University of Huddersfield Repository

Abdusslam, Shukri Ali

Detection and Diagnosis of Rolling Element Bearing Faults Using Time Encoded Signal Processing and Recognition

Original Citation

Abdusslam, Shukri Ali (2012) Detection and Diagnosis of Rolling Element Bearing Faults Using Time Encoded Signal Processing and Recognition. Doctoral thesis, University of Huddersfield.

This version is available at http://eprints.hud.ac.uk/id/eprint/17806/

Detection and Diagnosis of Rolling Element Bearing Faults Using Time Encoded Signal Processing and Recognition

Shukri Ali Abdusslam

A thesis submitted to the University of Huddersfield in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Huddersfield

December 2012

LIST OF CONTENTS

LIST OF FIGURES

6

7

9

LIST OF TABLES

ABSTRACT

This thesis presents a systematic study of using TESPAR (Time Encoded Signal Processing and Recognition), which presently is in use as an effective tool for speech recognition and shows great advantages in computational demands and accuracy, to develop a new technique for rolling element bearing fault detection and diagnosis.

The fundamentals of rolling element bearings are presented in line with different failure modes and relevant monitoring methods in the time domain, the frequency domain, the envelope spectrum and the wavelet analysis. These reviews show that vibration measurements are a proven and widely accepted data source for bearing monitoring of machinery.

This research thus has focused on developing TESPAR based approaches using vibration signals which are generated from bearings under different severities of faults located at the outer race, the inner race and the roller element. It firstly examines the theoretical basis of TESPAR and examines the diagnosis performance with a number of different simulated signals, which confirms that TESPAR based methods are able to resolve different signals by using their statistics including S-matrix, A-matrix and epoch duration, which paves a frame work to process and interpolate the bearing signal.

With understandings of the insights of bearing vibrations and TESPAR approaches a signal processing framework is then suggested to analyse bearing vibration signals. It consists of a pre-processing step which removes possible noise in the signal, a TESPAR coding step which converts the signal into TESPAR representations-TESPAR streams, a feature calculation

step, which produces different TESPAR statistic parameters, and finally a diagnosis step which applies common statistics to TESPAR statistic parameters to obtain required results.

The TESPAR solution proposed in this thesis shows that discrimination between different bearing signal waveforms has been implemented successfully. TESPAR S- and A-Matrices were constructed for the cases tested and used together with statistical correlation to differentiate between the types of faults. However, the severities of bearing faults have been identified using another TESPAR feature called the mean absolute magnitude value calculated using epoch durations.

The performance of the TESPAR approach was then evaluated against the envelope spectrum; this being the most common method for bearing condition monitoring that is conducted in two terms; the process complexity and diagnosis performance.

A major contribution of this research programme is the development of a method that can provide improved detection and diagnosis of bearing fault types and severity of faults seeded into roller bearings.

DECLARATION

No portion of the work presented in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

DEDICATION

I dedicate this work to my dear parents for the lifelong and boundless love that I received from them as well as their unwavering support, also dedicated to my wife for her unfaltering encouragement throughout my study and my little daughter Lamar. Without their support, reaching this point would not have been possible.

ACKNOWLEDGEMENT

Few goals are ever achieved without the help of teachers, mentors, colleagues, family, and friends. For the assistance, support, and guidance I have received while completing my postgraduate studies I am indebted to ***Prof. Andrew Ball***, I would like to express my sincere gratitude to him. Also I would like to express my genuine appreciation to my second supervisor ***Dr. Fengshou Gu*** for his continuous direction, support and patience throughout my PhD research above and beyond the merely technical during my research, it was really my privilege to have such a supervisory team. I would like to acknowledge invaluable input and countless enlightening meetings, discussions and conversations with them.

My special appreciation goes to Mr ***Ashley Elkins*** for his kind help and assistance in this research.

My special thanks goes to my colleagues and friends, I was fortunate to have nice peers in my research group.

My thanks also go to the technical team for their kind help and cooperation.

COPYRIGHT

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns any copyright in it (the "Copyright") and he has given The University of Huddersfield the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.

ii. Copies of this thesis, either in full or in extracts, may be made only in accordance with the regulations of the University Library. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made.

iii. The ownership of any patents, designs, trademarks and any and all other intellectual property rights except for the Copyright (the "Intellectual Property Rights") and any reproductions of copyright works, for example graphs and ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such intellectual property rights and reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant intellectual property rights and/or reproductions.

PUBLICATIONS

Abdusslam, S.A., Raharjo, Parno, Gu, Fengshou and Ball, Andrew (2012) Bearing defect detection and diagnosis using a time encoded signal processing and pattern recognition method. Journal of Physics: Conference Series, 364. 012036. ISSN 1742-6596.

Raharjo, Parno, Abdusslam, S.A., Gu, Fengshou and Ball, Andrew (2012) Vibro-Acoustic Characteristic of A Self Aligning Spherical Journal Bearing due to Eccentric Bore Fault. In: CM 2012 and MFPT 2012: The Ninth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, 12th - 14th June 2012, London, UK.

Abdusslam, S.A., Ahmed, Mahmud, Raharjo, Parno, Gu, Fengshou and Ball, Andrew (2011) Time Encoded Signal Processing and Recognition of Incipient Bearing Faults. IEEE In: 17th International Conference on Automation and Computing (ICAC'11), 10th September 2011, Huddersfield, UK.

Ahmed, Mahmud, Abdusslam, S.A., Baqqar, Mabrouka, Gu, Fengshou and Ball, Andrew (2011) Fault Classification of Reciprocating Compressor Based on Neural Networks and Support Vector Machines. IEEE In: 17th International Conference on Automation and Computing (ICAC'11), 10th September 2011, Huddersfield, UK.

Raharjo, Parno, Abdusslam, S.A., Gu, Fengshou, Wang, Tie and Ball, Andrew (2011) An Investigation of Acoustic Emission Responses of a Self Aligning Spherical Journal Bearing. In: The Eighth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies CM/MFPT 2011, 20th - 22nd June 2011, Cardiff, UK.

Abdusslam, S.A., Gu, Fengshou and Ball, Andrew (2010) Bearing fault diagnosis using time encoded signal processing and recognition. In: CM 2010 and MFPT 2010 : The Seventh International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, 22-24 June 2010, Stratford-upon-Avon, UK.

17

Abdusslam, S.A., Gu, Fengshou and Ball, Andrew (2009) Bearing fault diagnosis based on vibration signals. In: Proceedings of Computing and Engineering Annual Researchers' Conference 2009: CEARC'09. University of Huddersfield, Huddersfield, pp. 93-98. ISBN 9781862180857.

Abdusslam, S.A., Ball, Andrew and Gu, Fengshou (2008) Diagnosis and Prognosis of Machinery Health based on Advanced Intelligent Computations. In: University of Huddersfield Research Festival 2009, 25 Feb-13 March 2008, Huddersfield.

CHAPTER ONE

Introduction to Condition Monitoring

This chapter reviews the importance of condition monitoring and fault diagnosis for the smooth running of industrial processes. It emphasises vibration measurement as this is both the most commonly used monitoring method in industry and the technique selected for the rolling element bearings used in this project. The sections of the chapter sequentially present the motivation for this research project, outline the aims and objectives and present the structure of this thesis.

1.1   Background

Condition Monitoring (CM) became widespread in the 1960's and since then the technology has developed rapidly [1]. When operating as intended all electrical and mechanical systems generate a characteristic signal. If operating conditions change so does the signal, even if only slightly, indeed the small differences from the normal or healthy signal can be indicative of incipient fault development. Machine CM is the procedure of monitoring a parameter (e.g. temperature or vibration behaviour) of condition in machinery, such that a considerable alteration in a machine's condition is indicative of a developing defect. The use of condition monitoring term permits schedule of maintenance or corrective actions to be taken place to evade catastrophic failure. However, a deviation from a normal value must happen to identify hidden faults. Today many books and journal papers are available in the field of CM and much industrial interest has been expressed both in research and provision of services [2-4].

However, these indicative changes in a machine's condition may be so small that they are hidden by the noise in the system. Current interest is to combine modern transducers and signal processing techniques to differentiate between noise and significant trends and detect the presence of a fault at the earliest stage and even predict likely time to failure [5]. In a CM system, the measured signal changes with the parameter being monitored and, if chosen correctly will be a measurement of the electrical and/or mechanical condition of the machine. Such signals will include vibration level, acoustic emission and temperature. In addition other more traditional measurements such as the condition of the lubricant will enhance the possibility of accurately determining whether a machine may be considered "healthy" or has a fault which needs to be addressed and whether immediate maintenance action needs to be scheduled. Determining maintenance action in this way has enormous benefits; it cuts maintenance costs, extends the life of the machine, and avoids catastrophic failure [6–9].

20

CM is both an administrative and technical activity with the goal of maintaining or restoring a machine or process to a condition to carry out its intended functions. The technique is based on detecting the presence of a fault, diagnosing the cause of the fault, assessing its level of severity and making arrangements for its correction. CM offers numerous advantages, including [10]:

1- Reducing maintenance costs by minimising the number of machine overhauls by eliminating unnecessary interventions.

2- Because CM provides advanced information on the nature of the fault to be repaired advance preparations can be made and the duration of intervention time is minimised thereby minimising production loss.

3- Because CM provides advanced information on the severity of the fault to be repaired the likely duration of the intervention required can be determined early and disruption to the production process is minimised.

4- Avoiding/eliminating catastrophic failure.

The advantages gained by applying CM have led to a vast number of techniques now being available [2-10].

The condition of all dynamic systems changes with time, and thus its signature signal (vibration level, etc.) also changes. These changes in the system signal provide information on the presence, or otherwise, of the onset of a possible failure mode which until recently were commonly masked by system background noise. Today there is a growing emphasis on applying signal processing techniques in order to separate significant trends from random variations because this provides an earlier diagnosis and hence a longer time-to-failure which can be very important where the item of plant or machinery is of crucial importance for the production process, or an important component in any safety procedure. In other words, CM

21

introduces a reliable method to warn of the potential failure of critical components so that downtime can be based on a just-in-time principle rather than on a routine maintenance basis.

1.2   The Importance of Condition Maintenance and Fault Diagnosis

Competition in the global marketplace is forcing industrial companies to make greater efforts to reduce costs and enhance product quality to maintain their competitiveness. Catastrophic failures are dramatic and sometimes fatal and thus can be headline news but chronic failures which are smaller and often less visible are much more prevalent and frequent than catastrophic failure [11]. The consequence of chronic failures is frequent interruptions of production schedules, reduction in product quality and increased production costs. Generally, however, total downtime due to non-catastrophic failures is greater than that due to catastrophic failures, extraordinarily expensive and greatly exceeds maintenance costs [12].

Every year billions of dollars are spent on plant maintenance operations by the world's industries [13] and it has been reported that maintenance costs may account for as much as one third the production costs of goods [14]. Substantial cost savings and increased profitability can be achieved by greater equipment availability and reliability. This requires the implementation of an effective CM and machinery maintenance programme [15]. If CM can be combined with regular and planned maintenance to minimize downtime it will increase productivity and should be cost-effective. Such a system would be expected to have contemporaneously improvement.

Correct and rapid fault detection and diagnosis can significantly improve machine availability. Thus many CM techniques have already been developed, in particular vibration monitoring using time and frequency domains to improve performance of both machines and systems [7, 8 and 16].

22

The earlier the detection and diagnosis the more time maintenance personnel have to take the necessary action to ensure production losses are avoided [10]. Because industrial processes tend to rely on specific machines, these machines are crucial for the smooth functioning of the system and their malfunction will result in costly production shut-downs. Preventive maintenance is vital for these machines and a significant amount of research is now being directed into that area [17, 18]. It follows that CM used with critical machines to provide an early warning of potential failure minimizes expensive downtime, greatly improving manufacturing efficiency, quality and safety [19].

Because mechanical failure can be catastrophic there is the risk of secondary damage to nearby components. CM minimizes such incidents. Traditionally industrial plant maintenance was reactive, unless the machine broke down it was not attended to - in common parlance: "if it isn't broke, don't fix it!" With the advent of cheap and powerful computers and advances in cost-effective sensor technology, by the 1990's many industries had introduced a strategy of preventive maintenance to identify problem machines and predict maintenance requirements. A more recent approach to enhance maintenance technologies is called pro-active maintenance [17] which uses a comprehensive maintenance programme which is a balance of predictive and preventive maintenance.

1.3   Condition Monitoring Using Vibration Analysis

Considerable effort has been spent developing reliable methods for gear fault detection. Techniques which have been proven successful include analysis of lubricating oil, the acoustic signal generated by the gearbox when in operation, temperature and performance monitoring, electrical motor current analysis, angular speed of crank and drive shafts and, most popular today, vibration analysis.

Unfortunately, no one technique is able to detect all machine faults. However, it has been suggested that vibration measurement, which is the most widely used CM technique in industry, can accurately identify 90% of all machinery failures by the change in vibration signals which they produce and the level of signal can give an accurate prediction of future failure [7]. The task is to diagnose the fault at an early stage so corrective action can be taken as early as possible to extend the life of the machine [20].

1.4   Condition Monitoring Application to Rolling Element Bearings

This research focuses on CM of rolling element bearings because they are the most widely used component in rotating machinery and the consequences of bearing failure are the cause of widespread and substantial economic loss and, sometimes, catastrophic failure, [21]. In addition these bearings have presented a difficult task for maintenance programmes for a long time [22].

Rolling element bearings are widely used in all sizes of pumps and motors, and surveys have shown that failure in these elements accounts for just over half of electric motor failures [12]. Bearing defects are often a warning of other faults in the machine because, for example, misalignment and/or imbalance can be the cause of the bearing defects. Thus the CM of rolling element bearings is a very important component of many industrial maintenance programmes [12].

Despite the advances made in bearing design which have significantly increased bearing life, that life remains dependant on well-known factors such as lubrication (or lack thereof), contamination of the lubricant, bearing load, bearing speed, the precision of the initial setting of the component within the machine, the precision with which the component was

manufactured, the bearing maintenance programme and many other environmental factors including ambient temperature [23].

It is known theoretically that the time between installing the bearing and the first signs of material fatigue is a function of the number of bearing revolutions and the magnitude of the bearing load. However, with so many other factors to consider the life of any individual bearing can only be estimated since operating conditions will rarely be ideal and, in any case, there will always be slight variations during manufacture. It is known that apparently identical bearings can show considerably different lifetimes under the same test conditions [24].

The stresses experienced by bearings will invariably be cyclic in nature and such stresses produce metal fatigue. With time such stresses generate sub-surface cracks which gradually extend to the surface of the bearing raceway. The problem is exacerbated because the rolling elements of the bearing run over the cracks causing fragments of the surface to break away and increasing the area and depth of the crack. This is known as spalling or flaking. The surface wear progressively increases until the bearing is no longer viable and must be replaced. Because industrial processes require large numbers of bearings and individual bearing failures cannot be predicted accurately, the run time to failure of the process could be quite short with the danger of catastrophic failure and closure of the entire production line. It is therefore important to implement a plan for appropriate monitoring of the condition of all those bearings in the process machinery which are a key to continued production, preferably continuously, to avoid unplanned downtime and expensive repairs due to bearing failures [25].

Using modern technology a bearing CM programme would be expected to detect and locate the presence of a fault and identify the cause of the damage well before the fault developed to a serious stage [26, 27]. Early detection of bearing faults requires a method sensitive to impulsive signals and the changes in them. Early bearing fault detection has the same problem as all early detection systems, that the signal to be detected is small compared to other similar sources in the machine. A number of signal processing methods have been applied to improve the early defect detection of rolling element bearings and it is often said that focusing on the higher frequencies of the signals should be more successful because sharp impacts generate high frequencies and this has the added advantage of less background noise interference [6, 10]. Additionally a CM programme would track the severity of the fault allowing accurate prediction of bearing failure and timely corrective action to extend bearing life and minimize associated costs.

1.5   Literature Survey for Bearing Monitoring Methods

Localised defects are the most common failure mode for rolling element bearings with many problems in motor operations due to faulty bearings. These typically occur during operation when a significant piece of material is dislodged from the contact surface. Fatigue cracks due to cyclic contact stresses are a common cause [28]. Consequently, in industry there is an emphasis on fault detection and diagnosis of localized defects.

Correct defect diagnosis depends on using suitable signal analysis techniques. The vibration signal from the transducer will be complex (the sum of many different sources), non-stationary (the fault is developing) and contaminated by background noise. Processing of the signal is necessary to extract useful information related to any bearing fault. They are many signal processing techniques available; all fall into one of three categories: time-domain, frequency-domain and time-frequency domain [29]. These are briefly reviewed below.

1.5.1 Time Domain Analysis

The time-domain signal represents the time history of the energy contained in the signal and is dominated by the most energetic or "noisiest" elements. Time-domain measurement is often considered the simplest of the measurement techniques and requires relatively inexpensive and unsophisticated instrumentation. In the time-domain, a defect condition is often detected and evaluated using statistic descriptors of the vibration signal, such as the peak value, RMS, crest factor or kurtosis [30]. For a fault to be detected from the peak level or RMS value of the time-domain signal it must have progressed sufficiently to be seen "through" the background noise.

Analogue time-domain signals are considered continuous but with digital sampling the magnitude of the signal is collected at a sampling rate which should be at least twice the maximum frequency of interest. Time-domain signals can provide large amounts of useful information but this will usually require further analysis of the signal to extract important characteristics not readily observable. Some of the more common techniques used are explained below.

The RMS (or possibly the peak) value of the vibration signal can be produced as an ink trace on a chart and analysed visually, or simple statistical analysis is carried out and the values obtained examined to determine that they are within acceptable limits. These statistical parameters are often trended to better detect the early presence of bearing damage.

The most commonly used parameters are RMS, peak value, crest factor and kurtosis [29]. These values for a damaged bearing tend to be greater than the values for a normal bearing. Both RMS and peak values of the vibration signal increase with the onset and growth of a fault and by comparing these with values determined for acceptable operation of a normal bearing, the presence of a defect and its severity can be found.

Crest factor and kurtosis are not amplitude dependent; they are measures of the variability or spikiness of the vibration signal. Crest factor and kurtosis are most useful in the early stages of bearing damage when the spikiness of the vibration signal increases noticeably. Unfortunately, as the damage increases, the vibration signal takes on a more random appearance and the magnitudes of crest factor and kurtosis reduce back to those for more normal operation. Crest factor and kurtosis are thus not suited to the detection of bearing defects at later stages. Skew values of the rectified vibration signal were used [31] for the early detection of bearing faults and found able to detect bearing faults at an early stage. It was also found that the results obtained were largely independent of load and speed.

Heng and Nor [32] investigated the application of crest factor, kurtosis and skew to sound pressure and vibration signals in order to detect defects in a rolling element bearing. These authors also compared statistical parameters derived from the beta distribution to separate healthy and faulty cases. It was found that statistical methods can be used to identify different defects present in bearings. It was also demonstrated that there was no significant benefit gained from using the beta function parameters rather than kurtosis and crest factor.

1.5.2 Frequency Domain Analysis

With time domain parameters such as the RMS value, the contribution of an incipient fault will usually be swamped in the overall vibration signal. This is a severe limitation of time domain analysis that is largely overcome by frequency domain analysis. Because each component of a gear has its own characteristic frequency so does any fault associated with that component. These frequency signatures, or signal spectral content, are a good key to determining bearing condition. These signatures will not be lost in the overall signal because they appear at distinct frequencies. The FFT of the vibration signal is the most common

method for converting the time domain to the frequency domain. Trending of the frequency components is a common approach [32].

As a fault develops in a bearing component peaks appear in the vibration spectrum at the bearing defect frequency and its harmonics associated with that bearing component. Around each peak there are side bands. The separation of the side bands is complex and can depend on such factor as transmission path and any periodicities in the load, the amplitudes of the frequency peaks increase with fault severity. For an old, worn system the signal to noise ratio (SNR) will be low and the spectrum will contain a very large number of frequency peaks. In these circumstances it becomes almost impossible to separate out the peak(s) due to a single fault. For frequency analysis using the FFT this is a serious problem [32].

The article [33] entitled "On Initial Fault Detection of A Tapered Roller Bearing: Frequency Domain Analysis" reported an investigation of the vibration frequency spectrum of a defective roller bearing subject to different loads. Both multiple and single point defects located at arbitrary positions were studied. It was found that the peaks at the defect frequency and its harmonics produced by a single point defect on the bearing had an "equal frequency spacing distribution" pattern. It was confirmed that around each peak there were sidebands whose frequencies were related to the periodicity of the load and were a function of the transmission path. For multiple defects the frequency characteristics were found to be the superposition of the frequency characteristics for each defect separately, though this linear relation may have been due to the limited magnitude of the defects.

The most common frequency analysis technique used for detection and diagnosis of bearing faults is envelope analysis [34][35][94, 96, 97]. The technique is reported in details in [34], consider a localised defect hitting a raceway, each time the defect hits the raceway an

impulsive force (wide frequency content) is generated. This will excite structural resonances in the transmission bath between the point of impact and the point of measurement. Envelope analysis is a mechanism for extracting the periodic excitation or the amplitude modulation of the resonance allowing the presence and location of a defect to be detected. The envelope analysis will be thoroughly investigated in Chapter Three in Section 3.3.3 as the envelope has been used as a method in this research because of its widespread and effectiveness use in bearing diagnosis.

1.5.3 Time-Frequency Domain Analysis

The FFT transforms a time domain signal to the frequency domain information about the time domain signal is lost. Time-frequency domain techniques, however, contain both time and frequency domain information which enables investigation of such transient features as impacts. Time-frequency techniques include, for example, the Short Time Frequency Transform (STFT), the Wavelet Transform (WT) and the Wigner-Ville Distribution (WVD). These techniques are claimed to have the ability to detect and diagnose bearing faults in rotating machines where the SNR is low and a large number of frequency components are present [35].

Li and Ma present the use of the Wavelet Transform for detection of local defects in bearings [36]. The uncertainty principle means that good resolutions in either time or frequency-domains can only be achieved by loss of precision in the other [37]. The WT provides a method of scaling the time window to match the frequency of interest and minimize any compromise between time and frequency domains. WTs have been successfully applied to the periodic structural ringing generated by repeated impulsive forces induced by a rolling element passing over a defect [38]. The method was confirmed using experimental data collected from bearings with defects at different locations operating under different load

conditions [38]. The wavelet transform will be looked at in more details in Chapter Three in Section 3.3.4. The wavelet transform is also used in this research for bearings detection and diagnosis.

The article [39] entitled "Wavelet Packet Feature Extraction for Vibration Monitoring" argues that Fourier analysis provides a poor representation of signals well-localized in time because the FFT assumes a stationary signal and contains no information on when an event occurred. The article used the WT as a means of extracting time frequency information from the vibration signal but used statistically based feature selection to discard a large number of features containing little or no useful information this left a feature subset with a substantially reduced number of parameters. Such an approach is useful because it can substantially reduce neural network (NN) classifier training times and increase the ability of the NN classifier to generalise.

1.5.4   Higher Order Spectral Analysis

Higher Order Spectra can determine the phase correlation between different frequencies present in a vibration signal. The presence of a defect in a bearing will generate large values of phase correlation between the harmonics of one (or more) of the defect frequencies [40].

In rolling element bearings bicoherence spectra have been used for automatic detection and diagnosis of localized defects by finding features that indicate the condition of the bearing [40]. The article discussed which features best describe the degree of phase correlation between three harmonics of the characteristic fault frequencies of the bearing [40]. Bicoherence theory was used to detect local bearing defects and it was shown experimentally that the method was effective in detection and diagnosis of incipient bearing defects.

Some researchers have discussed the use of bispectral and trispectral analysis for condition monitoring [40]. They investigated Higher Order Spectra analysis of machine vibrations to determine useful diagnostic features. The experimental work was carried out on a based on a small test rig subjected to bearing faults. The bispectrum or trispectrum of the vibration signal were used as input features to train a NN classifier to determine the condition of the bearings. The researchers then compared the technique with simpler statistical feature and power spectral extraction algorithms and undertook a more detailed investigation of the HoS of the signals to obtain practical features that can be easily estimated to provide readily available improved diagnostic information about the bearings.

There are many other methods for classification, common to that are data-driven methods which is the desired system output modelling using historical data, and not necessarily of the system mechanics, these techniques include conventional numerical algorithms such as that are normally found in the machine learning and data mining areas. The latter algorithms encompass artificial neural networks (ANNs), Support Vector Machines (SVMs), fuzzy logic and decision trees etc. We enumerate below the most popular data-driven methods for fault diagnosis are enumerated below. The review given in [41] provides an extensive overview over data-driven methods in the context of computational intelligence.

1.5.5   Neural Networks Approach

With the Neural Networks (NNs) approach, bearing fault diagnosis is viewed as a pattern recognition problem. Standard techniques are used to extract features from the vibration signals which are then used for training NNs to match the features to a specified condition of the bearing (typically normal, fault 1, fault 2, etc.).

The extraction method must be such that features obtained for different classes of fault form well separated clusters in the feature space. The performance attained with the NN approach depends strongly on the method of feature extraction.

Two researchers [42] report the results obtained from two NN based approaches; a multi-layer feed forward NN trained using an Error Back Propagation (EBP) technique and an unsupervised Adaptive Resonance Theory-2 (ART2) based NN. These two methods were applied to the automatic detection/diagnosis of local defects in ball bearings. The vibration signal was recorded for defective and normal bearings for a range of speeds and loads. Statistical measures of the signals were extracted and used as features to train the NNs. The outputs of each NN denoted the state of the ball bearing; the success rate of the EBP technique was more than 95%, while that of the ART network was even better, 100%. The latter was also found to be very fast.

One researcher [43] has described how bearing vibration can play an important role in the performance of driven motor systems. In many cases they found the accuracy of controlling devices were strongly dependent on the dynamic performance of the bearings of the motor. Thus, fault detection of a motor system is directly related to the diagnosis of the bearing assembly. These authors examined the use of the vibration frequency features of bearings for identifying bearing faults. Next, they discussed a method of fault diagnosis using NN and features from time frequency-domain analysis of bearing vibration. Using both real-world testing and simulation they demonstrated that such a system can effectively detect a number of motor bearing faults. Others have analysed the vibration spectra of bearings using fuzzy logic and showed that fuzzy logic could be usefully applied for accurate diagnosis of bearing faults if the input data was processed appropriately [44].

1.5.6 Support Vector Machines

Neural Networks (ANNs) have been proven that they are good classifiers, but yet they need huge number of samples for training, and that is not at all times true in practice [45]. Support vector machines (SVMs) are based on the theory of statistical learning and they focus on a smaller number of samples. SVMs have better generalisation than ANNs and assure the global and local optimal solution similar to that attained by ANN [46]. Recently, SVMs have been tested to be good and effective in numerous real-world scenarios [47, 48]. As it is difficult to gain enough fault samples in practice, SVMs have been used on applications for machinery defect diagnosis by different researchers in recent years [49-51]. Envelop spectrum of intrinsic mode function has been used as input data to SVMs for bearing defects classification, also improved wavelet transform and SVMs have been applied for bearing defects detection [51].

SVM is developed from the optimal separation plane under linearly distinguishable condition. Its fundamental principle can be demonstrated in two-dimensional manner. Generally, the SVM attempts to put a linear boundary between two categories of data and orients them in such manner that the margin is enlarged and the distance separates the boundary and the closest point of data in each category is maximal. The closest points of data are used to identify the margin and are named as support vectors.

1.5.7 Fuzzy Logic

The rule of Fuzzy based systems employ fuzzy logic for inference. Fuzzy logic method is based on the concept or the theory of fuzzy set in a way that binary set has been extended to take in partial membership that range between 0 and 1 [52]. By contrast to other similar techniques, fuzzy sets have gradual transitions between defined sets, which permit to directly

34

model the uncertainty associated with these concepts. The mapping of inputs data to outputs can be considered as a set of IF-THEN rules and that after identifying each model variable with a series of overlapping fuzzy sets, this process completely determined from expert knowledge, or from data. Nevertheless, unlike neural networks, fuzzy models tend to a rule explosion, for instance, the number of rules increase exponentially if the number of variables or fuzzy sets per variable increases that makes it complicated to identify the whole model from the knowledge of an expert only [53]. There are different automated techniques have been used recently for optimising fuzzy models [53], including neural networks and genetic algorithms.

The fuzzy sets and rules are depended on the knowledge-base of fuzzy model. Inputs to the fuzzy model are initially fuzzified via the knowledge-base; the rules are processed by a fuzzy inference engine via a fuzzy inference procedure [54], then the fuzzy surface solution resulted from the conduction of the rule-base is defuzzified to generate the system output(s). Fuzzy IF-THEN rules can also be comprised of functional consequents, typically of a linear or polynomial form [55, 56]. The crisp inputs are fuzzified according to the fuzzy set definitions, joint via the inference engine, and the functional consequents are weighted which result from the implementation of the rules. The overall outcome is the equations weighted average as more than a rule can positively fire during a single pass of the rule-base. Fuzzy logic has found numerous successful applications in which the fuzzy rule-based systems are capable to be built by processing historical data and therefore shaping a data-driven model.

