn feature extraction in surface grinding process. *Measurement*, 47:14–21, 2014.

[121] S. Osman and W. Wang. An enhanced Hilbert-Huang transform technique for bearing condition monitoring. *Measurement Science and Technology*, 24(8):085004, 2013.

[122] E.S. Carbajo, R.S. Carbajo, C. Mc Goldrick, and B. Basu. ASDAH: An automated structural change detection algorithm based on the Hilbert-Huang transform. *Mechanical Systems and Signal Processing*, page Available online, 2013.

[123] R.S. Carbajo, E.S. Carbajo, C. Mc Goldrick, and B. Basu. Real-time autonomous structural change detection onboard wireless sensor platforms. *Key Engineering Materials*, 569-570:970–977, 2013.

[124] S. Pavlopoulou, W.J. Staszewski, and C. Soutis. Evaluation of instantaneous characteristics of guided ultrasonic waves for structural quality and health monitoring. *Structural Control and Health Monitoring*, 20 (6):937–955, 2013.

[125] N.E. Huang, Z. Wu, S.R. Long, K.C. Arnold, X. Chen, and K. Blank. On instantaneous frequency. *Advances in Adaptive Data Analysis*, 1 (2):177 – 229, 2009.

[126] S. Qian and D. Chen. Joint time-frequency analysis: Methods and applications. *Upper Saddle River, NJ: Prentice-Hall*, 1996.

[127] M. Feldman. In *Hilbert transform applications in mechanical vibration*. Wiley, 2011.

[128] R.J. Gledhill. Methods for investigating conformational change in biomolecular simulations. *Ph.D. Thesis, University of Southampton*, page 201, 2003.

[129] N.E. Huang, Z. Wu, S.R. Long, K.C. Arnold, X. Chen, and K. Blank. On instantaneous frequency. *Advances in Adaptive Data Analysis*, 1 (2):177 – 229, 2009.

[130] D. Hilbert. Grundzüge einer allgemeinen, therie der linearen integralgleichungen. 1953.

[131] D. Gabor. Theory of communication. *Journal Institution of Electrical Engineers London*, 93(3):429–457, 1946.

[132] H.M. Teager and S.M. Teager. A phenomenological model for vowel production in the vocal tract. In *Speech Science: Recent Advances*, chapter 3, pages 73–109. College-Hill Press, San Diego, CA, 1985.

[133] H.M. Teager and S.M. Teager. Evidence for nonlinear sound production mechanisms in the vocal tract. In *International Conference on Acoustics, Speech, and Signal Processing*, volume D, chapter 55, pages 241–261. Kluwer Academic Publications, France, 1990.

[134] J.F. Kaiser. On a simple algorithm to calculate the 'energy' of a signal. In *International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 381–384, April 1990.

[135] J.F. Kaiser. On Teager's energy algorithm and its generalization to continuous signals. In *Proceedings of IEEE DSP Workshop*, New Paltz, NY, September 1990.

[136] P. Maragos, J.F. Kaiser, and T.F. Quatieri. Energy separation in signal modulations with application to speech analysis. *IEEE Transactions on Signal Processing*, 41(10):3024 – 3051, 1993.

[137] E. Kvedalen. Signal processing using the teager energy operator and other nonlinear operators. *Ph.D. Dissertation*, 2003.

[138] K. Worden and G.R. Tomlinson. Nonlinearity in structural dynamics: Detection, identification and modelling. *Institute of Physics Publishing*, 2001.

[139] K. Worden, G. Manson, and N.R.J. Fieller. Damage detection using outlier analysis. *Journal of Sound and Vibration*, 229(3):647–667, 2000.

[140] N. Dervilis, R.J. Barthorpe, and K. Worden. Comparative study of robust novelty detection techniques. *Key Engineering Materials*, 569-570:1109–1115, 2013.

[141] E.J. Cross. On structural health monitoring in changing environmental and operational conditions. *Ph.D. Thesis, Department of Mechanical Engineering, The University of Sheffield*, 2012.

[142] D.G. Goring and V.I. Nikora. Despiking acoustic doppler velocimeter data. *Journal of Hydraulic Engineering*, 128:117 – 126, 2002.

[143] S. Dou and B. Romanowicz. An automated despiking algorithm for seismic normal mode data. *Annual report, Department of seismology, Berkley University*, 2011.

[144] D.L. Donoho and I.M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81:425 – 455, 1994.

[145] E.J. Cross, K. Worden, and Q. Chen. Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 467(2133):2712–2732, 2011.

[146] R. Engle and C. Granger. Co-integration and error-correction: representation, estimation, and testing. *Econometrica*, 55:251 – 276, 1987.

[147] E.J. Cross and K. Worden. Cointegration and why it works for shm. *Journal of Physics: Conference Series, MPSVA*, 382:287–290, 2012.

[148] W. Fuller. Introduction to statistical time series. *New York, NY: Wiley-Interscience*, 1996.

[149] J. E. Cooper and K. Worden. On-line physical parameter estimation with adaptive forgetting factors. *Mechanical Systems and Signal Processing*, 14(5):705 – 530, 2000.

[150] L. Ljung and T. Soderstrom. Theory and practise of recursive estimation. *Cambridge, MA: MIT Press*, 1983.

[151] K. Worden, E.J. Cross, and A. Kyprianou. Cointegration and nonstationarity in the context of multiresolution analysis. In 9^{th} *International Conference on Damage Assessment of Structures*, 2011.

[152] G. Manson. Identifying damage sensitive, environment insensitive features for damage detection. In 3^{rd} *International Conference on Identification in Engineering Systems, Swansea.*, 2002.

[153] E.J. Cross, G. Manson, K. Worden, and S.G. Pierce. Features for damage detection with insensitivity to environmental and operational variations. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, pages 1471 – 2946, 2012.

[154] K. Worden, E.J. Cross, I. Antoniadou, and A. Kyprianou. A multiresolution approach to cointegration for analysis of time-varying systems and structures. *Submitted to Mechanical Systems and Signal Processing.*

[155] K. Worden, C.R. Farrar, G. Manson, and G. Park. The fundamental axioms of structural health monitoring. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 463(2082): 1639 – 1664, 2007.

[156] S. Mallat. A wavelet tour of signal processing, Second Edition. *Academic Press*, 1998.

[157] D.E. Newland. An introduction to vibrations, spectral and wavelet analysis, Third Edition. *Longman Scientific and Technical*, 1993.