1.5.8 Model-Based Techniques

Model based techniques attempt to match a mathematical model to an electrical, mechanical or other physical system. Here a mechanical bearing would be modelled is that its output matched the vibration response of the bearing with and without faults present. A

comprehensive model would include effect due to unbalance, misalignment, etc., to produce the overall response of the system.

A model that described the vibration response of a rolling element bearing has been developed [57], the bearing was subjected to constant radial load with a single point defect on the inner race. This model was relatively sophisticated and included the effects of shaft speed, bearing geometry, load distribution and transfer function. Comparison of measured and predicted vibration spectra confirmed the model performed satisfactorily.

A general model has been established for the vibration signals produced by a faulty rolling element bearing [58]. For low shaft speeds they were able to derive the system's envelope-autocorrelation function. A simplified version of the model was used for bearing condition monitoring. Measured and simulated data were compared and the validity and effectiveness of envelope-autocorrelation confirmed for effective fault detection for shaft bearings at low speeds.

A one-dimensional, multiple degree-of-freedom model has been developed for fault detection using the vibration produced by rolling element bearings in a rotor-bearing system [59]. The faults were situated on the outer and inner raceways. Two outputs (inner and outer ring acceleration) were defined based on the theory of Detection Filter Design. It also has been reported that a robust monitoring system which integrated a sliding mode detector capable of isolating fault surfaces in the residual space of the detection filter [59].

1.5.9 Bearing Prognostics Using Vibration Signatures

An important area of condition monitoring research is the prediction of the remaining life of a bearing, which many bearing manufacturers take into account, using trends extracted from the vibration signal. However, relatively few papers have been published on bearing

prognosis because (i) usually, there is insufficient vibration data available and (ii) the difficulty in estimating the remaining life of a bearing even with an accurate vibration history [60]. Most papers in this field are concerned with statistical life estimation using data from laboratory experiments.

Currently most calculations of the fatigue life of a bearing are based on the Standard Life Rating formula provided by ANSI/AFBMA and which is based on the fatigue life theory [60]:

$$L_n = a_1 a_2 a_3 \left(\frac{C}{P}\right)^p$$

where L_n is the rolling contact fatigue life in $revolutions \times \mathbf{10}^6$,

$a\mathbf{1}$ is a reliability factor,

$a\mathbf{2}$ is a material factor,

$a\mathbf{3}$ is a lubrication factor,

C is the basic load rating of the bearing,

P is the equivalent load applied to the bearing, and p is $\mathbf{3}$ for ball bearings and $\mathbf{10/3}$ for roller bearings.

However, real life conditions may depart quite severely from those assumed in the above calculation so the expected and actual lives of a bearing can differ significantly. Thus estimation of remaining life based on online vibration measurement is receiving considerable attention.

A number of tests on bearings have been conducted in which ran for between 100 to 900 hours before failing [61]. The tests included bearings with inner race, outer race and ball defects. The RMS, peak, crest factor and kurtosis of the vibration signal in the frequency

37

band 5.0 -12.5 kHz were measured and trended in the tests. It was found that the RMS was particularly useful for predicting the bearing's remaining life.

The remaining life of a bearing has been estimated using a defect propagation model [62]. The relation of the RMS of the vibration signal in the frequency band 3.0-5.0 kHz to defect size was found in the form of a linear equation. By comparing measured and predicted defect sizes it was possible to develop an adaptive algorithm to fine-tune the parameters of the model. This allowed the rate of defect propagation to be determined from which it was possible to determine an accurate estimate of the bearing's remaining life.

1.5.10 Acoustic Emission Monitoring

Acoustic Emission (AE) has been defines as "transient elastic waves generated from a rapid release of strain energy caused by a deformation or damage within or on the surface of a material [63]. In the application to rotating machinery monitoring, AE are defined as transient elastic waves generated by the interaction of two media in relative motion. Sources of AE in rotating machinery include impacting, cyclic fatigue, friction, turbulence, material loss, cavitations, leakage, etc. For instance, the interaction of surface asperities and impingement of the bearing rollers over a defect on an outer race will result in the generation of AE."

Most AE sources are due to material damage of some form, thus monitoring of AE is commonly used to detect and/or predict material failure. In industry AE is widely used for the detection of leakage in high pressure vessels and piping systems. One of the advantages of AE monitoring is that it can detect the growth of subsurface cracks. Hence, it is an important tool for condition monitoring and it is one of the most effective methods [64].

Generally AE signal processing has as its goal either determining a suitable 'process model' from which the properties of specific variables can be used to define the wave state; or the

creation of a feature data base so that changes in the characteristics of the wave can be monitored. The important problem with processing AE signals is elimination of noise and to extract features that correlate uniquely with target process parameters [65, 66].

AE signals are often classified as continuous or burst-type [67]. In fact continuous AE signals are those where the bursts are sufficiently frequent for the overall signal to have a continuous appearance. Both types are seen during crack growth, material fracture or chipping.

Once AE signal parameters have been established for a healthy configuration, the RMS (or another suitable characteristic) of the AE signal may be monitored and compared to the nominally healthy values to detect the presence of an abnormal event. It has been shown that such monitoring can detect defects before the defects can be detected the vibration signal [68-69]. Also it has been shown that peak amplitude and count for the AE signal have been particularly useful in the detection of defects in ball bearings under load at normal and low speeds [70]. AE monitoring has been applied to detection of faults in rolling element bearings and has proposed the area under the time-amplitude plot as a preferred method for defect detection [71]. The RMS of the AE signal has been successfully used to determine milling tool breakage during the cutting process [72].

At its present stage of development the major disadvantage of commercial AE systems is that they can give only qualitative estimates of how much the test material has been damaged and only estimate the remaining life of a component. Consequently, other non-destructive testing methods are needed for a more rigorous examination and to provide quantitative predictions. While the frequency range of the AE signals used to monitor machinery and equipment is usually well outside the range of background noise nevertheless AE signals are usually weak and must be measured in the vicinity of their generation. Because the AE signals are weak

reduction of noise present in the signal and signal discrimination can be difficult but are necessary for successful application.

1.6   Research Motivation

Many novel and interesting CM methods have been developed in recent years [5]; however, the challenge will always remain of producing a CM system capable of detecting and identifying the presence of a fault at ever earlier stages of its development. This research is focused on the development of an approach based on advanced computations to detect and diagnose bearing condition.

The first motive for this project is that it is important to detect bearing defects as early as possible to provide maximum time for corrective action preventing further damage and extending bearing life. Without such early detection, maintenance staff must wait until the later stages of the fault by which time the damage has become both more severe and extensive.

The second motive is more personal, to be the first to attempt a novel approach to bearing fault diagnosis and detection. Time Encoded Signal Processing and Recognition (TESPAR) is a set of signal analysis and classification software that can be used to describe and classify band limited signals [73].

The use of TESPAR in the context of signal processing is not completely novel there is already work and some publications especially on speech recognition in electronic field as mentioned here: King R A, Gosling W, 'Time Encoded Speech', *Electronics Letters*, 20 July 1978, also King R. A. & Phipps T.C. "Shannon, TESPAR and Approximation Strategies September 1998 and Vu V V and King R A; "Automatic Diagnostic and Assessment Procedures for the Comparison and Optimisation of Time Encoded Speech (TES) DVI

systems" September 1989. In addition, in the context of condition monitoring the TESPAR S and A matrices have been used generally to analyse the input data there are few publications such as : Rodwell G M and King R A, "TESPAR/FANN Architectures for low-power, low cost Condition Monitoring Applications" July 1996, and George M H, 'Time for TESPAR', Condition Monitor Number 105, September 1995. However, the novelty of this research is the adaption of the approach using a new feature and statistics with TESPAR matrices to identify bearing conditions in both fault location and severity.

The history and development of TESPAR is described in Chapter 4. Here it is sufficient to say that originally TESPAR was developed to process speech waveforms which are known to be very complex and highly dynamic [74], an approach which could be useful for the analysis of impulsive vibration signals. The idea here is to use this technique to explore fault detection in a bearing rolling element and to compare the capability and effectiveness of TESPAR based methods with more traditional techniques in terms of identifying the onset of a defect in different bearing elements. These traditional techniques are described thoroughly in Chapter Three and will include time domain analysis, frequency domain analysis and wavelet analysis.

Specifically the idea of using the TESPAR approach as a focus for this research comes from the observation that almost all current commonly used bearing CM techniques are frequency domain or time frequency domain, and that there are insufficient studies in the time domain, of course, except the use of very basic statistics parameters, thus, this research considers a new focus in the time domain by investigating and exploring the application of TESPAR to fault detection and diagnosis.

During this research the following TESPAR benefits that exist over other existing methods have been identified:

- **Early Fault Detection & Identification**: TESPAR is able to utilise its classification capability to identify faults at the earliest stages of their inception. Once detected, any fault can either be repaired or continuously monitored to identify changes in its severity and if appropriate, predict the time duration before a catastrophic failure.

- **Low Power and Low Processor/Memory Solution**: TESPAR requires two orders of magnitude less processing power and memory than its frequency domain counterparts which opens up the prospect of bringing real-time condition monitoring to devices possessing much less processing power than a DSP (Digital Signal Processor). Moreover, a significant reduction in battery power is required which provides a greater scope for tackling problems previously consider infeasible because of the amount of operating current required. This reduction in processing power and memory has enabled a real-time TESPAR classification to be successfully implemented to a number of low-end processors such as the 8-bit Intel 8051 and the ARM 7 core.

- **Improve Existing Classification System by Running TESPAR Solution in Parallel**: If a traditional frequency-based system is already monitoring a piece of machinery, in many cases it is possible to also implement the TESPAR solution in parallel with the existing approach by using only the spare capacity of the current processor (e.g. DSP). This is achievable because of the small footprint and processing requirements of the TESPAR approach. This parallel methodology allows the advantages of TESPAR to be realised without having the expense or inconvenience of replacing the current system.

- **Applicable Across a Wide Range of Sensor Types**: TESPAR has been applied to signals from a wide range of sensors including Accelerometers, Vibration, Acoustic emission, Densitometers (nucleonic devices), Pressure, Temperature, Geophones / Hydrophones etc. An important advantage of TESPAR is that it does not rely on expensive high quality sensors in order to successfully achieve its objective – unlike some tradition condition monitoring systems. In a number of real-world case studies, Tespar have achieved similar high levels of performance from both cheap (£10) and expensive (£2000) sensors.

- **Low False Alarms and Misclassifications**: TESPAR achieves a low number of misclassifications and false alarms by coupling the flexibility of the TESPAR "coding table", the information rich TESPAR symbol stream / matrices, with appropriate highly accurate classification tools. The TESPAR solution effectively groups the faults into feature space clusters and by selecting the appropriate TESPAR processing configuration:

  o The elements within these clusters become more tightly grouped and

  o Separations between different clusters become more clearly defined

  o Thus causing a reduction in the number of misclassifications and false alarms.

- **Resistant to Noisy Signals**: TESPAR is highly resistant to noise that may be present and as a result has a significant advantage over traditional frequency based approaches. The TESPAR solution can extract signals from heavy noise contamination even at noise levels of 0dB. There are a number of TESPAR processing techniques that can be used to significantly reduce the effects of noise effects such as white noise and impulse noise. One such advantage over the traditional frequency based domain is in the presence of impulse noise – this typically presents itself across a large proportion of the frequency spectrum, which typically results in a

43

inferior classification accuracy. Conversely in the TESPAR time domain, the impulse noise only appears as one or two epochs in the Tespar vector and so can easily be excluded from the classification.

- **Accurate Identification of Fault Source**: Using TESPAR for fault detection and prediction allows the fault source to be accurately identified in a number of ways:

  o Each fault type to be detected can be assigned one or more of the various TESPAR feature sets / matrices

  o The output of different sensor types can, if required, be combined using higher-level logic to obtain more than one view of the potential fault

  o If a repetitive action is associated with the machinery under test, not only can the fault be detected by examining the TESPAR symbols, it will also be possible to identify its cyclic nature and therefore pinpoint the offending component. Moreover, if a piece of machinery is expected to operate at a known rotational speed, specific TESPAR matrices have been developed to indicate the extent to which the machine has deviated from the expected rotation speed.

- **Model Adaptation**: Once TESPAR models have been generated for each condition to be recognised, it is possible to easily adapt these TESPAR models (matrix archetypes) to ensure that any natural deviation in the machinery's operating condition can be incorporated in real-time into the model.

- **Fault Level Intensity Indicator**: A number of techniques exist that allow the output of the TESPAR classification stage to be represented on a linear scale in order that a meaningful level of fault intensity/severity can be calculated. I.e. The profile of the TESPAR matrices and other feature vectors can be displayed on a scale such that a 1 implies low severity and a 10 implies a high severity.

44

Provide Continuous Real-Time Monitoring: As TESPAR requires little processing overhead, this approach can easily support continuous real-time monitoring.

1.7   Research Aim and Objectives

The aim of this research is to investigate the development of an efficient technique using TESPAR for the more reliable, cost efficient and precise detection and diagnosis of incipient bearing faults.

The main objectives of this research are:

Objective one: To present and discuss machine condition monitoring and the applications of rolling element bearings.

Objective two: To describe the fundamentals of rolling bearing transmission, including bearing types and components, their failure modes and methods of monitoring.

Objective three: To review signal processing conventional methods and their parameters of bearing fault detection and diagnosis using vibration signal.

Objective four: To describe and explore the fundamentals of the TESPAR approach.

Objective five: To discuss and analyse the data outputs from an accelerometer using TESPAR methods applied to the time domain data collected for ten different bearing conditions.

Objective six: To perform a relative evaluation performance of the TESPAR application in order to explore its effectiveness and reliability on bearing detection compared to conventional techniques.

1.8   Thesis Organisation

The thesis begins by reviewing the importance of condition monitoring and fault diagnosis for the smooth running of industrial processes. It emphasises vibration measurement as this is the technique selected in this project for monitoring the rolling element bearings. The motivation for this research is described and the aim and objectives stated.

Chapter two describes the different types of roller bearing and their component parts. It also describes the typical bearing defects that occur and shows that, generally, such defects are due to inadequate maintenance or improper installation. The chapter then reviews dynamic response of roller bearings to local (or point) defects and shows that the typical symptoms are characteristic peaks in the vibration spectrum. Theoretical predictions of the characteristic frequencies for the test bearing are compared with measured results.

Chapter three introduces vibration measurement as a method of condition monitoring and defect detection in roller bearings. It reviews time domain, frequency domain and time-frequency domain analysis of the vibration signal from roller bearings and compares their performance regarding incipient fault detection and diagnosis. Various commonly used statistical parameters are described and their usefulness for incipient fault detection commented on. Chapter four describes the TESPAR technique as being a technique explored by this research, its background is given and an introduction to how it works. The chapter also gives examples of TESPAR in action. In order to provide a new insight into fault detection and diagnosis, simulation of TESPAR has been conducted to illustrate its principles. Finally, the TESPAR coding and implementation process is explained in details leading to the new procedure adopted.

Chapter five reviews the design and construction of a test rig suitable for measurement of parameters associated with incipient defect detection in a roller bearing. The parameters to be measured are discussed and details of the relevant transducers and data acquisition system given, as are details of measurement practice and data management. The manner in which defects are introduced into the bearings is described in detail. The chapter also reports the successful application of envelope and wavelet analysis to detection of faults seeded into a roller bearing.

Chapter six discusses the use of TESPAR and its results. The chapter presents a detailed discussion of the TESPAR findings and assessment of its performance; comparisons are also made with a popular method of bearing defect detection (the envelope spectrum) to evaluate the capability and effectiveness of the proposed techniques in real situations.

Chapter seven summarizes the findings and conclusions and also proposes future work based on this research. It presents the author's conclusions on CM of rolling element bearings using TESPAR approach. It also reviews the aim and objectives one by one comparing them to those set for this study and presented in Chapter one. These are followed by the contributions to knowledge made by this research. Finally the author makes suggestions for future work.

CHAPTER TWO

Fundamentals of Rolling Element Bearings

This chapter describes the different types of roller bearing and their component parts. It also describes the typical bearing defects that occur and shows that, usually, such defects are due to inadequate maintenance or improper installation. The chapter then reviews the dynamic response of roller bearings to local (or point) defects and shows that the common symptoms are characteristic peaks in the vibration spectrum. Theoretical predictions of the characteristic frequencies for the test bearing are compared with measured results and good agreement found.

2.1   Introduction

Rolling element bearings are used in their tens of millions in rotating machinery to minimise friction between adjacent parts moving at different speeds. They are not only one of the most critical components but also one of the first to fail. The problem is a major one because, as it has been explained, fewer than 20% of bearings achieve their design life [75] and other research has found bearing failures account for more than 50% of all motor failures [12]. The economic loss due to bearing failure in terms of machine down time, and even human life, is huge compared to the cost of the bearing itself [76]. Therefore, bearing condition monitoring (CM) and diagnosis has attracted substantial attention over the past four decades as economic pressure to increase machine speed has accelerated [77 - 97].

Before discussing bearing CM, and fault detection and diagnosis, it is necessary to understand the principles of operation of rolling element bearings and the signals produced by faulty bearings. Thus, this chapter introduces basic types of rolling element bearings and their major components, common fault generating mechanisms and bearing faults, and the signals to be looked for amongst the general background noise.

2.2   Types of Rolling Element Bearings

Bearings can be divided into two general categories; rolling element and sliding bearings (hydrodynamic oil film). We are concerned only with the former. Depending on the application there are many different sizes and designs of rolling element bearings which can be divided according to the shape of the rolling elements, though it is also possible to classify the bearings depending on the way the load is supported. The standard geometric shapes of the rolling elements that make up rolling bearings are listed in Table 2-1.

Table 2-1 Rolling element bearing types [98]

Rolling Bearing	Roller Bearing	Radial Roller Bearing	Cylindrical single row Cylindrical double row Tapered single row Tapered double row Needle roller bearings Spherical roller bearings
		Axial Thrust Roller Bearing	Cylindrical roller thrust bearings Tapered roller thrust bearings Needle roller thrust bearings Spherical roller thrust bearings
	Ball Bearings	Radial Ball Bearings	Single row deep groove Maximum capacity type Single row angular contact Duplex angular contact Double row angular contact Four points contact Self-aligning
		Thrust Ball Bearings	Single direction with flat back face Single direction with sealing race Double direction with flat back face Double direction with sealing race Double direction angular contact

The shape of the basic rolling element, ball or roller, determine into which of the two families any particular bearing is placed. Each family includes a variety of types, see Table 2-1 which will depend on such design requirements as physical space available for the bearing, magnitude and direction of load (radial, axial, and/or combined load), any misalignment or axial displacement, speed of rotation, whether quiet running is required, and so on. Ball bearings tend to be used in light to moderate load applications and high-speed operations. Roller bearings are used to support heavier loads so therefore, heavy engineering industries tend to use roller bearings rather than ball bearings.

Depending on the direction of the applied load rolling element bearings can also be classified as radial or thrust. These bearings are used to transfer loads between two objects (one rotating relative to the other). As its name implies a radial bearing carries a radial load, which is force

acting at right angles to the shaft, while a thrust bearing is intended to carry an axial load, which is force acting parallel to the shaft. In practice most rolling bearings carry both radial and axial loads so there is some overlap in their use.

Consider the bearing shown in Figure 2-1 whether it is considered to be under radial load or axial thrust is determined by the magnitude of the contact angle a_0. The contact angle is defined as the angle between a plane perpendicular to the bearing axis passing through the centre of the bearing (vertical in Figure 2-1) and a line joining the two contact points the ball makes with the inner and outer raceways. Increasing axial load increases the contact angle. Bearings where $a_0 < 45°$ are radial bearings and their ratings are given by radial load. Bearings with $a_0 > 45°$ are thrust bearings and are rated by axial load [76, 99].
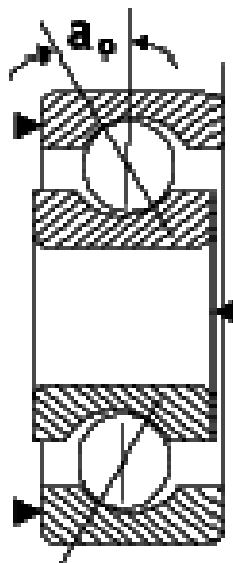
Figure 2-1 Schematic diagram showing contact angle, a0 [100]

Rolling bearings can be classified according to their functions. For example, ball bearings perform well at high speed under moderate radial and axial loads. They have low friction and can be produced with high precision and there is a quiet running variant. Thus they are preferred for e.g. small and medium-sized electric motors. However, in this study, the focus

51

is on roller bearings because this is the most popular rolling element bearing due to its high dynamic load capabilities and toleration of some misalignment. However, most studies based on roller bearings can be extended to other kinds of rolling element bearings.

2.3   Bearing Description and Components

Most rolling bearings have four basic elements: inner race, outer race, rolling elements, and cage or separator [76]. The inner race, outer race, and rolling elements support the bearing load, while the cage separates adjacent rolling elements to avoid friction between them, see Figure 2-2 and Figure 2-3. Thrust bearings which are designed to support an axial load are not considered in this thesis.
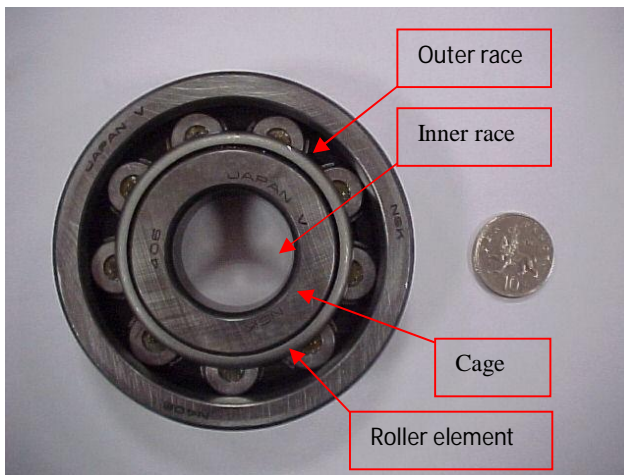
Figure 2-2 Rolling elements bearing, NSK type N406
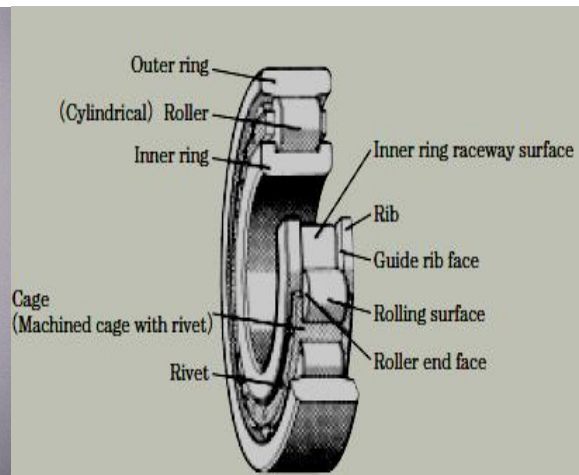
Figure 2-3 Schematic of cylindrical roller bearing [100]

2.3.1 Inner Race

The inner race is mounted on the shaft of the machine and so will usually be the rotating element. Depending on the type of rolling elements (spherical, cylindrical, or tapered) the raceway in which the rolling elements move will have a different form.

52

2.3.2 Outer Race

The outer race is mounted in the machine housing and so will not usually rotate. Again, depending on the type of rolling elements the raceway in which the rolling elements move has a different form.

2.3.3 Rolling Elements

The rolling elements used in rolling bearings are generally ball and roller bearings. The ball bearing transfers the load by point contact with the raceway, so its load-carrying capacity is lower than that of a roller bearing. Rollers transfer the load via line contact with the raceways. The rolling elements are usually made of "bearing steel" a type of carbon chromium steel.

2.3.4 Cage

The cage separates the rolling elements to prevent contact between them during operation which would restrict lubrication and increase wear very rapidly. Cages are made from cold rolled steel strip.

In addition to the four elements listed above there are two other important components. The seals protect the bearing from contamination and keep the lubricant inside the bearings and so are essential for long and reliable life. The guide races lead the rollers in the bearings so that they will rotate parallel to the shaft and distribute the load evenly to the raceway. These are essential in roller bearings that demand extremely high quality

2.4   Bearing Failure Modes

Rolling element bearings have been subject to extensive research over many years to improve their reliability. However, the large number of bearings associated with any given process increases the likelihood of system failure due to one of them failing, and such system failure can occur in a very short period of time. There are many reasons for early failure, including; excessive loading, inadequate lubrication, insufficient internal bearing clearance due to an excessively tight fit, etc.

53

[76][99-101]. The types of mechanical bearing failure and their relative frequencies are listed in Table 2-2 [102].

Table 2-2 Bearing defection by cause [102]

Failure Mechanism	Failure Frequency (all bearings)
Fatigue	2%
Other	10%
Dimensional discrepancies	29%
Lubrication (Corrosion and Over-rolling)	59%

The most frequent cause of bearing failure is corrosion, which is lubrication related. Chemical reaction occurs at the surface of the bearing most often with water but also possibly with any corrosive material present in the oil. Included as a lubrication failure is over-rolling, here contaminated lubricant deposits a foreign object within the bearing which then becomes trapped between the rolling element and the raceway, and is "over-rolled". The next most common cause is so-called dimensional discrepancies which is a collective term for damage prior to, or during service, due to, e.g., manufacturing flaws, or improper handling during installation such as forcing the bearing into position with hammer blows. These two mechanisms are estimated to cause more than four-fifths-quarters of all bearing failures [102]. Similar figures have been presented for bearing failures in aero engines and the authors draw attention to classical fatigue failure (cracks initiated at surface or sub-surface) as constituting a mere 2% of the total and explain this as being due to maintenance or manufacturing errors or mis-installation of the bearings [103]. Each factor will produce its own particular type of damage and leave its own special mark on the bearing and, in turn, creates secondary damage such as spalling and cracks. Failed bearings, when examined, usually display a combination of primary and secondary damage. In terms of bearing failures an understanding of the underlying mechanisms causing defects and the consequences of those defects helps in determining which features to look for to prevent early failure [100, 101, 103].

Note, smearing of ball bearings is not considered here, nor is electric current damage. Bearing faults can be categorised into three sets: primary damages, secondary damage and general damage [103].

2.4.1   Primary Damage

The primary mode of failure for a properly lubricated correctly assembled and normally operated roller bearing is spalling of the bearing element due to local fatigue rather than the bearing wearing out. This is because of high roller contact load and relatively low wear rate to which they are subject [103].

- *Wear*

Wear of bearings may be defined as the removal of surface material, see Figure 2-4. This may be due to abrasion, fatigue, corrosion or erosion. Normally there is little noticeable wear in rolling bearings, but rapid wear and subsequent failure results when foreign objects (such as grit) enter the bearing or when lubrication is inadequate. Vibration in bearings and/or misalignment can also cause excessive wear [76].

Wear can result from the presence of abrasive particles when foreign material which enters the bearing during cleaning or through a damaged or worn seal causes excessive wear of the bearing. If the wear is detected early so that there is only light surface bruising the bearings may be re-used after the bearing and housing are thoroughly cleaned and properly adjusted. Otherwise the bearing should be replaced. If a seal is damaged then the fine metal particles or grinding dust that are commonly found in factory environments can enter the bearing and cause abrasive wear. Within the machine metal components in moving contact will release fine metal particles into the lubricant, and if these find their way into the bearing can cause excessive wear of the roller body and races and, especially, on the roller end.

Figure 2-4 Wear of the surface caused by abrasive particles [110]

If the intruding particles are fine enough they act as a polish and mirror-finish surfaces result. Lack of effective seals, and lubricant contaminated by foreign particles (which may occur during installation of the bearing) causes small indentations on the rolling elements and around the raceways. Dull, worn surfaces are signs of this type of wear. The major problem here is that any wear changes the bearing adjustment and has been known to cause misalignment which directly affects other parts of the machine [104].

Wear can also result from inadequate lubrication, it is surprising how much wear is caused by too low lubricant levels or lubricant that has lost its lubricating property [105]. Metal to metal contact occurs between moving surfaces which removes the "tops" of the surface asperities (the peaks of the surface roughness) leaving abrasive particles which are not carried away and cause wear, as described above. This wear initially shows up as a mirror-like surface, over-heating results in the surface colour changing to blue/brown with subsequent bearing failure. Too thin lubrication films can result in a frosted appearance on the surface, usually of depth of less than 2.5×10^{-3} mm due to patches of flaking on a minute scale. Close examination reveals multiple hair-line cracks though the surface will not yet be flaking [106].

Misalignment will cause wear. The most common causes of misalignment are bent shafts, dirt on the shaft, shaft threads not square with the shaft seats, and locking nuts with faces that are not square to the thread axis. When a rolling element wear path is not parallel to the raceway

edges misalignment failure will usually be detected on the raceway of the non-rotating race. Extreme misalignment can cause heavy wear in the cage pockets and/or excessive temperature increases.

Noticeable wear of bearings resulting from vibration has been described [100, 104]. This is most usually the case when the machine has been inactive for a long time and there is no lubricant between metal surfaces. If nearby working machines generating high levels of vibration there will be a small movement between the metal surfaces in contact. As a result small particles can, eventually, break away and cause depressions in the metal surfaces known as a "false brinelling" because it occurs when the machine is apparently idle. In roller bearings, the rollers will produce rectangular depression marks, called flutings. In ball bearing the depressions are circular. The presence of the particles causes subsequent wear. Roller bearings appear more sensitive to this kind of wear than do ball bearings. Of course excessive vibration of the bearing it could lead to:

- *Indentation*

This is a form of plastic deformation of material either through foreign particles entering the bearing and causing indentations Figure 2-5, or if the bearing is exposed to excessive load when stationary, or if an excessive force is applied to the wrong race when mounting which is transmitted through the rolling elements.

Indentation caused by faulty mounting or overload is also called "true" brinelling, this defect is usually accompanied by an increase in bearing noise and vibration, and invariably leads to premature bearing failure. It is characterized by indentations in the raceways of both outer and inner races; the distance between the indentations is equal to the roller spacing. Common causes for this type of failure are mounting pressure applied to the wrong race, a sharp impact

57

incorrectly applied to the bearing during mounting or dismounting, overloading while stationary, and excessive hard drive-up on tapered seating.
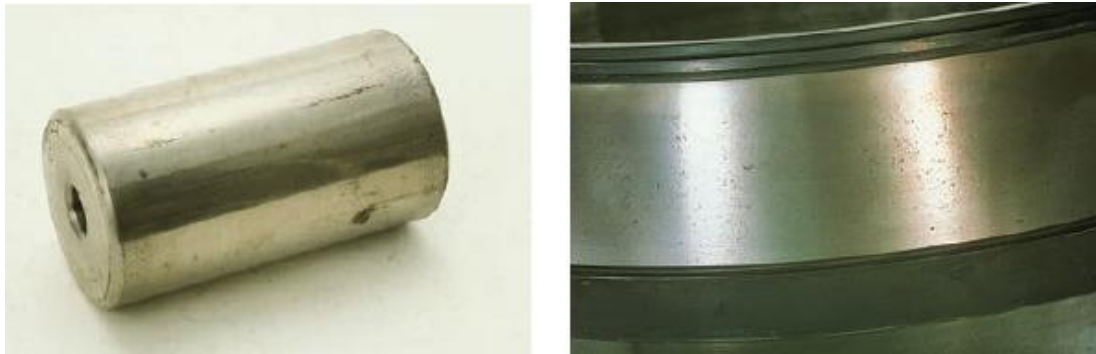
Figure 2-5 Bearing surfaces indentation [110]

Indentation caused by foreign particles (over-rolling) is a leading cause of premature bearing failure. Airborne dust and softer particles would be likely to cause small, shallow indentations distributed around the rolling elements and the inner and outer raceways. Deeper pitting and bruising is caused by dirt/sand particles and larger metal chips. The depth and number of the indentations determines whether the bearing must be replaced.

- *Smearing*

Smearing occurs when the lubrication between two moving surfaces is inadequate (too little or too old) with consequent contact between surface asperities and material is thus transferred from one surface to the other, see Figure 2-6. The surfaces are said to be smeared, and when such smearing occurs the material generally becomes very hot and local re-hardening takes place which may cause localized stress concentrations with surface cracking or flaking.

Smearing may occur on the guiding faces of the flanges or the ends of the rollers because of sliding under heavy axial loading with inadequate lubrication in cylindrical and taper roller bearings, and in spherical roller bearings with guide flanges. Roller ends and flange faces become scored and discoloured.

Smearing of roller and raceways is caused by roller acceleration when a roller enters the load zone. Raceway smearing appears at intervals corresponding to the roller spacing while transverse smearing marks (streaks) appear at intervals equal to the distance between the rollers in the raceways of cylindrical roller bearings.
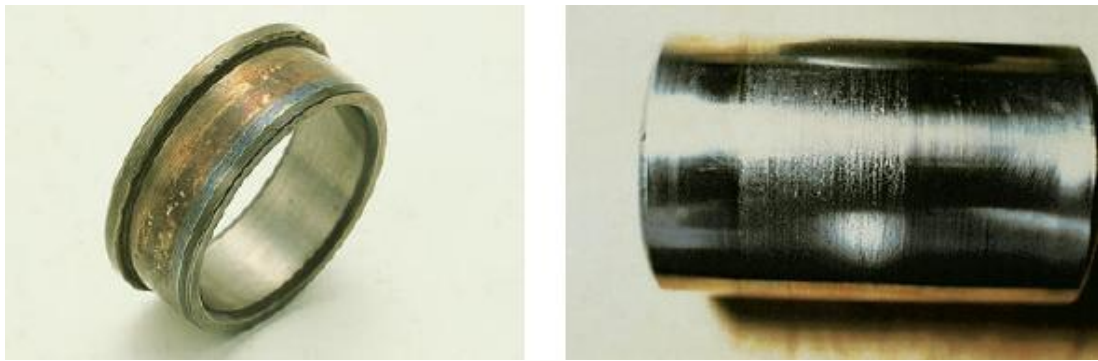
Figure 2-6 Bearing smearing [110]

This occurs when one ring is mounted out of line with the roller and cage assembly during the mounting operation. The surfaces of the rollers appear scored with discoloured areas at the start of the load zone in the raceways. Sometimes bearing rings can be so heavily loaded that they move relative to the shaft or housing. In such circumstances smearing, scoring and discolouration will appear on the outside surface of the outer ring or the surface of the inner ring bore.

- *Surface Distress*

The root cause of smearing and surface distress is the same: inadequate lubricant film. If, due to improper or inadequate lubrication between raceways and rolling elements the peaks of the surface asperities come into contact with each other small cracks form in the surfaces - known as surface distress. The initial stage of this type of defect is not visible to the naked eye, it increases gradually until shallow craters with crystalline fracture surfaces appear and

interfere with the smooth running of the bearing. Note that surface distress is not the same as fatigue cracks the latter originate beneath the surface and can result in flaking.

- *Corrosion (rust)*

Rust (chemical attack) will form if corrosive materials or moisture/water enter the bearing so that the lubricant no longer provides adequate protection for the steel surfaces of the bearings see Figure 2-7. Very quickly deep-seated rust will form. Corrosion is usually accompanied by an increase in vibration and wear, with subsequent increases in radial clearance or loss of preload.

Figure 2-7 Bearing rust and corrosion [110]

Any corrosive materials in the bearing will, given sufficient time, cause deep-seated rust (pitting). Usually, with roller bearings, grey-black streaks appear across the raceways coinciding with the rolling element spacing. As time progresses other surfaces will suffer pitting. Pitting is a serious defect so when pits occur, the bearing should be discarded before spalling has time to develop.

The layer of oxide forming the rust can act as a protective layer delaying further chemical reaction. Fretting causes the rust layer to be broken and, if allowed to develop sufficiently, these areas can become fracture notches. Fretting takes place due to a loose fit between bearing ring and shaft or housing, allowing relative movement. The result is that the bearing

rings may not be evenly supported and this produces a detrimental re-distribution of load in the bearings which exacerbates the process. Here, rust appears in the bore of the inner race or on the outside surface of the outer race.

2.4.2   Secondary Damage

- *Spalling (flaking)*

Pitting is due to the intrusion of foreign material, spalling occurs due to naturally occurring fatigue and so is also known as normal fatigue failure, see Figure 2-8. However, other factors can also cause spalling in bearings: preloading due to incorrect fitting, too great an external load, distortion of shaft or housing seating, and compression due to thermal expansion will all contribute to premature spalling. The depth of the crater with spalling can be between **20** μm and **100** μm and the crater can extend over as much as quarter to a third of the contact length of the roller bearing. For safety reasons the bearing is usually considered to have reached the end of its working life when the spall area exceeds about **6** mm^2 [106]. The decision is dependent on the application and in some circumstances the bearing it may continue to perform adequately. Spalling in bearings generates noise and vibration which indicates that it is time to replace the bearing.

For normal operation bearings are usually fitted to have a small clearance. For certain applications some rolling bearings have a negative clearance when mounted, to generate internal stress. However, some bearings can be accidentally preloaded as when a bearing race is fitted so tightly that it exceeds the internal radial clearance. If the internal radial clearance is lost there will be a rapid and possibly severe temperature rise and in such circumstances continuing operation will lead to spalling.
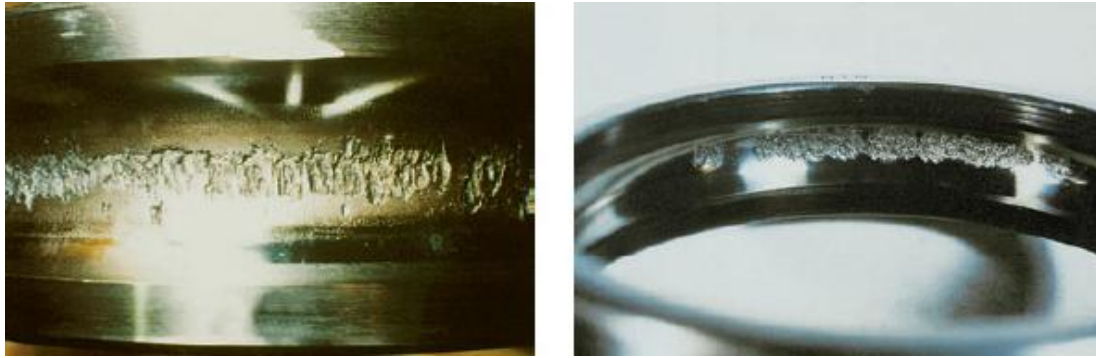
Figure 2-8 Bearing flaking [110]

Spalling will usually occur in the most heavily loaded zones and preload will be indicated by heavy rolling element wear paths in the raceways of both inner and outer races. Other significant causes of spalling are excessive drive-up on a tapered seating or taper roller bearings adjusted to give excessive preload.

It can happen that an element of the bearing and/or housing which should be circular is distorted. Usually the cause is the outer race being forced into place, but sometime it is the shaft which can be out of shape. This is a common defect in split housings and machine frames. The result is a deep spalling pathway and flaking on diametrically opposite sections of either bearing race.

Excessive load is another source of premature bearing failure due to consequent overheating, deep spalling and flaking, and heavy wear on ball paths. This fatigue spalling can occur over the entire face if the load is so heavy that the lubricant film reveals surface asperities, see Figure 2-9.

If a bearing is mounted incorrectly so that there is insufficient freedom of movement to accommodate thermal expansion, or where there has been excessive preloading of taper roller bearings, spalling due to axial compression can result. A common name for this is reverse loading failure. The spalling appears as a heavily marked raceway pathway pattern to one

62

side of the row of rolling elements. In some cases the appearance will be the same as for damage resulting from preloading.
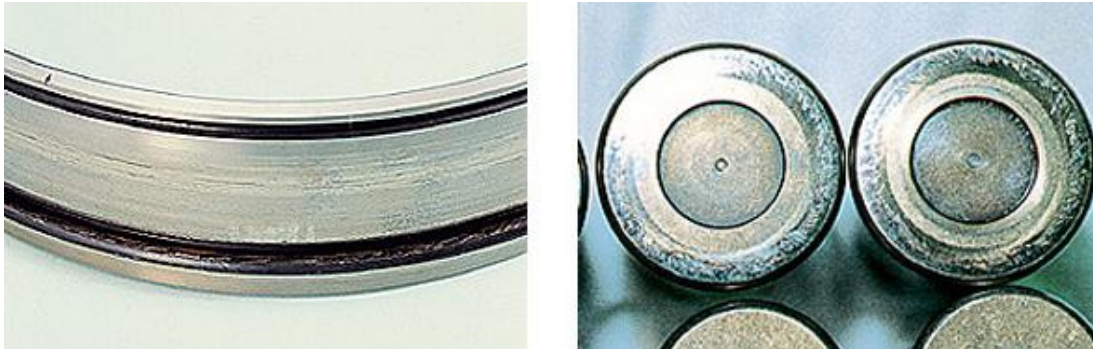
Figure 2-9 Bearing spalling [110]

Roller bearings can tolerate some misalignment but if the misalignment is excessive high edge stresses occur due to concentrated load, with subsequent fatigue spalling taking place on the bearing races. The greater the misalignment and the greater the speed the more rapidly the spalling progresses.

- *Cracks*

There are many reasons why cracks occur in bearing rings, the most common cause is rough treatment such as heavy hammer blows when the bearings are being mounted. The hammer blows can cause fine cracks which can grow and cause fragments of the ring to break-off when the bearing is put into operation, Figure 2-10 shows bearing cracks.

Another common cause of ring cracking is excessive tensile stresses in the rings of a tapered seating or sleeve due to excessive drive-up. When the bearing is put into service cracks appear.

For fretting corrosion, cracks in the outer race are generally longitudinal and for the inner race the cracks intersect. Lack of adequate lubrication generally generates cracks at right angles to the direction of slide.

Figure 2-10 Bearing cracking [110]

- *Bearing Cage Damage*

It is known that excessive vibration, speed, wear, and blockage cause cage breakage, see Figure 2-11, fortunately it does not happen often, but when it does it is usually difficult to identify the precise cause since the other components of the bearing are also damaged.

Figure 2-11 Bearing cage damage [**110**]

Vibration can cause accelerated fatigue and cracks to form in the cage material which develop over time and lead to cage failure.

Running bearings run at speeds well in excess of specification subjects the cage to substantial forces that can lead to fractures. The material of the cage is less hard than the other components of the bearing and rapid wear may be caused by insufficient lubrication or by abrasive particles. The cage is thus the first component affected when there is too little lubrication or the lubrication degrades.

It is possible that fragments of hard particles of flaked material may become wedged between the cage and a rolling element. When this happens over-rolling occurs or the rolling element may no longer rotate at all, both cases can lead to cage failure. Bearing cages subject to extreme acceleration and deceleration will experience considerable force between the surfaces in contact and heavy wear can result.

- *Scoring (galling)*

Metal-to-metal contact between the roller end and guide rib face can be another result of inadequate lubrication. This causes scoring which can change the relative geometry of the guide rib face and end of the roller with possible severe results: rollers can become distorted and broken, or may become welded to the races or the rib, causing the bearing to seize and severely damaging other components. Fortunately scoring on the roller ends is easily detectable in a roller bearing assembly.

Inadequate lubrication can also result in heating of the bearing (change in hardness) with damage to the roller ends rib and inner ring raceway.

- *Rollout*

Highly overloaded bearings experience considerable stress deep inside the race which may exceed the strength of the relatively soft core which will then plastically deform in the axial direction causing bearing rollout as shown in Figure 2-12 because the housing constrains it in the radial direction. As the core moves axially relative to the outer ring it causes circumferential cracks in the load zone.

Figure 2-12 Bearing rollout [110]

2.4.3   Fatigue Spalling

The primary failure mode for a healthy roller bearing is spalling of the bearing element due to local fatigue caused by the repeated application of stresses **[104]**, see Figure 2-13. Bearing material is not truly homogeneous and so is not everywhere equally resistant to failure which will occur at the weakest point. Thus, apparently identical specimens will show variation in time to failure even when operated under supposedly identical conditions. Nevertheless, major increases in bearing life and reliability have resulted from improvements in bearing materials, manufacturing technology and lubrication.

Figure 2-13 Bearing fatigue [110]

2.5   Dynamic Responses of Roller Bearings to Local Defects

A loaded healthy bearing is subject to many complex static and dynamic forces and moments so even a defect-free bearing will generate vibration. The static forces will include, for example, any preloading, the weights of the shafts and other moving components, etc. The dynamic forces will include centripetal forces acting towards the centre of rotational motion, frictional and traction forces due to the relative movement of surfaces in contact and fluid pressure, etc. Typically a healthy bearing rotating at constant speed and subject to a constant load will be subject to forces which are in quasi-equilibrium. However, when the rolling element meets a local surface defect, there is a sudden elastic deformation of the element due to the transient force experienced. These transient forces are detected by the resulting movements of the bearing components. The resulting movements will not normally be simple but can include impact and oscillatory contact between any two of roller, raceway and cage, as well as slippage between cage and roller. Nevertheless the periodic nature of the movement of the bearing will impose a periodicity on the signals generated, and this strongly suggests that frequency domain analysis of the signal will be a powerful tool in defect detection.

Many analytical dynamic models have been developed to represent the motion of the ball bearings, raceways and rolling elements generated by defects [108, 109]. Such models are important because the inaccessibility of bearing components makes measuring their angular and linear motion a difficult task with significant associated errors. In addition, for most rotating machinery simply detecting of bearing damage is not enough; it is just as important to assess its effect on bearing life.

2.6   Characteristic Symptoms for Bearing Failure

Increasingly researchers into vibration analysis of bearings are dividing the defects into point (or localised) defects and generalised defects [22]. This research focuses on local defects (scratches) made on the bearings and which are meant to be representative of spalls, pits or localised damage that appears on raceways and rolling elements. A major benefit of local or single-point defects is that they produce a characteristic frequency which can be calculated from the speed of rotation and geometry of the bearing.

Generalized roughness refers to such defects as deformations that occur during manufacturing, improper installation or wear. Such defects typically generate broadband machine vibrations and so require specialized techniques for their detection. As will be explained below crude time-domain parameters cannot be used directly to locate a point defect, for instance whether the fault is on the inner or outer race, or the roller.

Rolling element bearings have a unique frequency corresponding to the dynamic behaviour of each bearing component: the outer race characteristic frequency (f_o), the inner race characteristic frequency (f_i), the roller characteristic frequency (f_r) and the cage characteristic frequency (f_c).

Calculation of the fundamental point defect frequencies of rolling element bearings is useful because they can indicate what is failing in the bearing. For an angular contact ball bearing in which the inner race rotates and the outer race is stationary, the four characteristic fundamental defect frequencies are given in Equations 2.1 – 2.4 [109]. The number of balls in the race is n, f_s is the shaft rotational frequency, BD is the ball diameter, PD is the pitch circle diameter, and β is the contact angle.

For convenience, we assume that there is only one fault at a time in the outer race, the inner

race and one rolling element of the bearing. Each time the outer race axially rotates, an impulsive force is generated by the impact for every rolling element. The frequency of the impulsive force corresponding with an outer race fault will be f_o; the characteristic fault frequency of the outer race. In an exactly similar way for a single fault on the inner race the impulsive frequency will be f_i; the characteristic fault frequency of the inner race.

Outer race frequency:

$$f_o = \frac{n}{2} f_s \left[1 - \frac{BD}{PD} \cos \beta \right] \qquad (2.1)$$

Inner race frequency:

$$f_i = n \left(f_s - f_c\right) \qquad (2.2)$$

For a single fault on one rolling element, each time the rolling element rotates axially there will be an impulsive impact on both inner and outer races. The corresponding impulsive frequency f_r is the characteristic fault frequency of the rolling element.

Roller characteristic frequency:

$$f_r = \frac{PD}{BD} f_s \left[1 - \left(\frac{BD}{PD}\right)^2 \cos \beta \right] \qquad (2.3)$$

Similar to the above analyses, when there is one fault in the cage, the characteristic fault frequency of cage f_c can be calculated as:

Cage frequency:

$$f_c = \frac{1}{2} f_s \left[1 - \frac{BD}{PD} \cos \beta \right] \qquad (2.4)$$

As explained above, for most applications, the outer race housing is fixed in a housing and the inner race rotates at the same speed as the drive shaft so that each of the bearing defect frequencies are non-integer multiples of running speed, and the cage frequency is always lower than shaft rotating frequency.

Bearing type N406 is chosen to be used to conduct this research to represent a study on bearings' fault detection and diagnosis on the bases of its widespread in industrial machinery as well as its capacity of carrying heavy load. However, the author acknowledges that some other kinds of bearings could have been used for this thesis, nevertheless, that also could be a topic for future work.

For the SKF bearing type N406 as measured in the laboratory $n = 9$, ball diameter is $14\,mm$ and pitch circle diameter is $59\,mm$. Knowing the shaft rotational frequency to be 24.3 Hz; and taking β to be 0 as there is no axial load, we used the above equations to calculate the fundamental fault frequencies corresponding to the defects seeded individually into the bearing. Table 2-3 shows the experimentally measured characteristic frequencies at full shaft speed and those obtained using the above equations. It can be seen that the two sets of figures are in close agreement.

Table 2-3 Calculated defect frequencies and corresponding measured values for bearing type N406

Defect position	Measured Frequency (Hz)	Calculated Frequency (Hz)
Inner race	136.20	135.48
Outer race	83.50	83.52
Rolling elements	97.34	96.63
Cage	9.025	9.28

Table 2-3 shows frequency spectra for the vibration signals from bearings in which distinctive peaks can be seen for the faulty bearing compared to that of the healthy bearing.

Bearings with fatigue spall generate impulsive vibrations with characteristic frequencies and development to failure can be tracked by following the increasing prominence of one or other of these, as will be demonstrated in Chapter 3. Although the presence in the spectrum of a characteristic defect frequency is a good indicator of the location of a defect, automatic detection can only be achieved if these peaks can be detected in the presence of strong background noise. When a roller passes over a bearing defect in the races or cages, pulse-like excitation forces are generated that result in one or a combination of these defect frequencies [82].

CHAPTER THREE

Vibration Signal Analysis Techniques for Bearing Faults

This chapter introduces vibration measurements as a method of condition monitoring and defect detection in roller bearings. It reviews time domain, frequency domain and time-frequency domain analysis of the vibration signal from roller bearings and compares their performance regarding to fault detection and diagnosis. Various commonly used statistical parameters are described and their usefulness for rolling element bearing faults detection commented on.

3.1   Introduction

The use of vibration as a non-intrusive diagnostic tool is well established in engineering, has been widely applied to the CM of rotating machines and can be used to obtain information about sub-systems which would otherwise be inaccessible [111, 112]. The ability of vibration measurement to detect and diagnose a wide range of faults in machine elements is a good reason why it is often chosen as a preferred and cost-effective method for CM [113]. It is widely accepted that vibration measurement is able to identify 90% of all machinery faults by the change in vibration signals which they produce [7] and is reported to be the chosen method by most researchers to study bearing defects [114]. Given the faults to be studied in this project and the impacts they generate, vibration measurement has been chosen as the measurement method.

Local discontinuities in the material (this might be a deliberate scratch) forming the different surfaces of the roller bearings generate impulses in the vibration signal as described in Section 2.6. As these impulses pass through the bearing to the sensor they will be modulated by their transmission paths and be superimposed onto background vibration. In the early stages of defect development (e.g. spalling) the amplitude of the impulse signal generated will not be large and due to the natural damping provided by, e.g. the lubricants quickly decay before the next impulse is generated. However, patterns (in both time and amplitude) do exist in the signal which allows identification of the severity and location of the defects.

The vibrations of a machine during normal machine operation will change with changes in the dynamic properties of the machine, which could be an indication of component failure and simultaneously point to the problem. A healthy machine operating properly will generally produce a constant level of small amplitude vibrations. When a fault is developing

73

in a component, but is no longer incipient, it will usually lead to significant changes in machine dynamics and generate greater vibration, often with different spectral patterns which can be used to identify the fault [115, 116, and 117].

Vibration monitoring of rolling element bearings has consistently produced good results because of developments in signal processing techniques. Pattern recognition techniques have been investigated and shown to be able to diagnose faults in machines for various operating conditions [118, 73]. In this thesis, we extend pattern recognition analysis using TESPAR with the aim of increasing the reliability and sensitivity of the method [73]. TESPAR is described in some detail in Chapter 4.

When using pattern recognition techniques for fault detection and diagnosis, the method used for feature extraction, and the features selected, can be critical [119]. Here, knowledge of the system being tested will suggest the features to be used, and the better the system is known, the easier the diagnosis should be. Ideally, the features selected will uniquely represent the characteristics of the system being investigated, but such selection is not necessarily straightforward and will depend on the system being monitored. The features selected for monitoring should be robust to changes in operating conditions and any noise generated.

3.2   Vibration Measurement

CM can be successful only if the measurements taken are sufficiently accurate; i.e. the sensors are properly mounted, the instrumentation is properly calibrated, data collected from the same location under conditions which are largely identical save for the parameter being investigated [120]. Then the spectrum of the healthy bearing is compared to the one where a fault may be developing and the differences in the spectra used to detect and diagnose any possible fault. Such spectral comparisons are a very common technique in CM.

74

Nearly all available transducers measure the vibration movement of one or more surfaces of a machine, using displacement, velocity or acceleration. Today, with the development of modern electronics the accelerometer is the most used transducer. The signal is integrated once to provide velocity and integrated twice to provide displacement [121]. Because of their relative physical size accelerometers have a wider frequency range than their equivalent velocity or displacement sensors. Accelerometers were the sensors used to collect the data for this project.

The piezo-electric accelerometer is the most popular vibration measurement transducer in use today [122]. These are light in weight (can be less than 10 g) with a consequent wide frequency range (20 kHz is readily obtainable), have a good dynamic range (typically from about 0.020 to over 106 ms^{-2}) and will withstand both high and low temperatures (-20 to +120 0C for most common piezo-electric materials) [123]. The signals from the accelerometer are processed in suitable ways to extract those aspects of the signal which can be used to detect and diagnose any faults in the machine.

3.3   The Theory of Vibration Analysis Techniques

3.3.1   Time Domain Parameters Theory

- *Root Mean Square, Peak Value and Crest Factor*

The RMS value is a measure of the energy level of vibrations. Peak designates the maximum amplitude of vibrations. For sinusoidal signals the peak value, $Peak = \sqrt{2}.RMS$, but no such simple relationship exists for real signals. For digital sampling the random vibration signal $x\{t\}$ can be described as a discrete time series $\{xi\}(i = 1,2 \dots N)$, and the RMS and peak values are given by:

$$RMS = \sqrt{[(\sum_{i=1}^{N} x_i^2)/N]} \qquad \textbf{(3.1)}$$

$$Peak = max \{x_i\} \qquad \textbf{(3.2)}$$

Where there are N samples and xi is the value of the ith sample.

Attempts have been made to assess fault damage by detecting changes in peak values and counting the number of peaks above a certain threshold in the measured signal and a signal with a Gaussian amplitude distribution, but the results were not promising [124] and today peak values are rarely used except in conjunction with other more reliable parameters.

RMS is a simple measure of the energy content of the vibration signal and is widely used to detect deteriorating bearings. With incipient bearing damage, the impact signals occur as discrete peaks but the RMS value of the signal remains virtually unchanged [125]. An obvious statistical parameter to use here is the non-dimensional Cf:

$$\textbf{Crest facto} \quad = \quad \frac{\textbf{Peak}}{\textbf{RMS}} \qquad \textbf{(3.3)}$$

Cf is will be an effective indicator of the spikiness of the vibration signal [126]. The larger Cf the more and sharper will be the peaks in the signal, and the more likely the signal contains repetitive impulses. However, peak values are very sensitive to random variations so it is essential to perform an averaging process [125].

As the fault/damage becomes more severe the RMS value increases, but without necessarily increasing the peak value, i.e. the Cf value rises to a peak after which it falls. Thus, the Cf value could be an indicator of the severity and stage of a bearing defect. However, single Cf

76

values cannot be used for detection of bearing faults because even with averaging, Cf values will vary over a wide range depending on such operating factors as load and speed.

Combining the measured values of RMS, peak and Cf could be an effective method of determining the presence of a fault, but only if they can be compared with the baseline healthy values for the given system operating under the same conditions. Normalized values of RMS, peak and Cf have been proposed [123] to allow for the operational condition and non-defect induced vibrations:

$$R_V \quad = \quad \frac{RMS}{RMS_0} \qquad\qquad (3.4)$$

$$P_k \quad = \quad \frac{Peak}{RMS_0} \qquad\qquad (3.5)$$

Where $RMSo$ is the reference value for a healthy bearing, for roller bearing used in fixed machinery, $RMSo$ would be the value for bearings in good condition under ordinary operating conditions.

- *Kurtosis and Impulse Factor*

The first and second moments of a distribution of sample values are the mean value and the standard deviation of the distribution respectively. The third moment is skewness and the fourth moment is kurtosis. For real signals the mean and skewness will usually be close to zero, suggesting that the amplitude distribution is symmetrical. On the other hand, the higher the order of the even moment the more sensitive it is to impulses present in the signal.

Time-domain statistical parameters in addition to those described in the previous section are used to detect the presence of incipient bearing damage. Commonly used non-dimensional vibration amplitude parameters are the crest factor (Cf), kurtosis (Kv) and impulse factor (If) [127]. The impulse factor is a measure of how "sharp" the PDF of the amplitude distribution is. A few large spiky peaks in the time-domain vibration signal will give a large value for If, but a flattish time-domain vibration signal will have a low value for the If.

Non-dimensional measures such as clearance factor (Clf) have also been described [127] but as these are not in common usage they are not discussed here. Kv and If have been described [111] as:

$$Kv = \frac{[\sum_{i=1}^{N}(x_i - \bar{x})^4]/N - 1}{(\sigma^2)^2} - 3 \qquad\qquad \textbf{(3.6)}$$

$$If = \frac{max\ x_i}{\sigma} \qquad\qquad \textbf{(3.7)}$$

Where the signal consists of N samples of amplitudes $x1, x2, \dots xN$. σ is the standard deviation of the N samples, $\bar{x}$ is the mean value of the N samples

The factor "-3" is introduced into the term for kurtosis to normalise the expression so that $Kv = 0$ for the Gaussian/Normal distribution. If $Kv < 0$ it means the distribution is flatter than the Gaussian, if $Kv > 0$ it means peaks in the spectrum are sharper or more pointed than the Gaussian. A high value for kurtosis tends to mean infrequent large peaks rather than frequent modestly sized peaks.

Kurtosis has been chosen as a parameter in this work as a compromise between insensitive lower moments and too sensitive higher moments. It has been found that for healthy bearings the distribution of amplitudes tends to follow a Gaussian or Normal distribution with a kurtosis value of about 0

78

[128]. It has been argued that kurtosis while sensitive to incipient faults in bearings is insensitive to bearing operating conditions and a value of **3.0** as signifying incipient bearing damage has been suggested [128].

Kv and Cf are measures of the spikiness of the signal and are largely independent of the average magnitude of the signal. These two parameters have values of about **0** for healthy bearings and increase rapidly when a fault representing fatigue spall is introduced..

However, Kv and Cf do not increase linearly with the damage as it progresses. With an inner race defect, for example, the initial extent of a circumferential crack (fatigue spall) is likely to be significantly less than the distances between two rollers. Here the bearing is subject to a discrete impact every time a roller passes over the damaged area with a corresponding peak in the vibration signal. As the crack propagates around the inner race the vibration becomes less impulsive and more continuous and the vibration signal becomes more random. In the latter case the values of both Kv and Cf will decrease, and eventually will become close to those for the healthy condition. Thus, these two parameters cannot always provide accurate information on the level of fault damage. However, a significant increase of overall vibration energy – as measured by the RMS - will usually accompany more serious damage. Hence a combination of Kv and Cf with the RMS could be a suitable set of statistical parameters for identification of the fault level.

It has been said [128] that: If gives similar results to Kv and , and that all of Cf, If and Kv are sensitive to incipient fatigue spalling, and this opinion has been confirmed [129]. Kv was found to be most sensitive parameter for onset faults but the least robust to operating conditions. It shows that these parameters are sensitive to early spalling but may lead to inconsistent results if used in isolation.

The statistical parameters obtained from the time-domain of the vibration signal can be calculated from the signal with no frequency filtering, or the signal can be passed through a high pass filter in an attempt to reduce unwanted low frequency background noise; or even filtered into frequency bands to reduce unwanted signal noise. The time domain parameters are then calculated for each band.

For established fault detection the *RMS* value of the time domain of the vibration signal accompanied by other statistical parameters works well. Nevertheless, for the detection of incipient defects all time domain techniques, whether supported by frequency filtering or not, suffer from the same essential drawback: the energy level of the signal must be discernible over the "noise" from all other components of the machine.

3.3.2 Frequency-Domain Techniques

The Fourier transform (in one of its many forms) is by far the most common technique for transforming a varying time-domain signal into its frequency-domain representation. The frequency-domain of the signal, its spectrum, presents the energy in the signal as a function of frequency and allows the contribution of individual components to be seen. For example, when using time-domain analysis the signal from a small fault developing in a roller bearing will be hidden in general bearing noise, but in the frequency-domain there will be a distinctive peak in the spectrum at the characteristic fault frequency. The challenge is to detect the presence of an incipient fault at the earliest possible time.

Spectral peaks appear at characteristic frequencies which can be harmonics of the fundamental, and these signals can be enhanced by filtering out other undesired signals. Thus frequency-domain analyses often use low-pass, high-pass, and band-pass filtering. Today enveloping is becoming widely used because the bearing defects can impose distinct

characteristics on the spectrum [130]. However, efficient and effective use of filtering and enveloping techniques requires prior knowledge of the frequency band of interest and this may not be known beforehand.

To address the limitations of traditional frequency domain analysis for early defect detection, other analysis techniques in the frequency-domain have been developed in recent years including the digital Fourier transform (DFT) [131, 132] which has largely replaced the analogue Fourier transform due to the universal availability of the digital PC. A powerful new technique, bicoherence (also known as bispectrum) analysis is being used in CM for the processing of signals, in particular where traditional linear analysis cannot provide sufficient information. Bicoherence is most efficacious when used to investigate faults in rotating machinery which show themselves as non-linear transformations of the vibration signal with non-linear coupling between frequencies. Bicoherence analysis detects and quantifies the presence of the non-linearity in the signal and thus can indicate the severity of the fault [133, 134].

However, bicoherence analysis requires large data sets to produce a reliable evaluation and thus is computationally intensive and not considered appropriate for on-line applications. Today the most popular method for deriving the frequency-domain signal of e.g. roller bearings is to transform the time-domain signal using the discrete fast Fourier transform [135].

3.3.3   Envelope Analysis

Envelope analysis (amplitude demodulation) is considered a powerful and reliable technique for the detection of faults in rolling element bearing and is gaining in popularity for detection and diagnosis of such local faults as spalling and/or cracks in rolling element bearings [136].

Envelope analysis dates back over one hundred years to the "cat's whisker" detectors used in the early years of radio communication, so its mathematical basis is well established. In amplitude modulated (AM) transmission, the signal carrying the information is used to modulate the amplitude of the carrier wave transmitted by an antenna. With radio transmission, the carrier wave – which serves only to carry the transmission of information from one place to another – will be a radio frequency wave and the modulating wave will usually be at audio frequencies below about 20 kHz.

Amplitude modulation can be viewed most simply in terms of the time-domain of the signal. Using an analogue description for easier understanding we can say that if the carrier wave is $yc = Ac.coswct$ and the information carrying, modulating wave is $yi = p(t)$ then the modulated wave is $Yam(t) = yc.yi = p(t)Ac.coswct$ [136], see Figure 3-1.

How does this relate to identifying local faults in rotating machinery, particularly bearings? It is well known, and described earlier in this chapter, that whenever the defect in a rolling element bearing strikes or is struck it produces high-level pulses of very short duration.
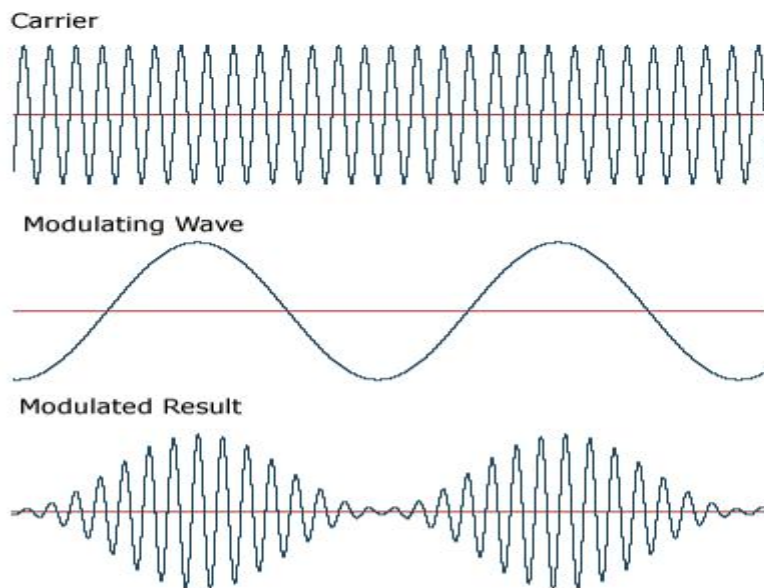
**Figure 3-1 Simple example of modulation of carrier wave
to produce a modulated wave (here signals have amplitudes = 1) [136]**

82

Every time a rolling element passes over the defect an impulse is generated, a process which is controlled by the geometry of the bearing and the speed of the machine so that the rate at which the impulses are produced will be characteristic. The spectrum of a repeated impulse, repetition period $= T$, is a discrete line spectrum throughout the frequency range of interest with a spacing of $1/T$ Hz. Such a signal is quite different from those induced by faults in other machine components, which should mean that the characteristic repetition could be an easy identifier of a bearing fault [137].

In fact many bearing manufacturers provide a bearing fault frequency calculator. The user inputs the bearing type number (which identifies the dimensions of the bearing, number of balls or rollers, etc.,) and the speed of the shaft and calculator outputs the corresponding potential bearing characteristic frequencies (BCFs) as described and discussed in Section 2.5 and Table 2-3.

The frequencies in Table 2-3 are truly constant only for ideal bearings because real roller elements not only rotate but also slide and the speed of rotation of the shaft can vary and it should be checked using a stroboscope or similar device [138]. If the first harmonic changes by 1 Hz the Nth harmonic will change by N Hz. In addition there are likely to be other local faults in the bearings each with an associated frequency. It is common to find that the harmonics smear and higher harmonics may even merge. Notwithstanding these difficulties the simple frequencies mentioned remain a valuable source of information on bearing fault identification [139].

Fundamental to the bearing envelope analysis is the fact that each time a defect in a rolling element bearing under load makes contact with another bearing surface, a short duration vibration impulse is generated, and the shorter the duration of the pulse the wider the

frequency range over which the impulsive energy is spread. Thus the impulsive energy is spread over a wide frequency band at a very low level and it is this wide distribution of energy which makes bearing defects so difficult to detect by conventional spectrum analysis – particularly in the presence of other vibration sources.

Over most frequencies the bearing structure acts to damp the impact energy but, nevertheless, for each impact some force is transmitted into the machine and, because the signal is wide band, it will excite resonances along its transmission path to the vibration transducer. Nevertheless any vibration measured on the machine's surface will contain harmonics of the repeated impact. The transducer will also pick up signals from other vibration sources within the machine, general background noise and even vibration from nearby machines [138]. The resulting complexity of the measured signal makes detection of the periodicity of the original impulse signal difficult, especially in the early stages of fault development.

In practice, then, the amplitude of the induced vibration at the characteristic bearing fault frequencies picked up by an accelerometer on the machine frame will be very small compared to other sources, with all signs of the BPFs hidden in what looks like noise even though in this case the severity of the bearing defect was rated "very high" [139].

Consider how, in principle, amplitude modulation could be used to enhance the signal. The impulse generated by the impact of the fault will excite resonances in the machine. Select the resonance which has the largest amplitude [140]. Experience has shown that these frequencies will be one of the machine modes and will often be in the region of 8-12 kHz [141]. In an ideal case where the signal is band pass filtered to remove any extraneous frequencies the output from the accelerometer would look very similar to the bottom plot in

Figure 3-1 where the carrier is the resonant frequency of the machine and the envelope is the characteristic bearing frequency corresponding to the particular fault.

The results of envelope analysis will depend on the frequency band chosen and the band pass filters must be selected to include those frequencies where the defect signal is strongest [79, 141, 152]. However, this band may not be known initially and also may change as the bearing operating condition changes. Thus envelope analysis is likely to be more successful if the operator has prior knowledge of the carrier frequencies before selecting the band pass filters, however, a degree of trial and error in determining the best of the resonant frequencies to use is sometimes necessary [138]. In practice, however, the raw signal can be very noisy and even after band pass filtering may still contain undesirable components which increase the difficulty of identifying the envelope spectrum. Envelope analysis is unsuitable for use with bearings with severe faults as the damage is likely to have spread to such a degree that the dynamic response is more random noise than clear impacts.

Figure 3-2 is a schematic diagram representing the digital demodulation process. The raw signal is filtered to remove unwanted low frequency background noise. The usual practice is to view the signal caused by the bearing fault in a frequency range associated with a machine or bearing resonance [138]. It is recommended that the structural resonant frequency is selected as the centre frequency of the band pass filter, thus the band pass filter will typically be 500 Hz to 10 kHz [142] though the upper limit may be higher if required to increase the signal-noise ratio. The band pass filter also removes unwanted high frequency noise. This could be unwanted machine and structural resonances and are separated out by passing the raw vibration signal from the bearing through a high-pass filter, which is set at about 10 - 12 kHz, a frequency determined by empiric testing to identify the most suitable resonant frequencies of the test system [143].

To recover the modulated signal, full wave rectification is applied to the band pass-filtered signal; If this rectification is not done the modulation enveloped would, after filtering, cancel itself out [144].

The remaining unwanted high frequency components the signal is low pass filtered. A spectrum analyser (e.g. FFT for time-invariant systems) is used to extract the spectrum which should contain the BPFs. It is because of the inability of the FFT to detect faults which exhibit non-stationary that there is a need to seek alternative methods such as wavelets or TESPAR.
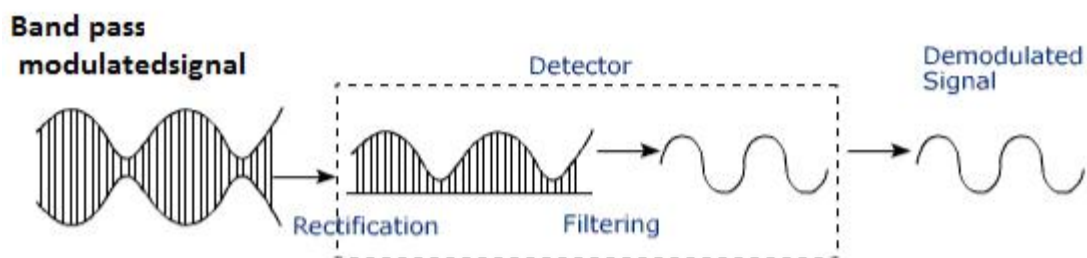
Figure 3-2 Schematic of digital demodulation [80]

A researcher used the envelope analysis routine contained in MATLAB to detect faults introduced into the outer race of type 6209 ball bearings [145]. The sampling frequency was 20 kHz. The drive motor had a constant speed of 24.55 Hz. The band pass filter was 500 Hz – 7.5 kHz, and the envelope spectrum obtained is shown in Figure 3-3. The bearing pass frequency of 89.3 Hz and its harmonics are clearly visible. The presence of harmonics can and should be used as confirmation of the fundamental frequency.

The first peak is visible at 24.55 Hz which corresponds to the shaft rotational speed. The bearing characteristic frequency of 89.3 Hz can be clearly seen – the calculated value was 93.0 Hz but a slide factor of 0.96 was included to give a "calculated frequency" of 89.7 Hz.
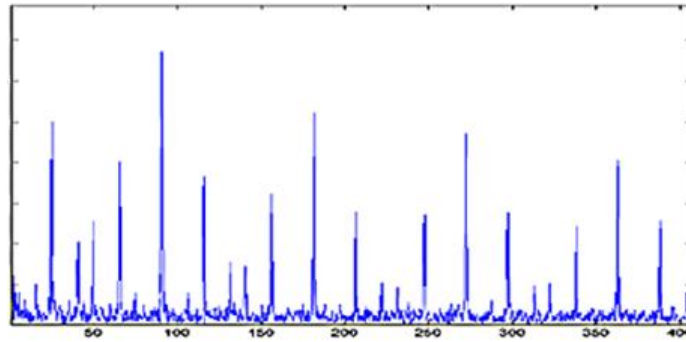
**Figure 3-3 Envelope spectrum (Hz) for bearing
with outer race fault** [80]

The results presented show that the version of MATLAB used is important [139, 141]. For example, in Figure 3-3 there is a large unidentified peak at 125 Hz (and numerous smaller peaks) in the MATLAB results which is not present with a dedicated instrument such as the Bruel & Kjear Multianalysator PULSE 3560-B-110. Incidentally, one researcher [141] gives a good example of averaging to reduce random errors, rather than average the time-domain signal he averaged the frequency spectrum.

FFT with envelope analysis is definitely able to reveal bearing defects but the results will not be consistent for all types of defects. If the bearing is subjected to a constant (radial) load a fault in the outer, stationary race of the bearing will produce uniform impulses.

However, a fault on the rotating inner race will be subject to a varying load, where the frequency of the variation in load will be the frequency of rotation of the inner race. But the fault moves in and out of the bearing load zone and while in the load zone, the defect produces vibration at the ball pass frequency, but when it is out of the load zone the amplitude of the vibration is reduced.

Consequently, the envelope signal will be amplitude modulated as shown in Figure 3-4 and this will appear in the envelope spectrum. As a result of considerations such as this [138] it

87

has been pointed out that the envelope analysis is able to give excellent results for bearings having outer race defect but that detection of roller and inner race defect may not be so easy.

Analysis of the envelope signal in Figure 3-4 will show sidebands either side of the modulation frequency and the presence of such sidebands may be a method of bearing fault identification.

There are different ways to extract the envelope; traditionally band pass filtering, rectifying and low pass filtering was used to carry out the demodulation. However, modern digital techniques tend to use the Hilbert transform for effective and efficient envelope detection [145].
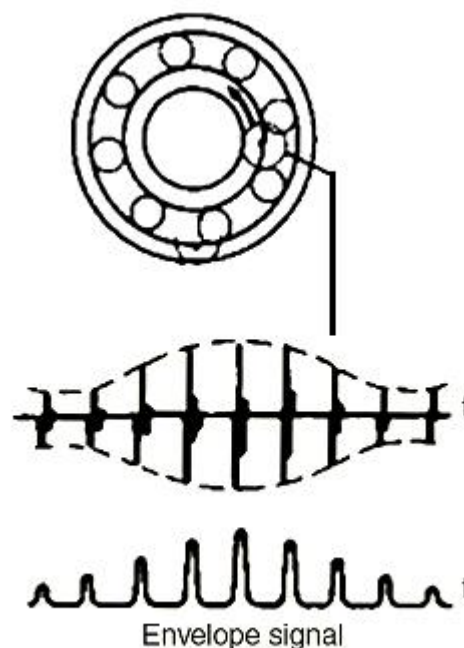
Figure 3-4 Amplitude modulation of the envelope signal
for fault on inner race [145]

Hilbert transform is a multiplier operator that provides an output with a 90^0 phase shift from the input [117] [146]. All negative frequencies of a signal get a $+90^0$ phase shift and all positive frequencies get a -90^0 phase shift. In the simplest case, if the input is a cosine wave the Hilbert

transform will create a sine wave out of it without any change in amplitude of the wave. However, an important consequence of this is that if the input signal is modulated the output will be the modulating signal, the so-called "envelope of modulation". In the case of the roller bearing with a fault in, say, the inner race, the output will be a series of pulses occurring at the roller passing frequency.

3.3.4   Time-Frequency Analysis - The Wavelet Transform

The Fourier transform is probably the most widely used technique to transform the time-domain signal into the frequency-domain but in the process time information is lost. The Fourier transform of a signal contains no information as to when a particular event occurred. This is not a problem for stationary signals which repeat themselves, but can be important when investigating/monitoring the growth of faults where transitory phenomena often constitute the most significant part of the signal.

In an effort to correct this deficiency, the Fourier transform was adapted to analyze only a small section of the signal at a time [147] - a technique called windowing the signal. Essentially Gabor divided the time-domain signal into a series of contiguous sections and the signal contained in each section (or window) is then transformed into the frequency domain. Of course, mathematically the process is not so simple, as the window is a multiplying function which must be tailored to avoid introducing spurious results.

Today a large number of such windows are available and most have the property that the initial and final values are close to zero so that the transform does not "see" the time-domain signal as a series of step functions. This process is known as the Short-Time Fourier Transform (STFT) and produces a series of spectra, one for each window. As the duration of

each window is known the spectra can be layered into a three-dimensional plot (frequency, time and amplitude).

The two drawbacks with this method are: (i) Once a particular window size has been selected it cannot then be changed but many signals can require more precision in determining frequency at some times than at others, (ii) The uncertainty principle means that good resolutions in both time and frequency-domains cannot be achieved simultaneously [148]. For example, if the signal to be analyzed is of short duration, obviously a narrow window should be selected, but the narrower the window the wider the associated frequency band and the poorer the frequency resolution.

The logical step was to develop a transform procedure which incorporated variable-sized windowing. Wavelet analysis allows the use of windows of longer duration when more precise frequency information is needed, and shorter duration when more precise time information is required. Because the term frequency is reserved for the Fourier transform we do not speak about time-frequency domains when discussing wavelets, instead we speak about time-scale representations where scale can be thought of as an inverse function of frequency.

The use of wavelet analysis is attracting considerable attention as a tool for diagnosis of the condition of bearings. Unlike the Fourier transform which expresses the time-domain signal in terms of sine and cosine functions, wavelet analysis expresses a signal over the whole spectrum in terms of wavelet functions of different scales which makes it more useful for the extraction of transient features contained in a signal. Because wavelet analysis uses wavelets of variable scales it can examine the entire spectrum when extracting a defect signal, and so

does not require a signal with a dominant frequency band as is needed for frequency-domain filtering.

Thus the windows used with wavelet transforms have the extremely important and useful feature that the width of the window can be changed ("scaled") and the transform is determined for each spectral component. The initial or so-called mother wavelet is the function $\mathbf{W}_{1,0}(\mathbf{t})$ and is defined for a certain time interval, with $\mathbf{W}_{1,0}(\mathbf{t}) = \mathbf{0}$ outside the time interval. To show that the window progresses with time the mother wavelet is written as $\mathbf{W}_{1,0}(\mathbf{t} - \mathbf{u})$, where u is the term that translates or shifts the window along the signal. With a time-domain signal u will have units of seconds.

The scaled version of the initial template is $\mathbf{W}_{s,u}(\mathbf{t})$ where s is a simple mathematical operation that either extends or compresses the window. Generally, the template function $\mathbf{W}_{s,u}(\mathbf{t})$ can be expressed in terms of the mother template function $\mathbf{W}_{1,0}(\mathbf{t})$:

$$\mathbf{W}_{s,u}(\mathbf{t}) = \frac{\mathbf{1}}{\sqrt{\mathbf{s}}}\, \mathbf{W}_{1,0}\left(\frac{\mathbf{t} - \mathbf{u}}{\mathbf{s}}\right) \qquad \textbf{(3.10)}$$

Where $\frac{1}{\sqrt{s}}$ is for energy normalization, so that the transformed signal will have the same

energy at every scale.

$$\int_{-\infty}^{\infty} \mathbf{W}_{s,u}^2(\mathbf{t})\mathbf{dt} = \int_{-\infty}^{\infty} \mathbf{W}_{1,0}^2(\mathbf{t})\mathbf{dt} \equiv \left\|\mathbf{W}_{1,0}(\mathbf{t})\right\|^2 \qquad \textbf{(3.11)}$$

In a linear signal space, the set of all template functions $\left\{\mathbf{W}_{s,u}(\mathbf{t})\colon \mathbf{s} \geq \mathbf{0}, \mathbf{u} \in \mathbf{R}\right\}$ forms a continuous frame Γ_c spanned by the scale s and time u. sampling at $\mathbf{s} = \mathbf{k}\ \text{and}\ \mathbf{u} = \mathbf{mkL}$, where k and m represent discrete format of the continuous s and u, a discrete function $\mathbf{W}_{k,m}(\mathbf{t})$ (which is a simplified notation for $\mathbf{W}_{k,mkL}(\mathbf{t})$) is obtained. The set

91

$\{W_{k,m}(t): k \text{ or } k^{-1} \in N, m \in Z\}$ in which N is the set of all non-negative integers and Z is the set of all integers, forms a discrete frame Γ_d, spanned by the parameters k and m. $k^{-1} \in N$ corresponds to s < 1. Frame Γ_c **or** Γ_d is said complete in the linear signal space, if any signal function $f(t)$ can be written in the continuous frame Γ_c:

$$f(t) = \int_0^\infty \int_{-\infty}^\infty C(s,u) \, W_{s,u}(t) \, ds \, du \qquad (3.12)$$

Or in the discrete frame Γ_d

$$f(t) = \sum_{k=0}^\infty \sum_{m=-\infty}^\infty C(k,m) W_{k,m}(t) \qquad (3.13)$$

A complete frame is defined to be constructed with all the necessary template functions to express any signal function in the space [149]. In Equations (3.12) and (3.13), the coefficient of the template functions $C(s,u)$ or $C(k,m)$, is a measure function and shows to what extent the signal function $f(t)$ is correlated to the template function W of scale s and at a specific time u.

Wavelet transforms naturally fit into the generalized frame. The finite time interval at different scales makes the template functions in a wavelet transform more effective than the exponential (sine and cosine) function of infinite duration in a Fourier transform to express a non-stationary or transient disturbance signal. The Wavelet transform can be expressed as:

$$C(s,u) = \int_{-\infty}^\infty f(t) \, W_{s,u}(t) dt \qquad (3.14)$$

Where the term **C(s,u)** is called the wavelet coefficient in the wavelet transform and the wavelet function **W**$_{s,u}$**(t)** is defined as in Equation (3.10). Due to the intensive computational load and redundancy in expressing a signal, the continuous wavelet transform is seldom used. There are different kinds of wavelet functions such as Harr wavelet [149], the family of Daubechies wavelets [150] and the biorthorgonal wavelet [151]. Each has a unique shape that can be selected to best match the features of interest in a signal **f(t)** in order to express the signal by template functions at only a few scales, see Figure 3-5 which shows four common Daubechies wavelets. In this project Daubechies' fourth order (db4) is used as it gives the best match to the bearing signals.
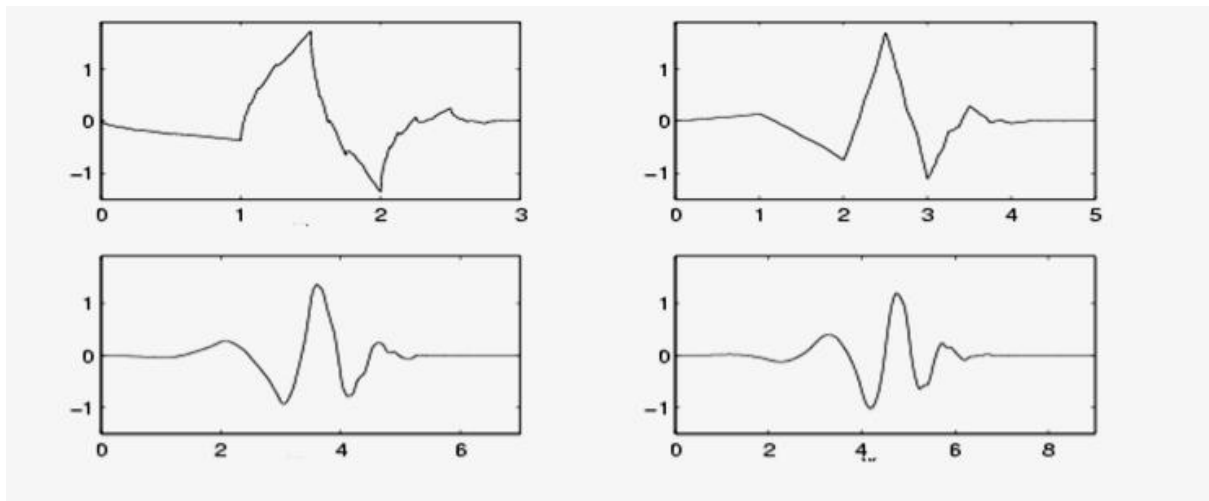
Figure 3-5 Four common Daubechies wavelets, clockwise from top left db = 2, db = 3, db = 8 and db = 4, [149]

The Harr wavelet is actually the 1st order Daubechies wavelet, with template function $\psi^{(1)}_{s,u}$ **(t)**. It represents a wavelet with a simple rectangular waveform. Symmetric biorthogonal wavelets are popular in image processing. Daubechies wavelet functions are non-symmetric in waveform, which makes them suitable to approximate the non-symmetric impulse response of the bearing structure to the defect impact. In addition, different orders of Daubechies wavelets provide a library for a user to try out and compare the different results

in individual applications. Daubechies wavelets have been used in bearing diagnosis by many researchers [152, 153, and 154].

When using wavelet analysis undesired signals and noise may appear and be so significant that it is difficult to identify the pattern of the signal from the fault which makes it difficult to determine what kind of fault it is and what is its magnitude. Statistical parameters such as RMS and peak values have been used to assist in the evaluation of the results [152, 15].

Recent books have described wavelets as becoming established as the most widely used investigative technique for detection and quantification of local bearing faults because of their flexibility and relatively efficient computational processing [155]. They do not, however, specifically refer to incipient fault detection and the problem remains that even if the chosen frequency band does contain the necessary resonances, detection can still fail due to the low amplitude of the fault during its first stages of development.

It is clear that further work is required regarding the theory of incipient fault detection; in fact further work may always be required as the boundary of detection is pushed ever closer to fault initiation.

CHAPTER FOUR

Time Encoded Signal Processing and Recognition (TESPAR)

TESPAR is a method of describing waveforms digitally. It is a collection of signal analysis and classification techniques that was originally used to depict and categorize speech signals. This chapter introduces the approach used for this research; it explains the TESPAR concept by giving a background to TESPAR and its origin. Next TESPAR alphabets and advantages are described then TESPAR S and A matrices are presented and examples of them showing the TESPAR in action are given. TESPAR simulation to signals associated with machine and bearing vibrations is discussed in details. Finally TESPAR coding process is presented in detail with the adaptations made by the author for this research project.

4.1   Introduction

Because of its commercial potential speech recognition has become an important and rapidly developing technology. The attraction of speech recognition is in the simplicity and ease of use of machines and devices given to the operator: e.g. mobile phones, household appliances and computers. The similarities between speech recognition and fault detection are considerable: key features are extracted from the input speech signal and interpreted to give a classification (meaning) to the speech. Bearing defects are impulses in a noisy background and speech is a sequence of sounds also in a noisy background. Of course, speech recognition is more complex than fault detection because the same message can have many meanings depending on the personal and social context.

Not only are the speech signals very susceptible to noise interference they are highly dynamic and stochastic in nature. After World War II Professor King (and colleagues) at the University of Bath (and later at Cranfield University) developed a new approach to speech recognition: TESPAR (Time Encoded Signal Processing and Recognition). This system was patented and is now available through Tespar dsp Limited, which owns the global intellectual property rights. TESPAR is straight forward and codes speech signals without the need to use complex Fourier transformations. TESPAR has the ability to code time varying speech waveforms into optimum configurations.

Clearly a speech recognition technique which is reported to be highly successful at simultaneously reducing the complexity of signal analysis while improving the quality of the information extracted should be of interest to those concerned with extracting relevant information from the vibration signals for faulty roller bearings [156].

Simply, TESPAR is a method of describing waveforms digitally. It is a collection of signal analysis and classification techniques that can be used to describe and classifying a wide range of complex band limited signals, such as those obtained by CM techniques. TESPAR is a technique which classifies time domain signals according to certain shape parameters which can be used as a condition indicator. Various methods can be used to analyse and compare the signals depending on the way in which the symbol stream is presented. In wide use since the late 1990's the technique has been theoretically and practically demonstrated [157, 158].

This author came to the conclusion that TESPAR could be used as part of a new approach to condition monitoring and as a focus for this research.

4.2 TESPAR Coding Background

4.2.1 Infinite Clipping

Infinite clipping is a signal processing technique that was originally developed over sixty years ago [74]. Investigations of human speech showed clearly that while clipping removed amplitude information from the signal and caused big differences in the sound quality heard, there was little reduction in word recognition. A research has been developed [74] to eliminate all amplitude information in the speech signal by "infinite clipping" so that the amplitude of the speech signal waveform was simply absent (reduced to zero) or present at a constant level (a "one") [74]. This binary transformation preserved only the zero-crossing points of the speech signal, see Figure 4-1 [157]. However, the speech remained highly intelligible. These results strongly suggest that speech intelligibility lies in the time durations between the zero crossings of the speech waveform signal, which depend only on the waveform. These observations provided the basis for the development of TESPAR.
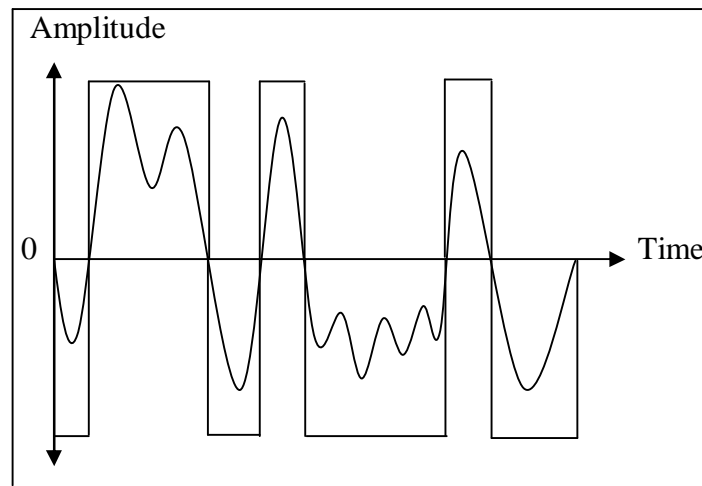
Figure 4-1 Infinite clipping: only zero crossing information preserved [74]

4.2.2 Zero-Based Analysis of Signals

Bond and Cahn in the late 1950's showed that the limitation of the infinite clipping format (e.g. signal distortion) could be overcome by introducing so-called "complex zeros" [159]. A curve that crosses the x-axis only once requires at least a single order equation, two crossings require at least a quadratic expression and crossing n times at least an nth order expression, see Figure 4-1.

The complex zeros are associated with perturbations in the signal shape – e.g. turning points, points of inflexion, etc. - that occur between the well-defined real zeros. The presence of every turning point adds an order to the expression. While it is relatively easy to identify the real zeros that correspond to the signal zero-crossings by a host of numerical techniques (e.g. Newton's method), it is much more difficult to identify the complex zeros [160].

This is a non-trivial problem that requires the solution of a trigonometric polynomial of order 2TW; where W is the bandwidth and T is the time duration of the waveform [74, 157]. Usually $2TW$ exceeds several thousand and the analytic solution of equations of such a high order is generally not possible, and though the computation of numerical solutions is

98

achievable with sufficient computer power, such power was not available at the time and that prevented the further development of the zero models.

However, TESPAR used a method in which the complex zeros were classified directly from the waveform by a series of approximations based on segmentation of the waveform between successive zeros. Of course, the locations of the zero's in the time domain must be equal to the locations of the real zero's in the TESPAR zero domain.

The TESPAR technique uses a representation in which a waveform is described in terms of the locations of its real and complex zeros. This contrasts with conventional techniques based on amplitude sampling at regular intervals, e.g. Fourier analysis. However, the TESPAR and Fourier transform are equivalent in that both require $2TW$ of digital sample data points to describe the waveform [73, 75, 157, 161]. A TESPAR vector quantisation procedure has been developed that codes the signal in terms of approximations to its real and complex zeros [171] and presents the analysis in terms of S and A-Matrices.
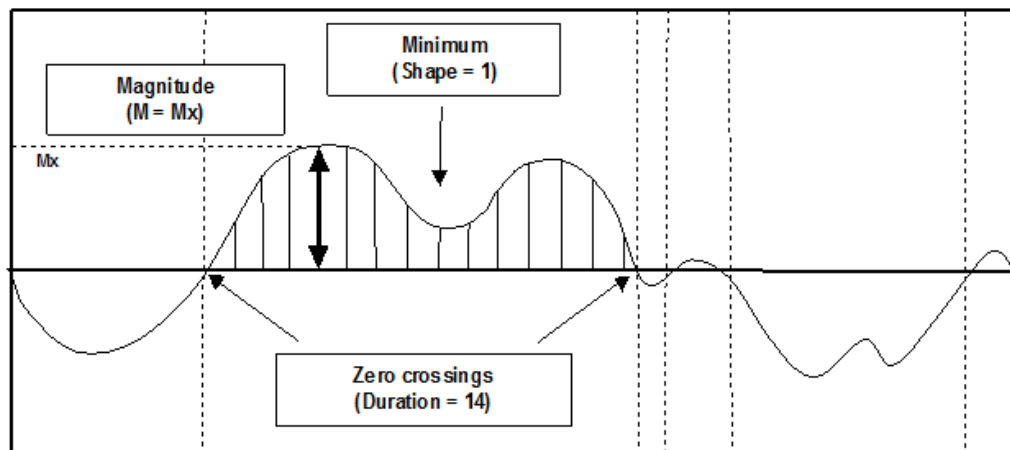

Figure 4-2 TESPAR: single epoch with D = 14, S = 1, and M = Mx [73]

The region between any two adjacent zero-crossings is called an epoch, see Figure 4-2, and every epoch is described in terms of three parameters: the duration (D) which is the number

of intervals between the samples taken over the epoch, the shape (S) which is the number of minima occurring during the epoch (positive values of amplitude, number of maxima for negative amplitudes) and the peak amplitude (M) which is the largest sample amplitude.

4.3   TESPAR Alphabets

In fact most applications can be coded and expressed as a small series of discrete numerical descriptors known as the TESPAR symbol stream. The standard TESPAR symbol alphabet consists of 28 symbols, as shown in Table 4-1, and this has proven sufficient to represent most signal waveforms to an acceptable approximation [162]. The 28 symbol alphabet is presented in Table 4-1 as a look-up table which can be used to convert the epochs comprising the waveform into an equivalent TESPAR symbol stream.

The look-up table was developed empirically and groups various epoch shapes together in the way shown to symbolize an appropriate summary of the waveform, whilst excluding combinations of shape and duration that cannot occur within the bandwidth of the signal. By this way it reduces the "alphabet" required to describe any single epoch waveform to just one of 28 symbols (the integers 1 to 28).

In fact the 28 'word' TESPAR alphabet is referred to as the Standard TESPAR alphabet. Signals with a bandwidth ratio of approximately 10-20:1 produce a natural alphabet of about 200-300 words – each a unique combination of duration and minima. This natural alphabet may then be used to create a smaller quantised alphabet of TES symbols by mapping the natural alphabet onto a two dimensional array of epoch duration on one axis and number of minima on the other. The mapping from natural alphabet to quantised alphabet does not need to be linear – in practice it has proven advantageous to apply a logarithmic quantisation to epoch duration. For example, the logarithmic mapping of epoch duration used in the 28

100

symbol TES alphabet shown in the following table ensures an approximately constant fractional accuracy in terms of symbols per octave.

Table 4-1 TESPAR symbol "alphabet"

Duration (D)	Shape (S)						(SF 20kHz) Fundamental Frequency (Hz)	(SF 96kHz)
	0	1	2	3	4	5		
1	1	1	1	1	1	1	10000	48000
2	2	2	2	2	2	2	5000	24000
3	3	3	3	3	3	3	3333	16000
4	4	4	4	4	4	4	2500	12000
5	5	5	5	5	5	5	2000	9600
6	6	6	6	6	6	6	1667	8000
7	6	6	6	6	6	6	1429	6857
8	7	8	8	8	8	8	1250	6000
9	7	8	8	8	8	8	1111	5333
10	7	8	8	8	8	8	1000	4800
11	9	10	10	10	10	10	909	4364
12	9	10	10	10	10	10	833	4000
13	9	10	10	10	10	10	769	3692
14	11	12	13	13	13	13	714	3429
15	11	12	13	13	13	13	667	3200
16	11	12	13	13	13	13	625	3000
17	11	12	13	13	13	13	588	2824
18	11	12	13	13	13	13	556	2667
19	14	15	16	17	17	17	526	2526
20	14	15	16	17	17	17	500	2400
21	14	15	16	17	17	17	476	2286
22	14	15	16	17	17	17	455	2182
23	14	15	16	17	17	17	435	2087
24	18	19	20	21	22	22	417	2000
25	18	19	20	21	22	22	400	1920
26	18	19	20	21	22	22	385	1846
27	18	19	20	21	22	22	370	1778
28	18	19	20	21	22	22	357	1714
29	18	19	20	21	22	22	345	1655
30	18	19	20	21	22	22	333	1600
31	23	24	25	26	27	28	323	1548
32	23	24	25	26	27	28	313	1500
33	23	24	25	26	27	28	303	1455
34	23	24	25	26	27	28	294	1412
35	23	24	25	26	27	28	286	1371
36	23	24	25	26	27	28	278	1333
37	23	24	25	26	27	28	270	1297

Figure 4-3 shows the TESPAR symbol stream conversion that is extracted according to the Table 4-1, for instance, epoch 1 is represented by D = 4 and S = 0, from the Table we can look up the alphabet symbol matching (D=4, S=0) and that symbol is 4. Epoch 5 is represented by (D = 9, S = 3), so from the table the corresponding alphabet symbol is 8.
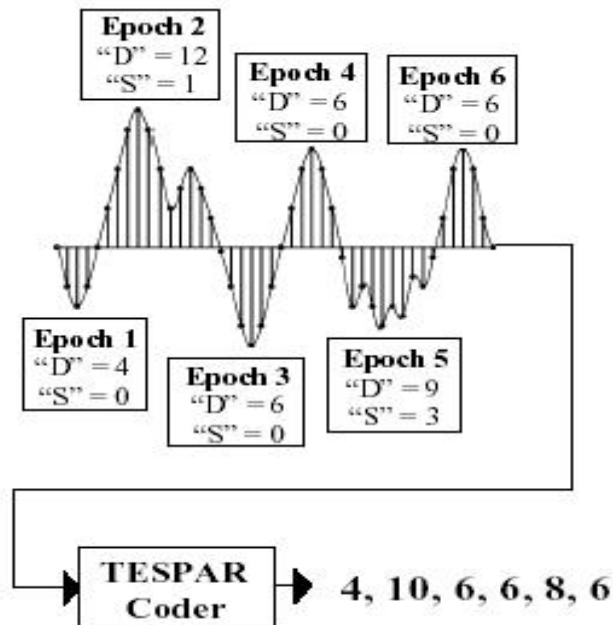

Figure 4-3 TESPAR symbol stream conversion [157]

The six epochs comprising the waveform in Figure 4-3 have a TESPAR symbol stream of 4, 10, 6, 6, 8, 6. Conversion of the TESPAR symbol stream into a compact TESPAR matrix is simple to achieve and allows signals to be represented entirely in terms of their TESPAR symbol distribution.

4.4   TESPAR Advantages

TESPAR is reported to remain reliable even in noisy environments [163], which is a requirement for real-time machine CM in industrial environments. It is because of this robustness that TESPAR has been specifically developed for use in security systems where the speaker has to be identified with a high degree of reliability [164-166]. The 28 codes for

an individual epoch are claimed to make it sensitive and able to differentiate between waveforms with similar spectral content in, e.g., rotating machinery giving TESPAR the possibility of reducing false alarms because of its improved recognition and classification of faults [167].

The technical benefits of the TESPAR approach over other conventional systems include, for example, recognising and classifying conditions and events in band-limited signals of any shape and duration. Also, TESPAR is very much less complex and TESPAR solutions can be implemented at typically 100 times lower cost because it typically requires very much less storage space and it is less computationally intensive; typically 10-100 times fewer MIPS are required [168], in addition TESPAR is very adaptable to changing needs and requirements as can be seen from the wide range of applications listed on the TESPAR website. The software is easy to install and also does not need trained personnel [167-169]. TESPAR can run on 8 bit A/D systems, though 12 bit A/D systems are more common. Of course, this advantage arises at the cost of the results being approximations rather than analytical solutions to the waveform equation

4.5   TESPAR Data Matrices

One of the key characteristics of TESPAR is that it produces descriptors of fixed size and dimensions regardless of the variability of the waveform being coded. Many of the complex methods associated with frequency domain analysis and that involve intensive computations can be discarded; this has the advantage of relative simplicity but at the expense of loss of some information.

TESPAR demands that the signal segment being examined is limited in time by end-points which can be varied - the time between the end points can be easily set and the TESPAR

analysis is restricted to the signal between these two points. The encoding result of the input signals into its symbols stream $\tau(n)$ each of symbol has an amplitude $a(n)$, where n being the nth symbol in the stream. The allocation of quantised TES symbols will depend upon the features of the input signal.

The standard S-Matrix is useful for classification purposes and for providing a visual representation of the signal under test. It is also common practice to divide the signal under test into fixed sized lengths and for an S-Matrix to be produced for each timeslice. By plotting these S-Matrices in a 2 dimensional array, it is possible to visualise changes to the signal over time and can be used to identify additional discriminatory parameters that can further separate the different conditions.

An A-Matrix is a histogram of the number of occurrence of *pairs* of epochs. The epochs that are included in the epoch pair are determined by the lag setting. The default lag setting is a lag of 1 - the epoch pair is made up of adjacent epochs; however a different lag can be specified by using the lag parameter. In general odd numbered lags are preferred as these results in an epoch pair with both +ve and -ve epochs.

Adjusting the A-Matrix lag parameter can be useful when working with signals in which events occur cyclically e.g. rotary machinery. For example, if in the normal running state the same TES symbols occur every cycle, and each cycle is of the same length, then the lag can be set so that the events in the A-Matrix center around the diagonal. The lag is set so that the epoch pair made up of corresponding epochs in consecutive cycles. As the signals deviate from normal, and corresponding symbols in each cycle are no longer the same, the events in the matrix will move away from the diagonal. The degree to which events move away from the diagonal can provide a measure of how far the signal is moving away from normal.

104

The following provides the mathematical underpinning of the S and A-Matrices.

4.5.1   The S-Matrix

The S-matrix is presented as a histogram of TESPAR symbols which records the number of times each TESPAR alphabet symbol appears in the TESPAR symbol stream. That TESPAR matrices are fixed length structures is the main advantage of processing signals using the TESPAR method over traditional frequency domain methods [168, 169].

Given such stream codes T_i, it is possible to build a one-dimensional histogram or S-matrix (in its easiest shape), of the occurrence frequency of quantised TES symbols:

$$S_i = \frac{1}{N} \sum_{i=1}^{M} T_i \qquad 1 \leq i \leq M \qquad\qquad (4.1)$$

where:

$$T_i = \begin{cases} 1 & if \ T_i = N \\ 0 & otherwise \end{cases}$$

4.5.2   The A-Matrix

The A-matrix is defined as a two-dimensional histogram of the occurrence frequency of couples of quantised TES symbols which allows more temporal information to be included and so is generally more useful than the S matrix. For a given symbols sequence in the TESPAR symbol stream T_i, it is possible to build an A-matrix [168, 169] as follows:

$$A_{ij} = \frac{1}{N} \sum_{i=l+1}^{M} T_{ij} \qquad 1 \leq i, \qquad j \leq M \qquad\qquad (4.1)$$

where N represents the total number of epochs in the signal waveforms encoded, l is the lag and $X(n)$ characterizes the given function:

$$T_{ij} = \begin{cases} 1 & if\ \tau(n) = i\ and\ \tau(n - l) = j \\ 0 & otherwise \end{cases}$$

In this way the A-matrix produces a representation of the TESPAR data stream that provides information on periodicities contained within the original signal. The magnitude of the A-matrix at the point (i, j) symbolizes the total occurrence of the TESPAR symbol pair (i, j) where the second number in the pair is delayed relative to the first by the number of epochs identified by the lag. Thus the A-matrix represents the temporal relationship between consecutive pairs of alphabet symbols (integers). It is normally a two dimensional 28×28 vector matrix $(0, 1, ..., 28\ elements)$ that records the number of times each pair of symbols in the "alphabet" appears n integers apart in the integer stream [167] (some researchers have used larger matrices, see example below). l is known as the delay or lag between symbols; l less than 10 characterizes a fast oscillating signal, while l greater than 10 characterized a signal with a slowly oscillating content [169].

Obviously if there is no lag the A matrix will be equivalent to the S matrix, with the entries appearing in a diagonal line across the A-matrix.

4.6 TESPAR in Action

- *Example 1: S-Matrix*

TESPAR was used for speaker identification application [168]. Here the speech was first encoded as an S-matrix: a single dimension $1 \times N$ vector (a histogram), where $Nr.App$ in Figure 4-4 is the number of times the symbol appears in the data stream. It was found that without exception the uni-dimensional description was not as powerful in discriminating

106

between speakers as the two-dimensional A-matrix; see Figure 4-5 below. Eventually, the S-Matrix was discarded as a means of speaker verification and relied only on the A-Matrices.
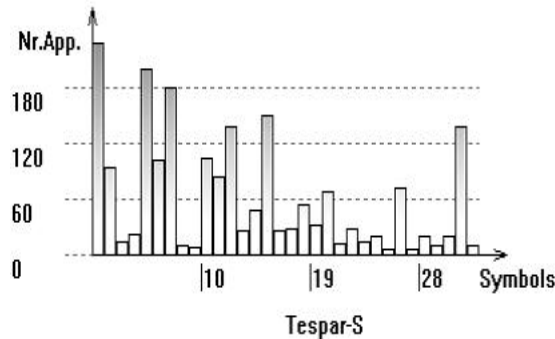

Figure 4-4 TESPAR S-matrix used in speaker recognition [168]

- *Example 2: A-Matrix*

As mentioned above, it was reported that the bi-dimensional A-Matrix when used for speaker identification gives a much greater discriminatory power between two speakers [168]. As can be seen in Figure 4-5 the number of distinctive peaks and their positions in A-matrices can be a feature for classification.
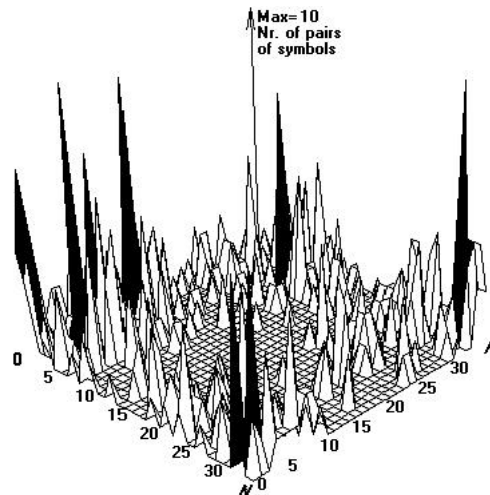

Figure 4-5 TESPAR A-matrix used in speaker recognition [168]

TESPAR and its A-Matrices particularly are capable of classifying signals with amplitude modulation, and because the vibration signals from faulty bearings often exhibit these

107

periodic characteristics, so the TESPAR A-Matrices should be able to be used for CM of bearings, that what will be investigated and discussed thoroughly in Chapter Six.

4.7   Simulation of TESPAR to Detection of Bearing Defects

As explained earlier in the section 3.3.3 envelope analysis bearing faults or defects give rise to signals at clearly defined frequencies but which, certainly in their early stages are swamped by the background noise in the system. Envelope analysis approaches the detection problem indirectly, instead of trying to detect the weak bearing defect signal. Envelope analysis detects the modulating effect of the bearing signal on more easily detectable machine resonances. However, can TESPAR achieve this or similar? In this section a theoretical study of TESPAR S and A-Matrices will be conducted in order to investigate how TESPAR works.

4.7.1   Numerical Study of TESPAR S-Matrices

To clarify and illustrate the TESPAR approach, a simulation of the TESPAR is conducted in this section. A simple sine wave 100 Hz signal of unit amplitude is passed through the TESPAR code. 100 Hz has been chosen simply because it can be taken to represent the BCFs shown in Table 2-3.

There will be 100 cycles per second, so the epochs (periods between zero crossings) will last for one half of one period, and number of samples in each epoch $= Durations$ $\mathbf{=}$ $\mathbf{0.5}\ x\ \mathbf{0.01}\ x\ Sampling\ rate$.

In this particular case the sampling rate was chosen to be $\mathbf{96,000}$ Hz. So the number of $Durations$ $\mathbf{=}$ $\mathbf{480}$ (put another way at a sampling rate of $\mathbf{96,000}$ S/second there are 480 samples taken in one half cycle of a 100 Hz signal). For comparison, a simple sine wave 800 Hz signal of unit amplitude has 60 samples in a one half cycle.
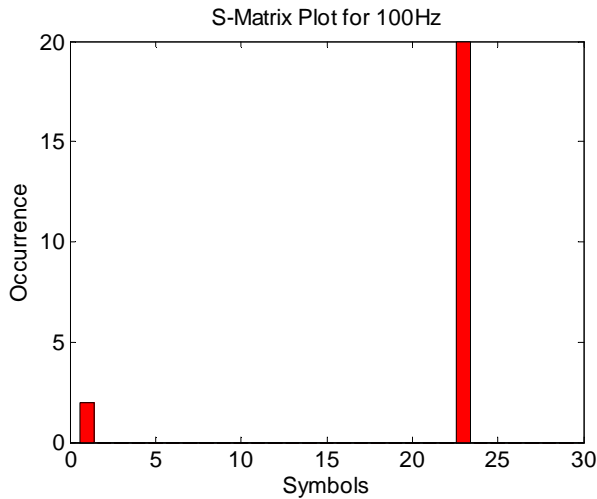
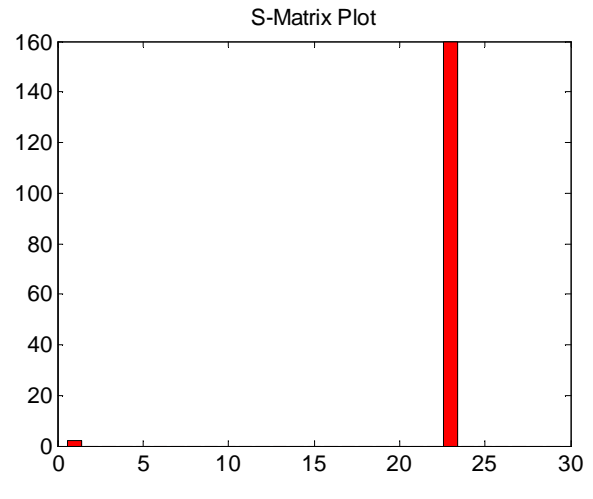**Figure 4-6 TESPAR S-Matrix for simple 100 Hz signal of unit amplitude**

Figure 4-7 TESPAR S-Matrix for simple 800 Hz signal of unit amplitude

For the 100 Hz signal D **= 480** and S **= 0,** since there is only one peak in each epoch. However, the TESPAR coding/alphabet Table 4-1 has a maximum value of D **= 37**. So when D **> 37**, TESPAR sets a default value of 37. The TESPAR coding for D **= 37** $and\ S$ **= 0** is **23**. Similarly for the 800 Hz signal D **= 60** $and\ S$ **= 0**, and again TESPAR sets a default value of 37. The TESPAR coding for $D = 37$ and $S = 0$ is *23*. The TESPAR coding is the same for the 100 Hz and the 800 Hz signals see Figures 4-6 and Figures 4-7, note that in both Figures an entry can be seen in symbol position 1. This is due to the edge effects associated with the signal.

This simulation demonstrates that the TESPAR coding depends on the sampling rate. The bigger the sampling rate the larger the number of Durations and the greater the symbol (x-coordinate) assigned to that epoch.

Because the TESPAR alphabet was designed for speech, the frequency band of interest was typically 300 Hz to 3 kHz, and the Table 4-1 was constructed using a sampling rate of 20 000 S/s. If the 800 Hz signal was sampled at a rate of 20,000 S/s then the Duration D **= 12** (actually 12.5, but fractions are rounded down, see Figure 4-2), S **= 0,** so the coding would

109

be 9. Obviously then, if two signals are to be compared there must be the same sampling rate for each signal.

However, the simulation reveals a second point. A maximum value for the Duration of 37 imposes a lower limit on the frequency analysis – for a sampling rate of 96 000 S/s a frequency below about 1250 Hz will give $D > 37$, which will then be set to 37 by default. However, the bearing characteristic frequencies (BCFs) of the bearing being investigated were all well below 1250 Hz, see Table 2-3. Each of the BCFs in Table 2-3 would be assigned the same coding of 23 for sampling frequencies which meet the Nyquist criterion for machine resonant frequencies. Thus while analysis of raw data using TESPAR could detect the appearance of BCFs it may not be able to discriminate between them.

Extending this approach it is considered that the TESPAR coding acts as a coarse frequency filter. For $D = 1 \; and \; S = 0$, the frequency corresponding to a sampling frequency of 96 000 S/s is 48 kHz. The column on the right hand side of Table 4-1 lists the frequencies associated with Duration but is should be noted that the TESPAR codings for $D > 5$ extend over several Durations. For S = 0 a code of 11 covers a frequency range of 2.7 kHz to 3.4 kHz. Generally, the higher the value of the coding the lower the precision in determining the corresponding frequency content of the signal. It is also correct that the lower the sampling rate the finer the filter, but for condition monitoring – as will be shown later – a sampling rate of 96,000 S/s is reasonable.

In Figures 4-6 and 4-7 both the 100 Hz and 800 Hz signals had same TESPAR codes but the S-Matrix gives a value of 160 to the 800 Hz signal but a value of only 20 to the 100 Hz. The reason for this is that the ordinate does not represent the maximum amplitude of the signal (Mx) but the number of epochs in the signal. A signal of 0.1 seconds long is chosen here, an

110

800 Hz signal has 160 epochs in 0.1 seconds, so the corresponding amplitude or occurrence is 160, a 100 Hz signal has only 20 epochs in 0.1 seconds so the corresponding code 23 has occurrence of 20. By default, S-matrices give equal weighting to all codes; for each occurrence of a code the corresponding entry in the matrix increases by one. Thus the S-Matrix contains no amplitude information, only information on the frequency content of the signal.

Suppose that the machine resonance frequency is 8000 Hz and repeat the above TESPAR analysis for an 8 kHz sine wave. Theoretically exactly 6 samples will be taken in one half cycle of an 8 kHz signal with sampling rate 96,000 S/second. Since $S = 0$ the code will be 6; and the ordinate will be 1600. This is what the result shows in Figure 4-8. Once again a small rectangle at symbol = 1 which was not predicted and again it is due to edge effects. However, the rectangle at code 5 in Figure 4-8 is more interesting because it represents "rounding issues" within the system. An 8000 Hz signal gives a calculated Duration of exactly $6; if\ S\ =\ 0$, and then the symbol value should be 6.
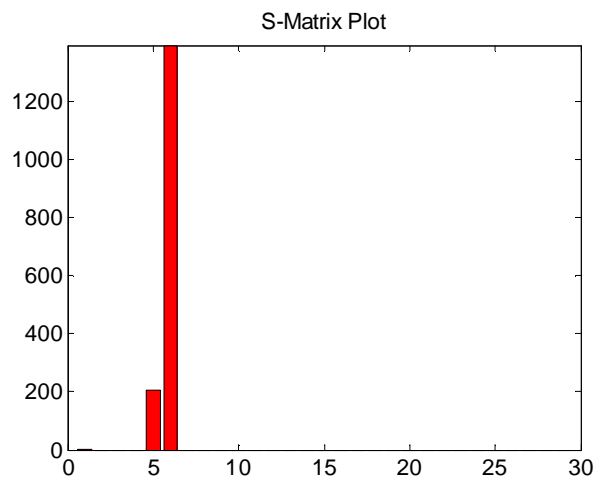

Figure 4-8 TESPAR S-Matrix for simple 8000 Hz signal of unit amplitude

But the system is not perfect and uses approximations so that sometimes the crossing can be rounded down to $D = 5$. So $D = 5$ and $S = 0$ so the coding is 5. Note that the total count is 1600 – as predicted, but 200 of this count are at code 5 and 1400 at code 6. This implies there is either significant "wobble" on the 8000 Hz sine wave or in the TESPAR processing.

By simulating the presence of an incipient bearing fault with BCF 100 Hz, Figure 4-9 below shows the S-Matrix for white noise (assumed to be background noise) and Figure 4-10 shows the S-Matrix for background noise containing a 100 Hz signal with one tenth the amplitude of the noise. No significant difference can be seen. The 100 Hz signal (representing the incipient bearing fault) is effectively hidden.
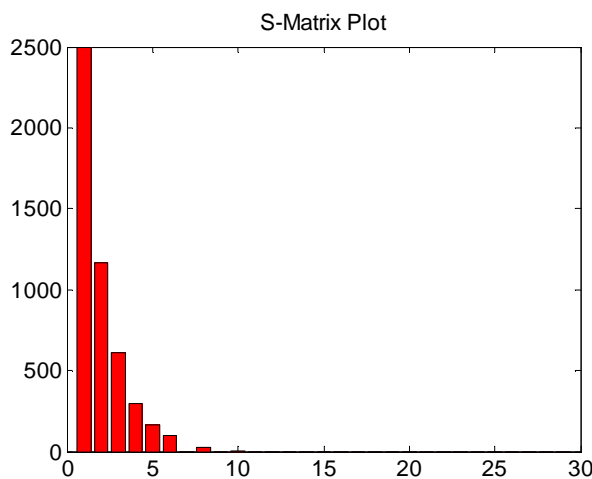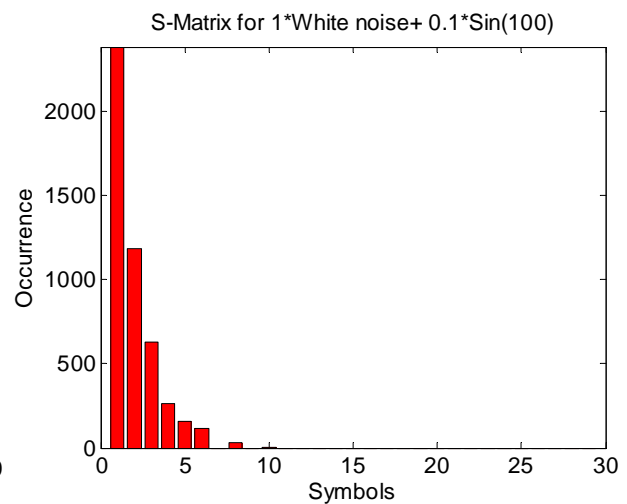

Figure 4-9 S-Matrix for white noise

Figure 4-10 S-Matrix for 1 tenth of white noise with 100 Hz

4.7.2   The S-Matrix and Amplitude Modulation

Considering the amplitude modulation, the two vibrations are then physically interlinked by the process described in section 3.3.3, mathematically, the modulation behaved as the multiplication of the two signals. One immediate simplification arises as a result of the multiplication the relative amplitudes of the signals are not important. Once the BCF exists (an incipient fault) it will modulate the machine resonant frequency. By using the two

112

frequencies above again, the modulated signal is now the product of two sinusoids, one of 100 Hz and the other of 8000 Hz see Figure 4-11.

By simplifying the product trigonometrically using the formula below:

$$sin(2\pi A) \times sin(2\pi B) = \frac{1}{2}\left[cos\big(2\pi(A-B)\big) - cos\big(2\pi(A+B)\big)\right] ----- (4.3)$$

Then:

$$sin(2\pi * 8000t) * sin(2\pi * 100t) = \frac{1}{2}\left[cos(2\pi(8100)t) - cos(2\pi(7900)t)\right]$$

The appearance of a modulated signal is expected which contains sidebands (7900 Hz and 8100 Hz) either side of the 8000 Hz carrier frequency.
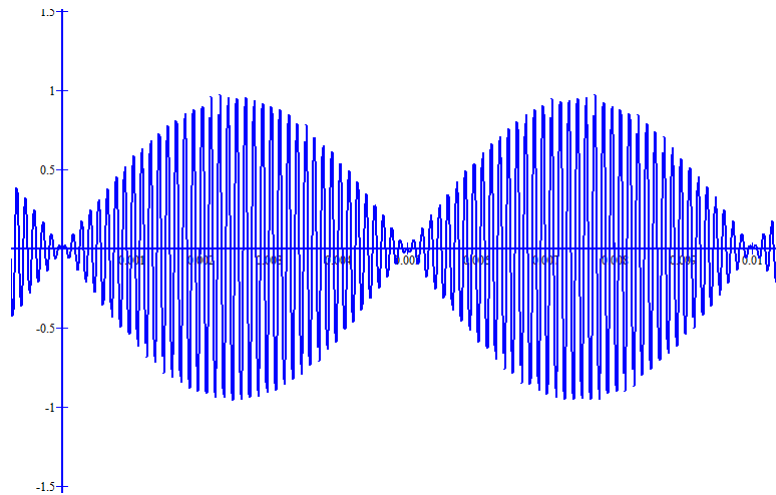

Figure 4-11 The modulated signal of 100 HZ and 8000 HZ

For the 8100 Hz signal the $Duration$ **= 5.92**, which should be rounded down to 5 and with S **= 0**, the code should be 5. For the 7900 Hz signal the $Duration$ **= 6.**08, which should be rounded down to 6 and with S=0, the code should be 6. For the 100 Hz signal the code 23 should have zero ordinate (the BCF does not appear on the S-Matrix). In duration of 0.1 seconds there should be 1600 zero crossings and in the S-Matrix the ordinate values do indeed total 1600, see Figure 4-12.
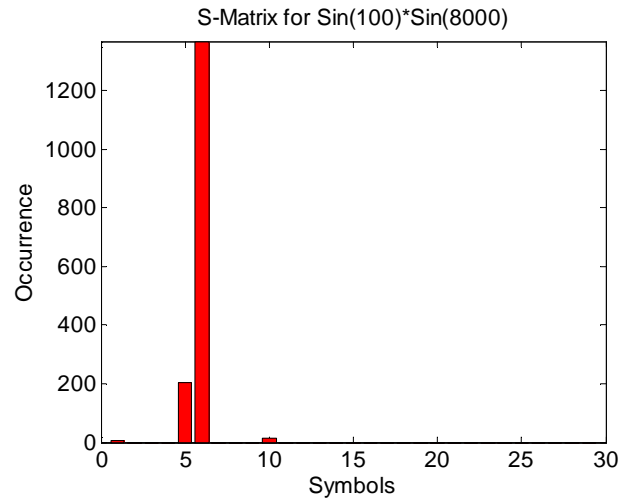
Figure 4-12 S-Matrix for the modulated signal

The ordinate at code 1 is – as previously – due to edge effects in the system. The rectangle at code 10 occurs is possibly due to random fluctuations and/or sampling errors at points such as at $t = 0.001$ seconds, causing the curve to fail to just "touch" the x axis and to not quite reach a zero value, TESPAR will then see a duration of twice the value calculated $(2 \, x \, 6.67 = 13, approximately)$ and then $D = 13$ and $S = 1$ which has a code $10.$

Compare Figures 4-8 which represents the "healthy" condition and Figure 4-12 which represents the condition with the BCF present ("faulty" condition), the two sidebands at 7900 Hz and 8100 Hz which occur by fault existence fall within codes 5 and 6 as for the "healthy" situation that shows TESPAR is a coarse filter.

The major difference between the two plots is the emergence of a small rectangle at symbol 10. This suggests that TESPAR can be used to detect amplitude modulated signals where the BCF modulates a machine resonance.

However, noise is present and it is necessary to simulate how TESPAR performs with the modulated frequencies in the presence of background noise. Figure 4-13 shows the S-Matrices for one tenth of the white noise and modulated signal with sampling rate of 96,000

114

S/s. The result of the background noise in the Figure shows the effect on the amplitude of the machine resonance, which can be problematic for fewer machines.
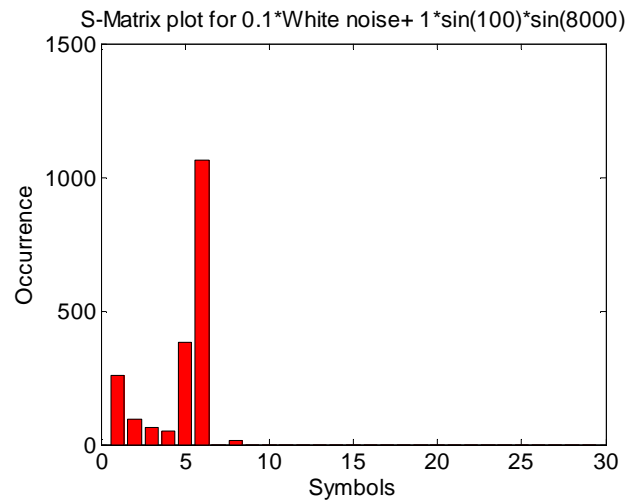

Figure 4-13 S-Matrix for one tenth of white noise and the modulated signal

It is confirmed, exactly as found by previous researchers investigating amplitude modulation as a means of fault detection, it is necessary for the machine resonance to be of much greater magnitude than the background noise.

However, an important point arises here concerned with sampling rate. It is commonly said that the sampling frequency must be at least twice the maximum frequency of interest (Nyquist criterion). But that is not adequate where the sum of two sine terms is concerned (non-coherent sources, very unlikely where a single bearing or drive shaft is involved). Assume having $\sin(At) * \sin(Bt)$ (where A and B are not multiples of one another) it can be then seen that the differences in durations of the epochs change from a minimum that can last as little as a few $\mu\,secs$ to a maximum which is the period of the higher frequency signal. To capture such information a high sampling rate is required.

Thus the higher the sampling rates the better for the kinds of signals likely to be used for bearing fault detection. The TESPAR handbook recommends a sampling rate of three times

115

the Nyquist criterion [172] but this work shows that where the equipment allows for faster rates they should be used.

4.7.3   Numerical Study of TESPAR A-Matrices

In section 4.5.2 the A-Matrix is defined as a two dimensional 28✖28 histogram that records the number of times each pair of symbols in the alphabet appears n symbols apart in the symbol stream, n is called the lag or delay which is the distance between symbols.

Suppose, just for example, the TESPAR symbol stream is short: 4, 8, 9, 7, 4, 13, 9, 2, 8, 5, 7, 4, 10, 9, 1.

Give *n* the value 2 – means looking at the symbol pairs that appear 2 apart. Looking through the stream it is noted that (4, .. ,9) appears 3 times. (8, .. ,7) appears twice (9, .., 4) appears once, (7, .., 13) appears once, (13, .., 2) appears once, (9, .., 8) appears once, (2, .., 5) appears once, (8, .., 7) appears once, (5, .., 4) appears once, (7, .., 10) appears once, (4, .. 9) appears once and (10, .., 1) appears once. Here an A-matrix has a square base, 28✖28. The height of the ordinate at (4,9) is 3, at (8,7) is 2, and the ordinates at (2,5), (4,9), (5,4), (7,10), (7,13), (8,7), (9,4), (9,8), (10,1) and (13,2) are all 1. All other ordinates on the Matrix will be zero since there are no more pairings.

Looking at the line of figures in the TESPAR stream above the pairs can be read from right to left. So the following is also true; the height of the ordinate at (9,4) is 3, at (7,9) is 2, and the ordinates at (5,2), (4,9), (4,5), (10,7), (13,7), (7,8), (4,9), (8,9), (1,10) and (2,13) are all 1. This means the A-Matrix has a line of diagonal symmetry which corresponds to there being no lag with the entries appearing in a diagonal line across the A-matrix and the line of symmetry being equivalent to the S matrix.

The A-Matrices of the simple 100 Hz sine wave is a single peak at (23, 23), see Figure 4-14. For the 8000 Hz signal (supposedly healthy condition) there is a major peak at (6, 6) with very small side peaks at (5, 6) and (6, 5), see Figure 4-15. For the modulated signal (supposedly fault condition) there is a major peak at (6, 6) with very small side peaks at (6, 10) and (10, 6), see Figure 4-16.
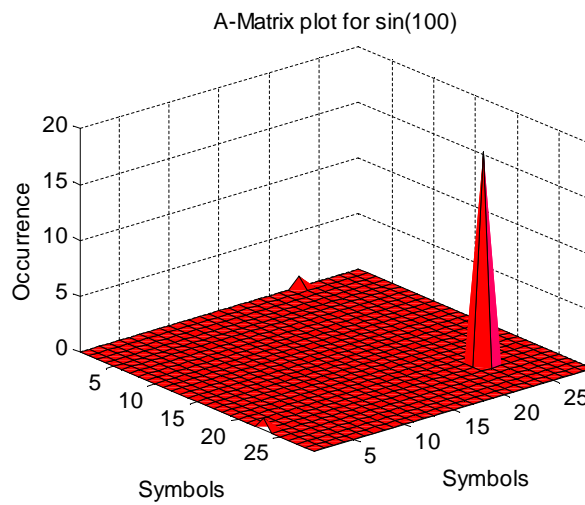

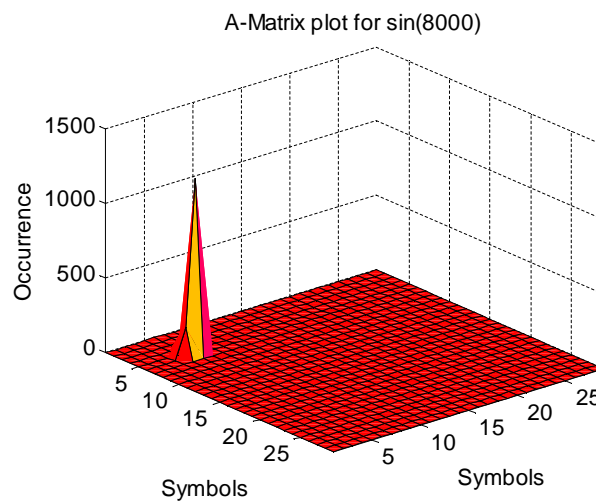**Figure 4-14 A-Matrix for 100 Hz sine wave with (96,000 S/s)**

Figure 4-15 A-Matrix for 8000 Hz sine wave with (96,000 S/s)

117
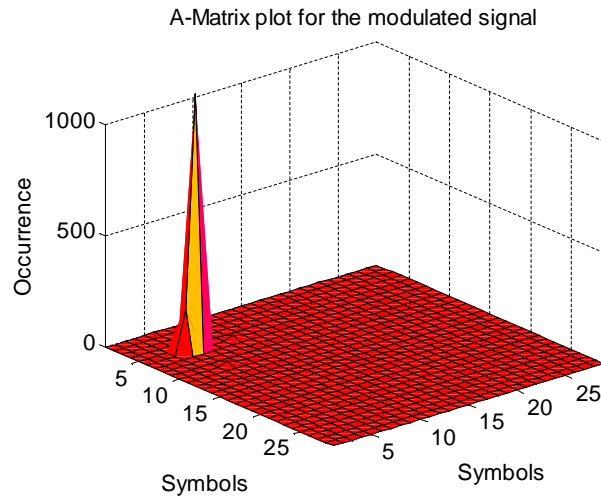
Figure 4-16 A-matrix for modulated signal with (96 000 S/s)

The A-Matrix for white noise, and is what would be expected a maximum at (1, 1) falling away smoothly and steeply shown in Figure 4-17. The A-Matrix for the simulated modulated signal with background noise is shown in Figure 4-18 and to the eye there appears very little difference to the Matrix with no fault present, the peak at (23, 23) disappears due to the TESPAR processes.

**Figure 4-17 A-Matrix for one tenth of amplitude
of random white noise**

The addition of white noise has very little effect on the A-Matrix when the resonant signal is 10dB above background, see Figure 4-19, but the BCF is no more discernible.
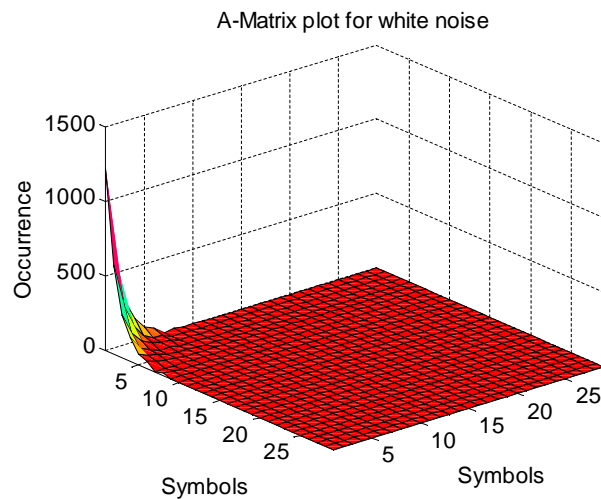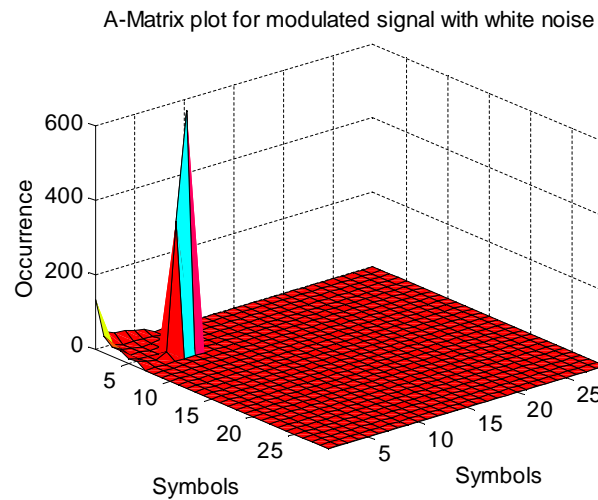

Figure 4-18 A-Matrix for modulated signal with white noise

But what physical meaning does this A-Matrix pairing mechanism have? If *n*=0, simply the original TESPAR stream emerges which is the S-Matrix appearing along the diagonal of the A-Matrix, $n = 0$ provides information on the total numbers of epochs of a given duration and shape but little information on the signal's temporal pattern. For $n = 1$ (adjacent symbols), an elementary insight into the temporal pattern of the signal is obtained, in the sense of a statistical measure of the order in which epochs follow each other.

The example of one tenth of white noise with modulated signal being examined here, the 8000 Hz signal should repeat itself every 12 samples and so the delay, n, was set to 12 and Figure 4-19 was obtained. There are significantly sharper peaks than for $n = 1$, at about (1, 5) and (5, 1). There is thus the potential for better fault discrimination if the delay is set to match the carrier frequency or, possibly, one of the modulated frequencies.

Therefore a series of tests were carried out on the test rig to investigate the effect of changing *n* on the A-Matrices for the different faults. The outcome was that $n = 1$ was the delay used

119

because in the trials as *n* was increased there were only small changes in the A-Matrix and in every case in the direction of smaller ordinate values.
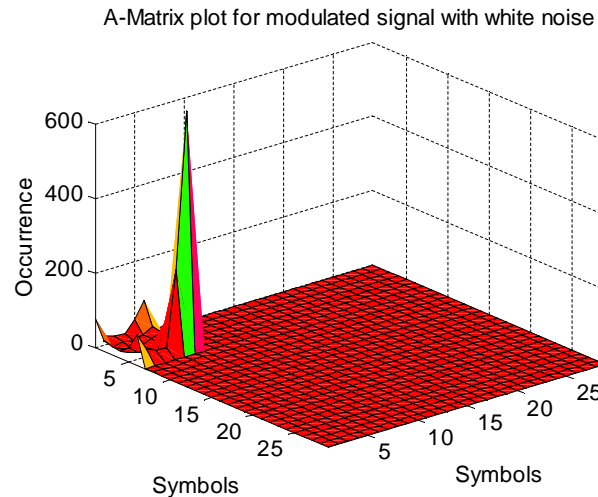
Figure 4-19 A-Matrix for modulated signal with white noise and delay = 12

4.8   TESPAR Coding and Implementation Process

Chapter Five describes how the raw signal is obtained from the CM sensor on the test rig and sampled in a way which ensures successful categorization and processing of the signal waveform for each of the ten bearing conditions investigated in this thesis. The TESPAR code was written to process the data so obtained.

Typically all TESPAR analyses comprise three main processing stages: pre-processing, TESPAR coding and classification. Figure 4-20 shows these steps. The pre-processing stage includes filtering the raw data so only the frequencies of interest are analysed during the TESPAR coding process. The filter applied was a *2nd* order Butterworth bandpass filter with high and low pass frequencies of 250 Hz to 17000 Hz respectively.
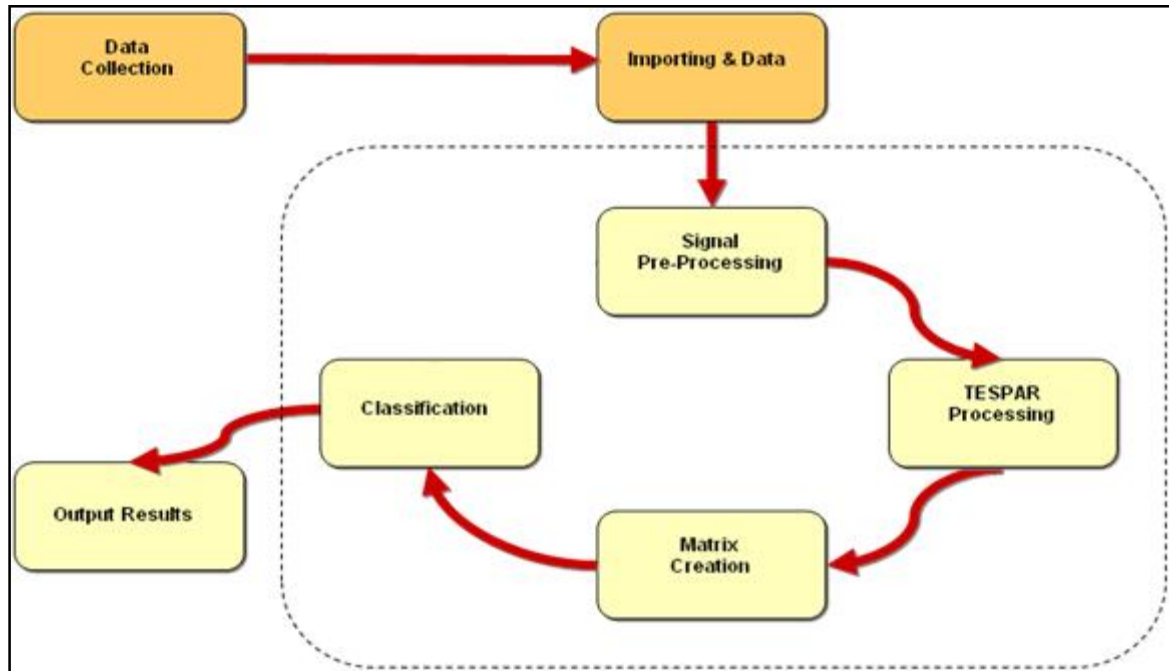
Figure 4-20 TESPAR coding and classification procedure action

The reason for using 250 Hz high pass filter cut off even though the lowest frequency that can be coded by TESPAR is 1297 Hz; is because of the low filter order being used. The filter order is only a 2nd order filter and therefore the filter roll-off will not be particularly steep. If a higher order filter of 8 or 16 had been used, the high pass cut off could have been moved much closer to the TESPAR coding limit of 1297 Hz (based upon the 96 kHz sampling rate). Nevertheless, it is believed that using a 2nd order filter significantly reduced the computational processing required by higher filter orders and yet was still sufficient to provide classification of fault locations and severities.

The TESPAR coding process comprises the following key steps:

- Create TESPAR stream

This is performed using the code:

```
TStream = tespar(FilteredSignal,FluctError);
```

121

Where the first parameter is the filtered signal, the second parameter is the fluctuation error (in this case set to 0). Epoch fluctuations that have a peak-to-peak value of less than 'FluctError' are not counted as extrema and do not contribute to the shape attribute of an epoch.

The output 'TStream' is a 3-by-N matrix where N is the number of epochs and:

Row	Description
1	Peak Epoch Amplitude
2	Epoch Duration (No. of samples)
3	Epoch Minima/Maxima

- Create TESPAR codes

This is performed using the code:

```
TCode = tescode(TStream,AmpThresh,stdalph);
```

Where the first parameter is the TStream (generated above), the second parameter is the amplitude threshold (in this case set to 0) which assigns the TES symbol for silence to epochs whose peak amplitude is less than the specified silence threshold. The third parameter confirms the standard TESPAR alphabet is used. This is simply the look up table that converts the duration and shape epoch components into a TESPAR code value between 1 and 28 (29 including the silence threshold).

The output 'TCode' is a 2-by-N matrix where N is the number of epochs and:

Row	Description
1	Absolute Peak Epoch Amplitude
2	TESPAR code value

- Create TESPAR S-Matrix and A-Matrix for each fault location example (as there are three examples of each of the four fault locations, twelve S-Matrices and twelve A-Matrices will be generated).

This is performed using the codes:

S-Matrix:

```
Mats{FaultLocnLoop,SigLoop} = smatrix(TCode,AmpWeight,28);
```

A-Matrix:

```
Mats{FaultLocnLoop,SigLoop} = amatrix(TCode,AmpWeight,Lag,28);
```

The first parameter is the 2xN vector TCode generated above, the second parameter specifies the epoch amplitude weighting to be applied. By default, S-matrices and A-matrices give equal weighting to all TES codes; for each occurrence of a TES code or pair of codes, the corresponding entry in the matrix is incremented by one. Thus, by this measure, the *magnitude* or *energy* of the segments of the signal forms no part in these descriptors. *Amplitude weighting* takes advantage of the peak epoch amplitude data stored as part of the TES code of each epoch. The corresponding entry in the matrix is incremented by a value proportional to the peak epoch amplitude. In this way, epochs with higher amplitudes are given more weighting than epochs of lower amplitude.

Depending on the application, amplitude weighting can result in improved discrimination between signals from different condition sets. The default value for weight is 0 indicating that no amplitude weighting will be applied, in which case, each occurrence of a pair of TES symbols causes the entry in the corresponding bucket of the A-matrix to be incremented by 1. If the weight argument is specified, the amount by which the corresponding bucket in the A-

matrix is incremented is calculated by adding together the amplitudes of both epochs, dividing this figure by the weight and adding 1.

The third parameter specifies the lag to be used during the A-matrix generation process by comparing TES symbol codes from epochs that are separated by 'Lag', in this research n = 1.

The fourth parameter simply fixes the size of the A-matrix output with a default setting of 28✖28 elements. For a given quantised TES symbol sequence $\tau(n)$, it is possible to construct a two-dimensional histogram using Equations (4.1) and (4.2).

4.9   Test Procedure Adopted

Assuming that the mean absolute magnitude of the lower frequencies in the spectrum provides an indicator of fault severity because it is anticipated that as fault severity increases the magnitude of the lower frequency epochs increases. However, the simulations above suggest these might be hidden by the traditional forms of TESPAR data presentation – the S and A Matrices - and so a modified approach was adopted. Firstly, all epochs that had durations greater than the average duration associated with the whole signal were identified. Then the code was adapted to generate the mean absolute magnitude value of all these longer than average epochs. The mean absolute magnitude value thus produced is taken to be a direct indictor of fault severity.

Then create baseline S and A-Matrices, called an archetype, for each of the four bearing conditions, one healthy and three different faults.

As dimensions of the S-Matrix are fixed at 1✖28 and the A-Matrix are fixed at 28✖28, it is very simple to generate an archetype matrix for each of the conditions by simple averaging

the individual matrices generated from the example associated with each condition the S and

A Matrix archetypes are given below:

$$\textbf{Archetype} = \frac{\sum_{i=0}^{n} \textbf{S Matrix(i)}}{\textbf{n}} \qquad \text{(4.4)}$$

and

$$\textbf{Archetype} = \frac{\sum_{i=0}^{n} \textbf{A Matrix(i)}}{\textbf{n}} \qquad \text{(4.5)}$$

When using TESPAR as a signal processing solution, it is not necessary to always use

matrices as a means of classification. It is possible to use other and possibly more appropriate

TESPAR features to identify the presence and classify the severity of a fault. There can be

advantages to using the output directly from the TESPAR or Tescode functions. Within the

code, the following lines are processed for each example:

```
Idx = find(TStream(2,:) >= mean(TStream(2,:)));
```

The code above identifies the indices of all of the epochs whose duration is greater than the

mean duration within the signal. Essentially the indices being identified are those associated

with the lower frequency components from within the signal.

```
MeanMag = mean(abs(TStream(1,Idx)));
```

This part of the code generates the mean absolute magnitude values from all of the indices in

the epochs identified above.

Essentially this means identifying the energy associated with the lower frequency

components within the signal.

```
AllMeanMags = [AllMeanMags MeanMag];
```

The part of the code above stores the mean absolute magnitude generated. A comparison then will be made between the stored values which can be used to identify the fault severity as will be shown later in Chapter Six.

It is proposed to identify the fault location during the classification stage of the data processing. This will be done by comparing each of the test S-Matrices A-Matrices with the four S and A archetypes. In this example, correlation was used as the classification procedure of choice. A 'winner takes all' strategy is applied, meaning that the archetype giving the highest correlation score when matched with the test Matrix is defined as the 'winning' fault. The correlation is performed by using the 'score' function:

$$C = SCORE(A,B);$$

This function returns the angle correlation score between matrices A and B: this is simply the square of the cosine of the angle between the two `vectors'.

$$C(a,b)_{ang} = \frac{sum((a.*b))^{\wedge}2}{sum(a.*a) * sum(b.*b)} \quad - - - - - \textbf{(4.5)}$$

The correlation scores produced by comparing each signal under test with each archetype are discussed later in TESPAR analysis in Chapter Six. Classification of the faults was done using both S-Matrix correlation and the A-Matrix correlation described above.

To determine which methodology (the conventional methods or TESPAR) is better; metric needs to be identified which can be used to compare the different techniques. During this research, three key metrics have been identified that can be used for this purpose. These are:

- Results accuracy.
- Processing overheads (memory used, current draw and number of processing cycles required).

126

- Ease of implementation.

Although other techniques can match the TESPAR performance, the TESPAR implementation requires significantly less processing overheads and is easier to implement as less underlying mathematical knowledge is required.

CHAPTER FIVE

Test Rig Facility and Experimental Procedure

This chapter introduces the test rig construction and the parameters measurement associated with incipient defect detection in a roller bearing. The parameters to be measured are discussed and details of the relevant transducer and data acquisition system given, as are details of measurement practice. The manner in which defects are introduced into the bearings and data acquisition is described in detail.

5.1   Experimental Facilities

The aim of this study was to detect specific faults introduced into given roller bearings and to diagnose differences between the different faults. This will be done by monitoring the vibration signal from an accelerometer placed on the bearing housing. The rig was simple and industrially relevant in which faults could be introduced in a repeatable and controlled manner and the vibration be measured to an acceptable level of accuracy.

An important element of CM research is the seeding of faults in a controlled manner into a practical system that will provide real data similar to that which might be recorded in an industrial situation. The choice of a roller bearing is because this component is in widespread use in a wide range of industries. The bearing chosen for this experiment was a NSK type N406 cylindrical roller bearing which is typical of many used in industry and the bearing faults could be introduced as required. The rig also had to include an accelerometer for measuring the vibration generated. The instrumentation and equipment used was selected on the basis that it was available within the Centre for Efficiency and Performance Engineering at the University of Huddersfield and had been found to be accurate and reliable. The advantage of using an existing rig that has been developed and used previously is represented in disassembling and re-assembling and that is due to its construction which has been done in a way to ease its connection to other experimental facilities such as sensors as well as its isolation from other sources of vibration. New instrumentation was also purchased from Sinocera on the basis that this was a company that produced quality products at competitive prices.

5.1.1   Test Rig Construction and Components

The test rig is shown in Figure 5-1. The drive power is from a 4kW, 3-phase, and 4 pole electric inductions motor. The speed and torque of the motor are controlled by a Siemens

129

Micro Master Controller so the drive shaft can be run at different speeds (to a maximum of 1460 rpm) with different applied torque loads to a maximum of 4.0 kW. Two pairs of matched flexible couplings couple the motor to the brake via three short cylindrical steel shafts. There are two bearing housings to support the shafts, the housing nearer the motor contains a MAC cylindrical roller bearing type N 406 and the other contains an SKF double row self-aligning spherical roller bearing type 22208 EK.

A radial load is applied to the central shaft using a hand pump with a pressure gate to pressurise a hydraulic ram frame, which is connected to a load cell. The Sinocera YD-5 piezoelectric accelerometer with frequency range up to 20 kHz was positioned on the housing of the N 406 bearing; it was mounted vertically to measure the vibration.
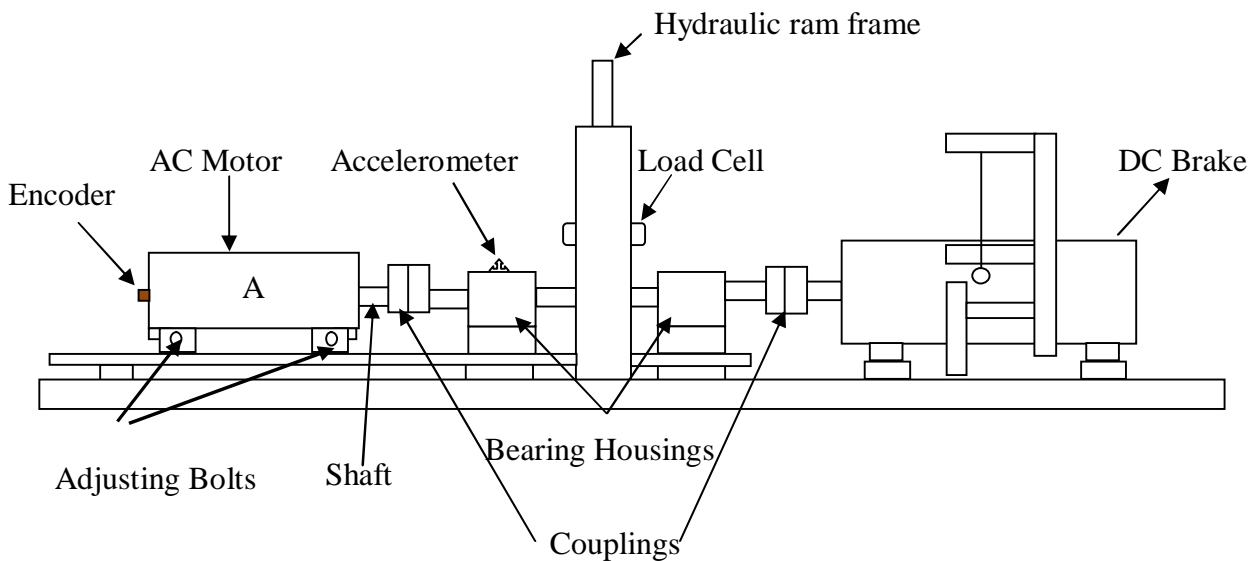

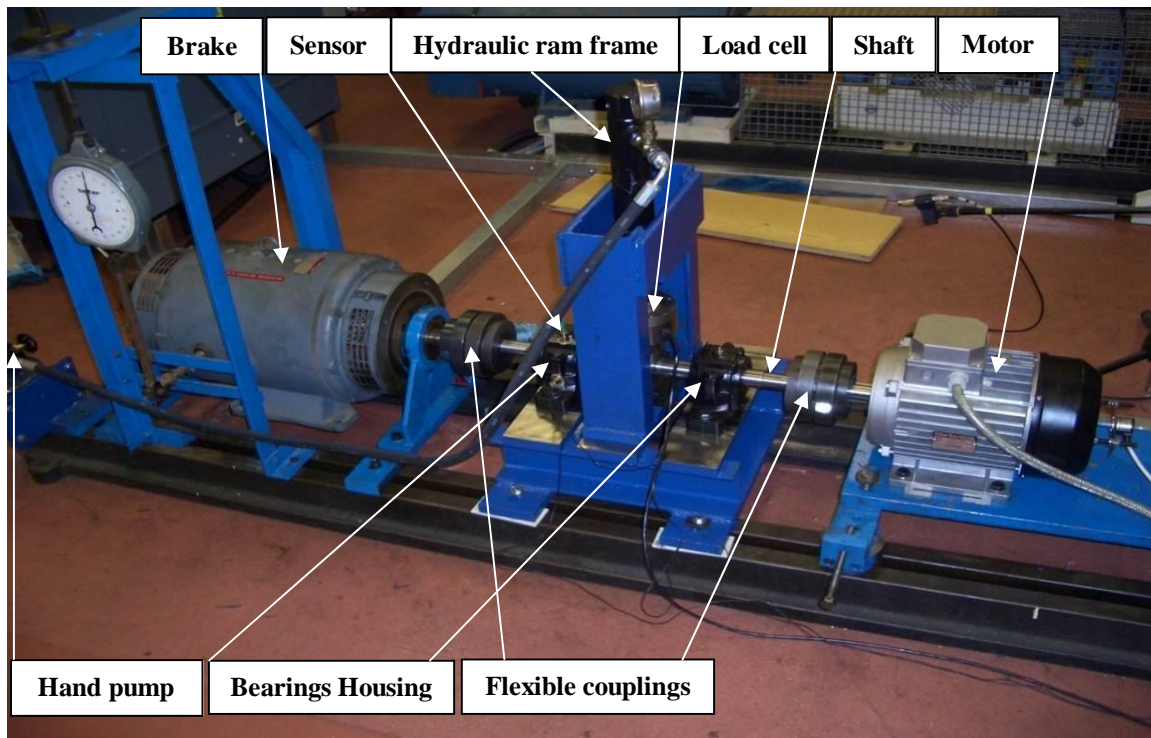**Figure 5-1 Schematic of the test rig construction**

Figure 5-2 Test rig construction

5.1.2 Data Acquisition System

A data acquisition system was designed to monitor and record data parameters such as temperature, vibration, sound, etc., but it also contains some basic data analysis tools including spectrum analysis for online data examination.

A Sinocera type YE6232B, 16 channel, 16 bit data acquisition system was used, see Figure 5-2. Essentially a 16-channel high speed data acquisition system was employed to record all the measurements at a sampling rate of 96 kHz. The hardware consisted of three parts: the piezoelectric accelerometer (sensor) which was connected via a charge amplifier to a data acquisition card which was then connected to a computer. The amplifier is necessary to amplify the vibration signal, which is often very weak, and it also electronically isolates the accelerometer from the processing and display equipment. The software is the data acquisition logic and the analysis software (with other utilities that can be used to access and

131

analyse data from the data acquisition memory of the computer). The data acquisition system

specification is shown in Table 5-1.

Figure 5-3 Sinocera YE6232B data acquisition system

Table 5-1 Data acquisition system specification

Data Acquisition System	Specification
Manufacturer	Sinocera YE6232B
Number of channels	16 channels, selectable voltage/IEPE input
A/D conversion resolution	16 bit
Sampling rate (maximum)	100 kHz per channel parallel sampling
Input range	±5V
Gain	Selectable either 1, 10 or 100
Filter	Anti-aliasing filter
Interface	USB 2.0

A data acquisition control programme developed for Lab-Windows was used and consists of

a main data inspection panel and parameters set-up panel. This software was based on a

Windows operating system and had the capability to carry out on-line data sampling.

Modifications such as the number of channel, sampling frequency, data length and filenames

can be chosen, recorded and stored on separate set-up page of the software package.

5.1.3 Vibration Transducer

A vibration transducer used to measure the structural vibration converts vibration energy into a measurable voltage. While velocity pickups and eddy current or proximity probes are still used for vibration measurement, accelerometers are by far the most popular vibration transducers used with rotating machinery [171].

Piezoelectric accelerometers can be described as a single degree of freedom spring-mass system which is so arranged as to generate a force equivalent to the amplitude and frequency of vibration. The force is applied to a piezoelectric element, which produces a charge on its terminals that is proportional to the vibratory motion. A charge amplifier is used to produce an output Voltage. The specification of the accelerometers is given below in Table 5-2 also see Figure 5-3.

Figure 5-4 Sinocera accelerometer model YD-5 [121]

Table 5-2 Sinocera accelerometer specification [121]

Accelerometer	Manufacturer Sinocera
Type	Piezoelectric
Model	YD-5 (SN. 4251)
Frequency range	0.5 Hz to 20kHz
Sensitivity (Calibration)	8.08 mv/ms^{-2}
Temperature	-20ºC to 120ºC

The accelerometer was attached to the bearing housing casing vertically by threaded bronze studs**.** The vibration transducer was connected to a charge amplifier and from the charge amplifier to the data acquisition system and then to the computer via a USB port.

5.1.4   Load Cell

The load cell, see Figure 5-4, is a sensor used to measure the force applied by the hydraulic ram onto the bearings. Specification of the Sinocera model CL-YB-11 5kN load cell is given in Table 5-3

Figure 5-5 Sinocera load cell [121]

Table 5-3 Sinocera load cell specification [121]

Load Cell	Manufacturer Sinocera
Name	Weigh transducers (Strain based)
Type	CL-YB-11
Sensitivity	1.5-mV/V
Linear range	0-5 kN
Overload	150%
Operating temperature	-20ºC to 60ºC
Material	Special alloy steel
Output type	Five core shielded cable

5.1.5   Speed and Torque Controller

The test rig is operated at different speeds and with different torque loads, so a speed and torque controller was required. In Figure 5-5, the Siemens Micro Master Controller was connected into the rig and proved easy to use and was able to deliver a known torque and speed accurately.
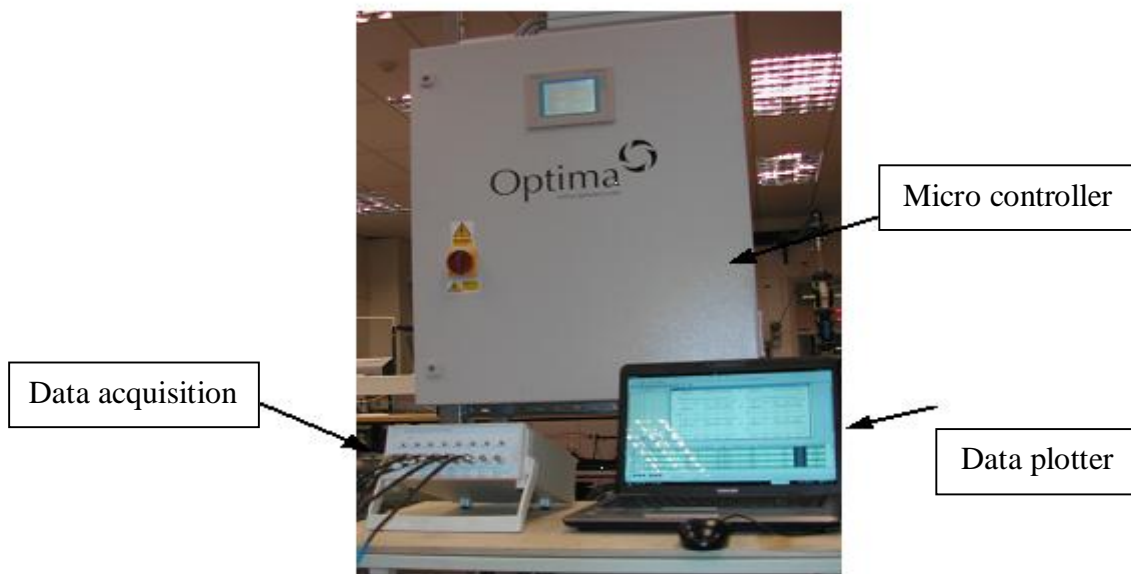

**Figure 5-6 Photograph of Siemens Micro Master Controller
and data acquisition system**

5.2   Bearing Sample Preparation

In this research, ten NSK type N406 cylindrical roller bearings were used and their specifications are shown in Table 5-4:

Table 5-4 Specification of NSK type N406 cylindrical roller bearing

Parameter	Measurement
Pitch Diameter	59 mm
Bore Diameter	30 mm
Roller Diameter	14 mm
Roller Number	9
Contact Angle	0^0

One of the bearings had no fault introduced and was considered a reference. Three kinds of faults were simulated; roller, inner race and outer race. The faults were scratches of varying degrees of severity introduced by Electrical Discharge Machining (EDM). The scratches were all rectangular slots of depth 0.1mm and width 0.18mm, but of varying lengths for all simulated faults used in this research. It is obvious that single scratches of uniform width and depth are not truly representational of real defects in bearings, but this approach has three important advantages: the faults are consistently reproducible across all the bearings; such a defect is expected to produce a clear periodic signal which will optimise the possible success of the TESPAR application; and it is possible to control the degree of the fault, i.e to introduce a small fault. As known a real defect depends on the nature of the cause of that defect, some of them start small in terms of depth, width and length then increase gradually, others are apparent from the start, sees section 2.4 in Chapter two.

It is accepted that a possible alternative would have been to overload and run to failure and this could be a topic for further work. This method was not the one chosen here because, firstly, the radial load on the bearing needed would have been higher than the rig capacity to bear as well as being time consuming in a time restricted research project. Secondly, the faults could be produced and simulated in the same way by electrical discharge machining.

Faults R_1, R_2 and R_3 were scratches introduced onto the rollers of the bearing see Figure 5-6 for example. The most severe fault (R_1) extended right across the roller (100% = 13.2mm length), the least severe fault (R_3) extended only 30% of the way across the roller (3.96mm length) starting from the edge of the ground area and the intermediate fault (R_2) extended 60% of the way across the roller (7.92mm length), again starting from the edge of the ground area.

Figure 5-7 Photograph of NSK type N406 cylindrical roller bearing

with 30% roller fault (R3)

Faults I_1, I_2 and I_3 were scratches introduced onto the inner race see Figure 5-7 as an example. The most severe fault (I_1) extended right across the race (100% = 12.00mm length), the least severe fault (I_3) extended only 30% of the way across the race (3.60mm length) starting from the edge of the ground area and the intermediate fault (I_2) extended 60% of the way across the race (7.20mm length), again starting from the edge of the ground area.

Figure 5-8 Photograph of NSK type N406 cylindrical roller bearing

with 60% inner race fault (I2)

137

Faults O_1, O_2 and O_3 were scratches introduced onto the outer race see Figure 5-8 as an example. The most severe fault (O_1) extended right across the race (100% = 18.50mm length), the least severe fault (O_3) extended only 30% of the way across the race (5.55mm length) starting from the edge of the ground area and the intermediate fault (O_2) extended 60% of the way across the race (11.10mm length), again starting from the edge of the ground area.
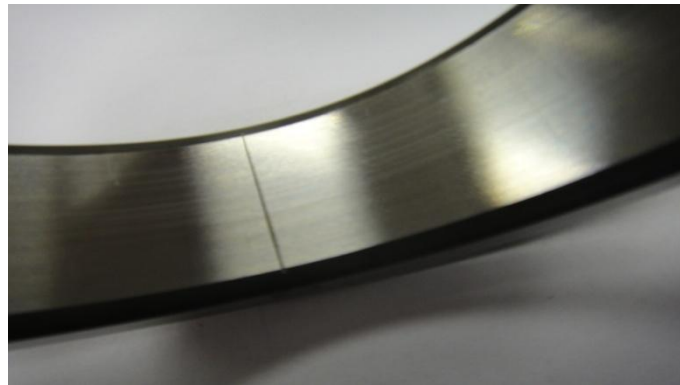

Figure 5-9 Photograph of NSK type N406 cylindrical roller bearing

with 100% outer race fault (O1)

Table 5-5 lists the specifications of the defects (scratches).

Table 5-5 Bearings defect specification
(defect width = 0.18mm and depth = 0.10mm)

Bearing	Fault	Length mm	Remarks
1	Roller = R_1	100% = 13.20	Starts from
2	Roller = R_2	60% = 7.92	Edge of
3	Roller = R_3	30% = 3.96	Ground area
Bearing	Fault	Length mm	Remark
4	Inner race = I_1	100% = 12.00	Starts from
5	Inner race = I_2	60% = 7.20	Edge of
6	Inner race = I_3	30% = 3.60	Ground area
Bearing	Fault	Length mm	Remark
7	Outrace = O_1	100% = 18.50	Starts from
8	Outrace = O_2	60% = 11.10	Edge of
9	Outrace = O_3	30% = 5.55	Ground area

5.3   Experimental and Data Collection Procedure

The bearing rig was allowed to run 5 minutes to warm up before starting data collection. Of course errors such as manufacturing error, speed relative errors and roller estimation error might be presented and the accuracy cannot be 100%, however, the ten bearing conditions were tested with the assumption that the bearings have no clearance, no lubrication and no slippage error. The bearing with no faults introduced was tested and used as a reference. Vibration signals were collected from the accelerometer mounted as already described on the bearing housing. Experiments were conducted to study bearing fault detection and diagnosis. Three sets of data were taken for each bearing condition, the three data sets of each bearing condition were taken in the same day, each set of data contains sex tests to ensure that the signals obtained were self-consistent, one of the sets was used to create the S and A archetype matrices and the other sets of data were used to create the matrices. All experiments were conducted in equipped Laboratory that maintains the same weather condition; therefore, all tests were supposedly carried out under the same conditions and also under full shaft rotational speed of 1460 rpm (24.3 Hz) to enhance the vibration produced by running on high speed. As known in real condition there are different kinds of vibration conditions; axial, torsion and radial loads, therefore, 50% of 4kW torsion load and a radial load of 450N was applied vertically on the shaft that is connected to the bearings in every test where the vibration is also measured vertically and picked from the housing that contains the bearing N(406) tested, the other supporting bearing was not subject to experiments because the study concerns heavy duty bearing and also avoid adding other modifications to the pre-excited rig (e.g. adding other kind of bearing or extra sensors) which would cause more difficulties and more time consuming to a fixed time research, and also to prevent potential disturbing for existing ones. These loads are applied to the rig, firstly to simulate real

139

conditions in industry, and secondly to be sure that they are permissible by the rig tolerance. The sampling rate was 96 kHz and data length was 960000 points with duration of 10 seconds. At this shaft speed the bearing would have rotated over 240 times in 10 seconds and it was considered that this was an adequate number (also 10 seconds is a convenient duration). The sampling rate of 96 kHz meant the upper frequency limit of the system was over 40 kHz; this was about 300 times the highest of the calculated and measured characteristic frequencies (inner race at 136Hz) and so would include all harmonics of interest.

The bearings experiments have taken two week to be conducted; and the time-consuming part of the experiments was changing bearings. Swopping a bearing takes three hours. First the rig needed to be dismantled and the shaft taken out. Then the first tested bearing was slid out and the next bearing to be tested slid onto the shaft. This part of the process required the use of a pressing machine. Then the bearing was replaced and the rig re-assembled. This procedure had to be repeated every time with all bearings.

The order of the experimental procedure was healthy bearing, bearing with outer race fault, bearing with roller fault and finally bearing with inner race fault. In each case the 30% fault was tested first followed by the 60% fault and then the 100% fault.

CHAPTER SIX

Results and Discussion

This Chapter reports the analysis of the bearing data collected using conventional methods and the TESPAR approach developed by the author to determine whether it is suitable for use with bearing condition monitoring. The results presented include the envelope technique, wavelet and the TESPAR. Eextraction of TESPAR streams and codes for the vibration signals from both healthy and faulty bearings is implemented. The construction of both S and A-Matrices is described for each fault. A method for assessing the similarity of the S and A-Matrices is developed using a correlation technique to determine both the presence and type of fault. This chapter reports the successful development of a fault severity classification based on the TESPAR system and shows that such an assessment can be a direct measure of fault severity. The results indicate that TESPAR can be a suitable method for detection and diagnosis of faults in bearings.

6.1   Initial Results and Discussion

The most obvious technique is to visually inspect the time-domain waveform. Figure 6-1 shows a 0.25 second sample of a vibration signal from two N(406) bearings where the sampling rate was 62.5 kHz and rotational speed was 1460 rpm (24.3 Hz). There are two traces present one is from a bearing with a small outer race fault (blue trace) and the other is from a bearing with a large inner race fault (red trace). It can be seen that the amplitude of the red trace is significantly larger than that of the blue trace.
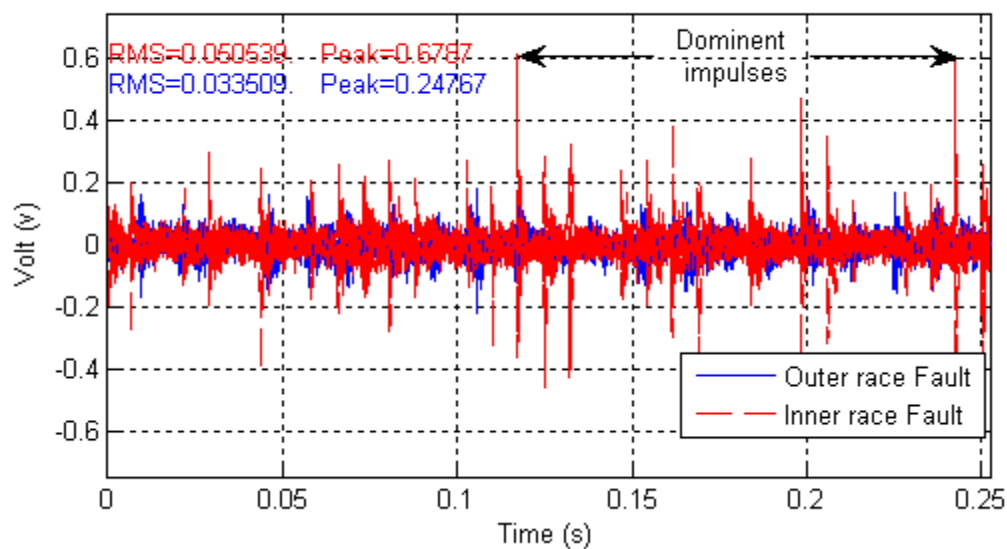
Figure 6-1 Time domain waveforms for small outer and big inner races faulty bearings

Figure 6-1 shows repetitive impacts at intervals corresponding to the time interval between the rolling elements passing the point on the outer race and the inner race where the faults are situated. It is also seen that these repeated impulses are amplitude modulated with a similar patterns repeated for every revolution.

Each impact of a passing element with the fault excites damped resonances in the bearing structure that die away rapidly. Unfortunately, vibration signals from bearings do not always produce a signal showing the impacts as clearly as in Figure 6-1. The vibration signal from a

142

sensor will be the sum of all the vibrations from the machine arriving at that point and will contain many components. It is unlikely that simply viewing the time-domain signal will detect, for example, a spalling defect in the bearing [123]. Certainly at the inception stage small impulses due to spalling will be masked by background noise generated by vibrations from other components. The present means of overcoming this problem is to extract certain characteristic parameters from the signal and trend them with time.

Typically vibration signals from real bearings are non-stationary and non-deterministic. Time-domain feature extraction is done by determining statistical parameters, which provides information about overall level and the "spikiness" of the signal associated with the defect induced impulses [124]. Commonly used time-domain parameters are RMS, peak value (Pv), crest factor (Cf), kurtosis (Kv) and impulse factor (If).

The probability density distribution PDF has also been found useful [124, 125]. The RMS will directly reflect the energy level of the vibration. The Peak value is the maximum value attained by the signal. Kv and Cf indicate the degree of spikiness of the signal and are useful with defect-induced impulses.

Consider two signals $S1$ and $S2$ with the same peak value. Suppose $S1$ has the greater energy but is less spiky, while S2 has less energy but is spikier. We should observe that $S1$ has the greater RMS value but S2 has larger Kv and Cf values.

Cf, Kv and If are non-dimensional and can be used to indicate incipient spalling due to fatigue. However, if the fault develops and becomes sufficiently severe, then Kv and Cf return to the values observed for a normal bearing. Thus these parameters are not often used in isolation.

When a fault develops in a bearing, the vibration signal will exhibit sharp impulses of relatively low amplitude, these will appear with incipient damage and as the fault develops the vibration energy level will first increase significantly and then substantially. Figure 6-2 shows vibration signals taken from a normal "healthy" bearing and a bearing with a fault in the outer race. It can be seen that the vibration amplitude of the defective bearing is significantly higher than for the normal bearing both in terms of peak value and RMS.
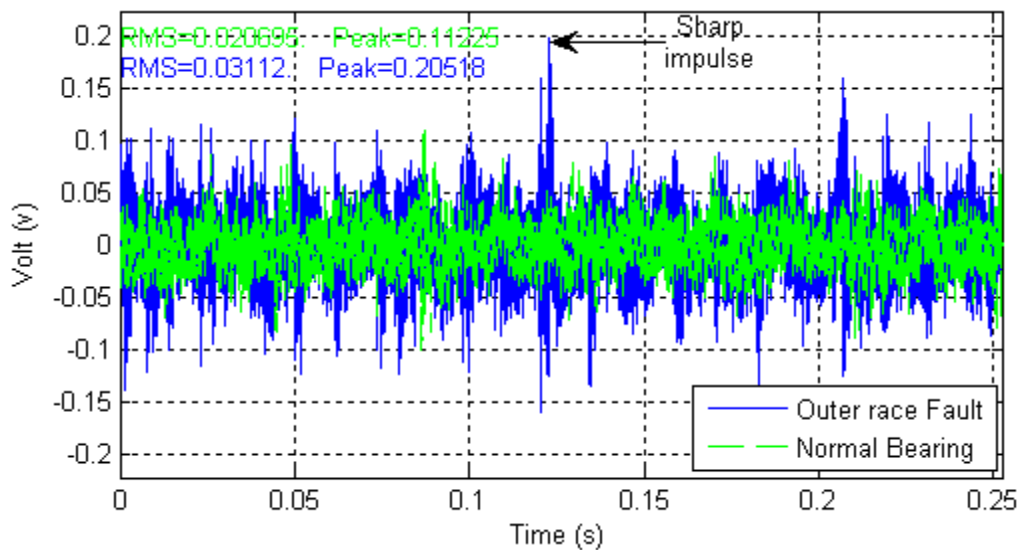
Figure 6-2 Vibration signals for normal bearing and bearing with outer race fault

In frequency domain, the presence of peaks in the vibration spectrum can be used to identify both the type of fault and its location. For detection and diagnosis of bearing faults, the characteristic frequencies associated with specific faults need to be known and specimen calculations of these are presented in Table 2-3. Figure 6-3 shows the vibration spectrum of a healthy / normal bearing and a bearing with a fault in the outer race. The initial dominant peak in the spectrum can be easily identified by the difference between the healthy and faulty spectra.

However, automatic detection of faults using characteristic bearing frequencies is not simple; it requires the spectrum be searched for specific frequencies and their harmonics and

144

sidebands [175]. One major difficulty is that the vibration energy of the bearing is wide-band noise which can easily bury specific frequencies; another is that impacts generate wide band noise which can excite resonances in the bearing and surrounding structure. Spectrum comparison has also been investigated for the purpose of signature analysis [176].
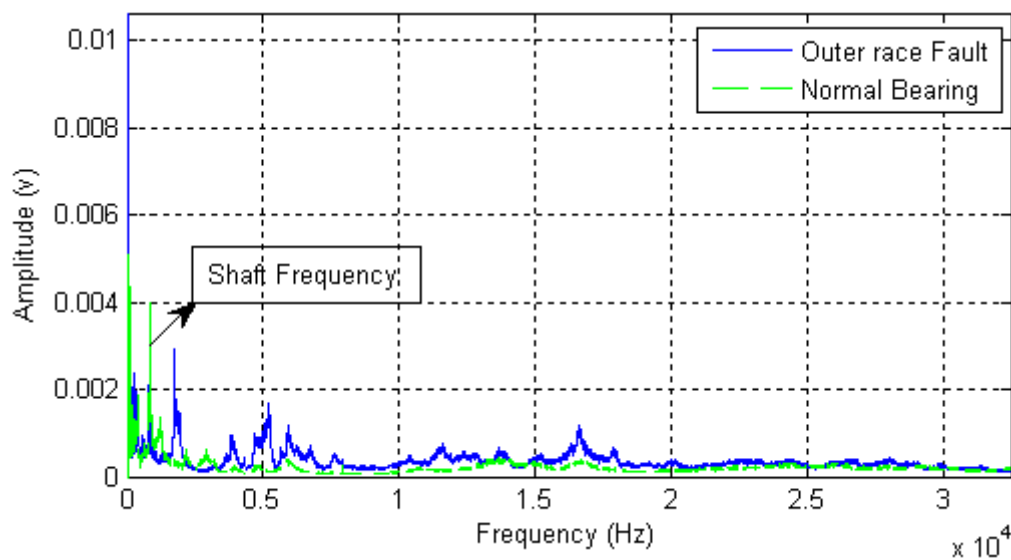

Figure 6-3 Vibration spectra of normal bearing and bearing with outer race fault

Here the baseline spectrum against which the others are to be compared is for a machine (or component) in a healthy condition. A difficulty here is that often there can be considerable divergence between the spectra for apparently identical healthy machines, and some form of averaging has to be undertaken. Any difference which subsequently develops between the baseline spectrum and the measured spectrum is taken as an indication of changes in the mechanical condition of the machine (or component). The comparison identifies frequencies at which significant changes in amplitude have taken place. In Figure 6-3 the baseline spectrum was determined for the healthy bearing. Then the measured signal with fault in the outer race is plotted on the same axes and compared with the baseline. Subtraction produces a 'difference' spectrum which forms the basis for decision-making. A weakness with this technique is that often incipient faults produce insufficient energy to make a noticeable

145

difference in the spectrum compared to the total machine component vibration. Another problem is that the vibration spectra from bearings vary with bearing speed and loading.

6.2   Base-Lining Analysis and Results

6.2.1   Envelope Spectrum

As explained in Chapter three, it is well known that a popular method for CM of bearings is to use the envelope spectrum. To check if the data contains the information of different faults a common envelope procedure is applied to the data to check the character frequencies are shown in the envelope spectrum. Moreover it will be used as a reference to benchmark the results obtained late from TESPAR analysis.
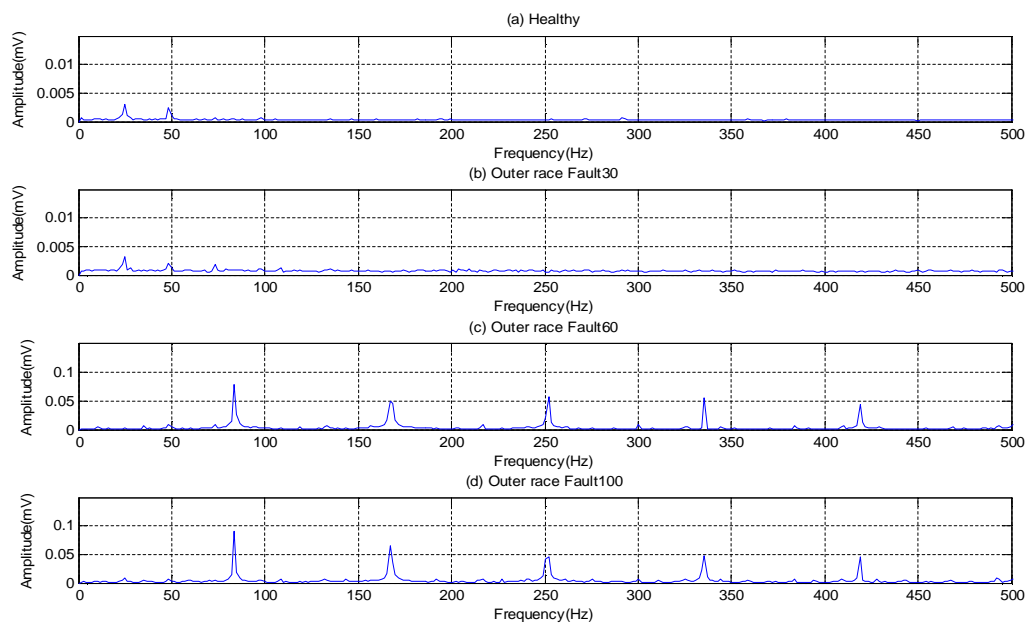
**Figure 6-4 Envelope spectrums for healthy and three different
degrees of outer race faulty bearings**

Figures 6-4 to 6-7 show the envelope spectrum of the datasets for all the bearings tested in the frequency band from 8 kHz to 15 kHz. With no fault present the envelope spectrum of the vibration signal from a bearing show very little as can be seen by the results in plots labelled

(a) in Figures 6-4, 6-5 and 6-6 compared with the plots obtained when a fault was present the envelope spectrum for the healthy bearing is very flat with no clear spectral lines to be seen. Similarly with Figure 5-9 (b) which represents the smallest of the outer race faults, there is little indication of the presence of a fault. However, Figure 6-4 (c) and (d) show the characteristic outer race fault frequency (83.3 Hz) clearly indicating the presence of an outer race fault.

Figure 6-5 shows the results obtained for roller faults. Here peaks at the characteristic fault frequency (97.3 Hz) can be seen even with the smallest fault. The amplitude of these peaks increased as the fault severity increased.
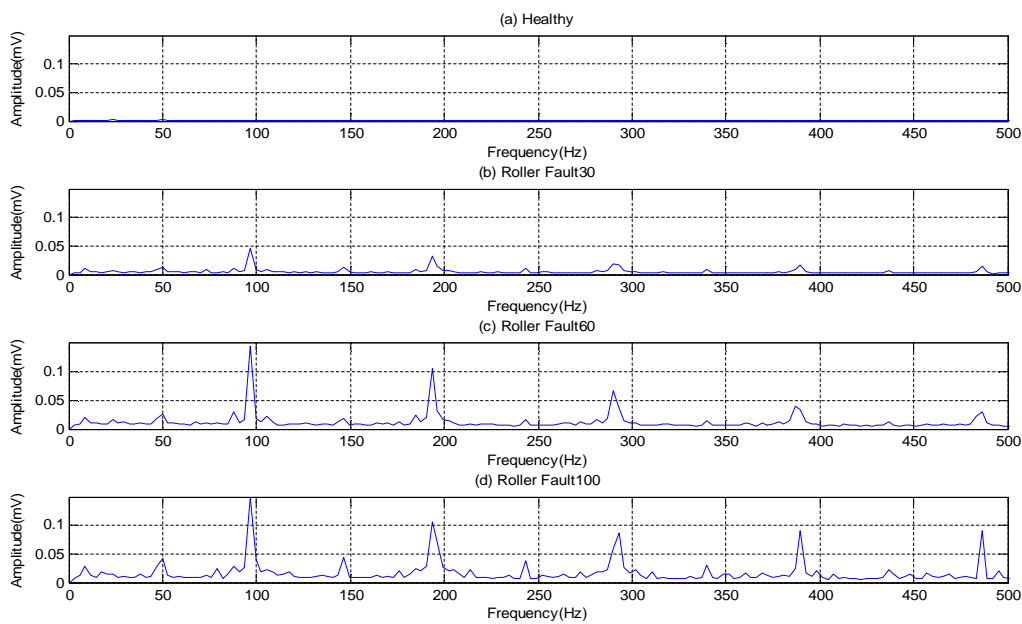
Figure 6-5 Envelope spectrums for healthy and three different degrees of roller faulty bearings

Figure 6-6 shows the results obtained for inner race faults. Here peaks at the characteristic fault frequency (134.4 Hz) can be just discerned even with the smallest fault. Once again the amplitude of these peaks increased as the fault severity increased.
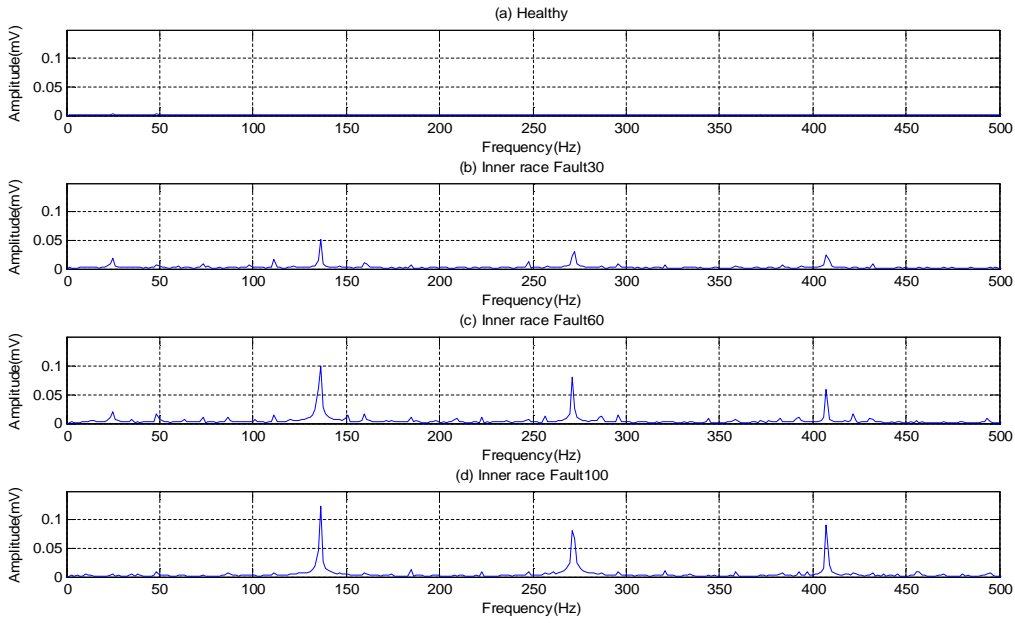
Figure 6-6 Envelope spectrum for healthy and three different degrees of inner race faulty bearings

In Figure 6-7 the spectra of the roller and the inner race faults are quite clear as shown in Figure 6-7 (c) and Figure 6-7 (d) respectively. The roller and the inner race characteristic faults' frequencies are identified at 97.3 Hz and 134.4 Hz respectively as mentioned above.
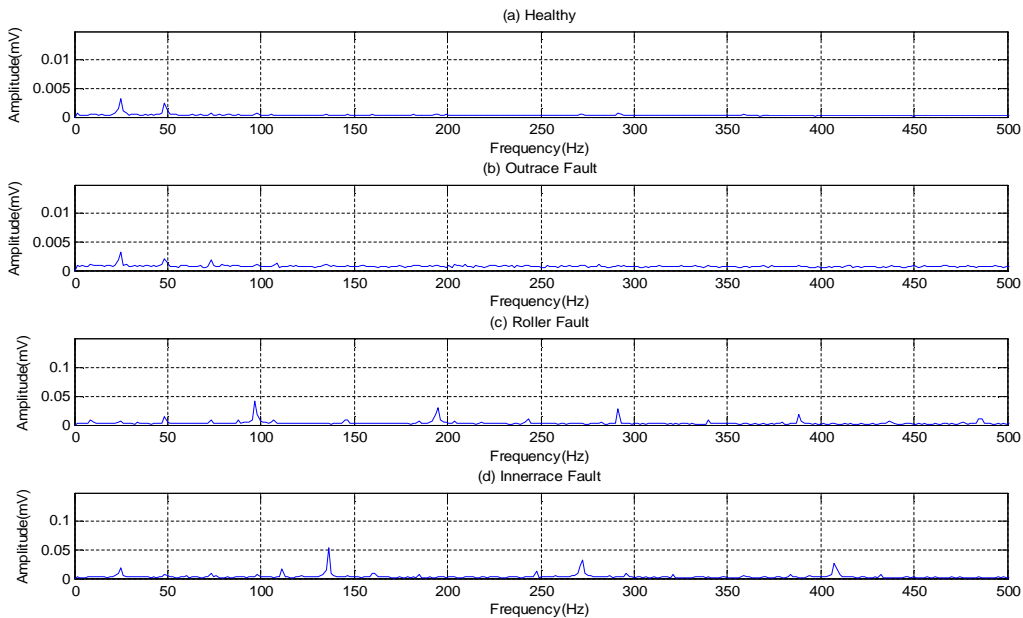

Figure 6-7 Envelope spectrum for healthy bearing and three different bearing faults

148

Moreover, the characteristic frequencies cannot be seen in envelope spectrum in 6-7 (a) which indicates healthy bearing spectra. This can be illustrated by the blue solid line in the Figure. In Figure 6-7 (b) the spectra representing the small outer race fault appears smooth and there is only a small amplitude which cannot be seen easily from the figure.

6.2.2 Wavelet Analysis

As explained in Chapter Three, there are different kinds of wavelets for different kinds of applications and the wavelet function of 4th order of Daubechie was chosen for the research because it gives the best match of the bearing defects signal which features strong non-stationary characteristics.

Figures 6-8, 6-9 and 6-10 show the results of applying wavelet analysis to the time domain datasets obtained for all the bearings tested.

Figure 6-8 shows the wavelet signal obtained for the healthy bearing and bearing with three severities of outer race faults. Figure 6-8 (a) shows the waveform for the healthy bearing and seems to be smooth and flat with no distinctive features except a very small peak at the shaft speed. Similarly in Figure 6-8 (b), the waveform of the bearing with a 30% outer race fault does not show any clear indication that a fault was present.
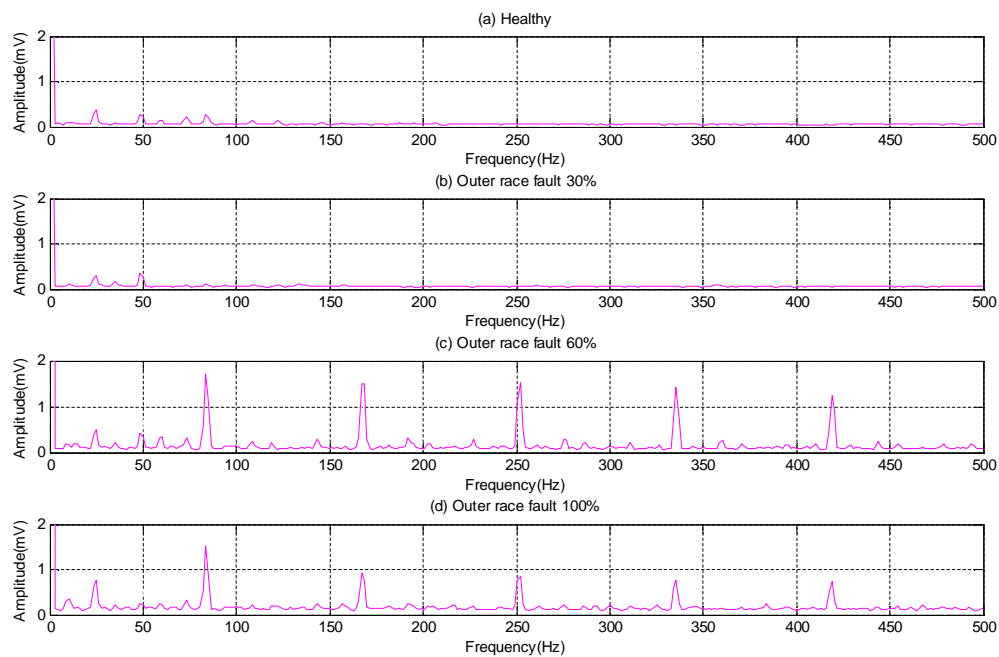
Figure 6-8 Spectrum of wavelet filtered signals for a healthy and three different severities of outer race faults

However, Figure 6-8 (c) and (d) which show the signal waveforms for the 60% and 100% outer race faults respectively clearly exhibit peaks at the characteristic outer race fault frequency of 83.3Hz and its harmonics. Unlike the envelope analysis the amplitude of the peaks does not increase as the severity of the fault increases.

Figure 6-9 shows the wavelet results obtained for roller faults. Here peaks at the characteristic fault frequency (97.3 Hz) and its harmonics can be seen even with the smallest fault, and the amplitude of the peaks increased with severity of the fault sufficiently for the severity of the faults to be differentiated.
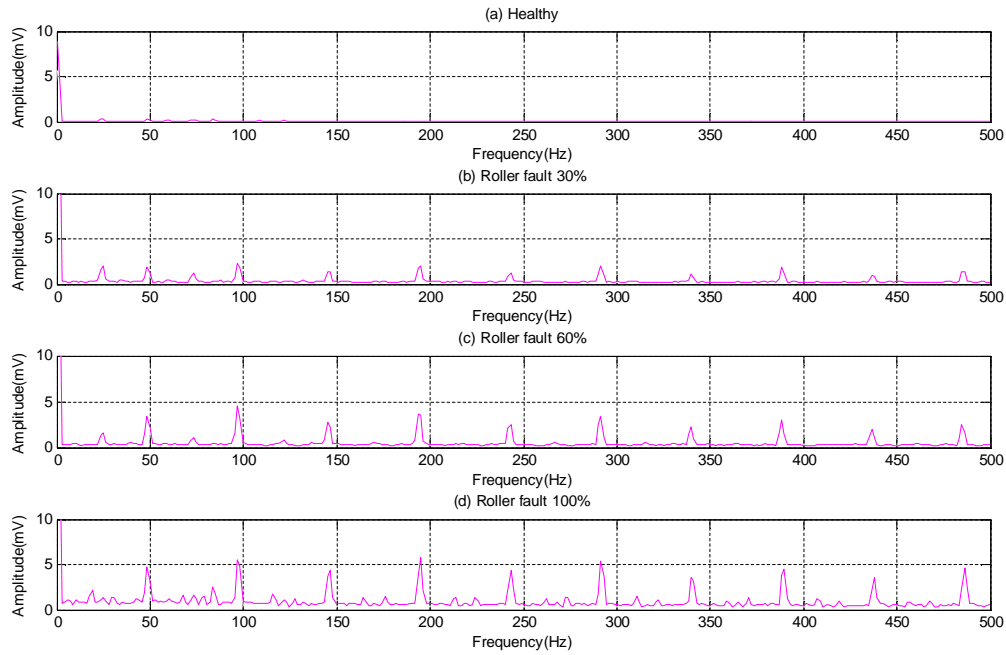
Figure 6-9 Wavelet signals for a healthy and three different degrees of rollers faults

Figure 6-10 shows the wavelet results obtained for inner race faults. Here peaks at the characteristic fault frequency (134.4 Hz) can be just discerned even with the smallest fault. However, the 60%, 100% inner race fault shows very similar waveforms which makes it very difficult to determine the severity.

From Figure 6-11 it can be seen that the least severe fault, in the case of the outer race the fault cannot be identified easily. In the case of the roller element fault characteristic defect frequencies could be located at 97.3 Hz and its harmonics. But with the inner race fault there is clear indication of peaks at the characteristic frequency of 134.4 Hz and its harmonics, the inner race defect seems to generate the largest amplitude peaks.
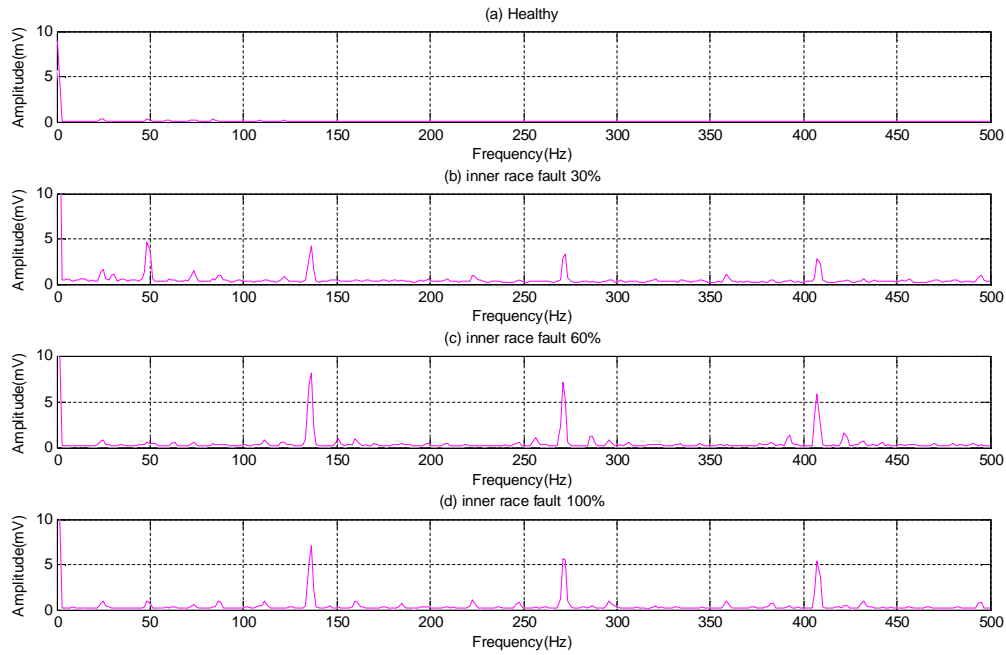
Figure 6-10 Wavelet signals for a healthy and three different degrees of inner race faults
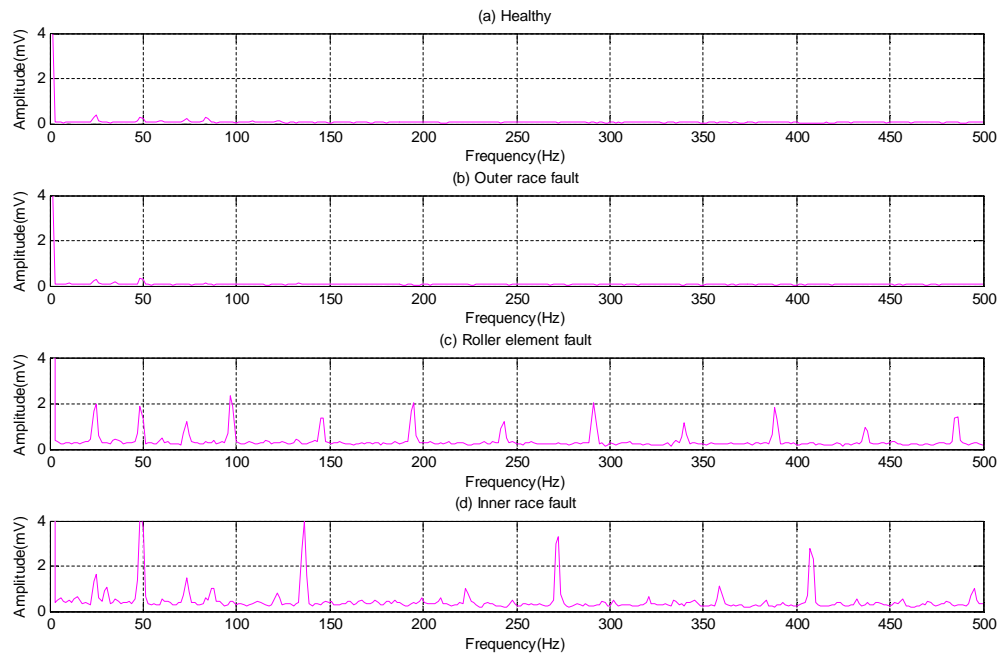

Figure 6-11 Wavelet signals for a healthy and three different small bearings faults

Thus, three conventional techniques of time domain, envelope spectrum and wavelet analysis have been used with the ten bearings to provide a base-line with which to assess the effectiveness of the TESPAR analysis.

152

6.3   TESPAR Processing Specification

The bearing data being processed, detailed in Chapter Five, represents ten conditions, with a healthy bearing as baseline and with three different fault locations; outer race, inner race and roller (each of these fault locations having three levels of severity).

The TESPAR processing applied to both the classification of fault location and fault severity comprises relatively few steps:

- Step 1: Load each raw data signal

- Step 2: Filter the signal using a 2nd order Butterworth bandpass filter with filter band of 250Hz–17000Hz.

- Step 3: Each of the filtered signals is passed into the TESPAR function to produce the TESPAR stream. Each epoch in the filtered signal is converted into a magnitude, duration and shape component, thus producing a 3xN length vector where N is the number of epochs within the signal.

- Step 4: Each TESPAR stream is passed into the Tescode function to convert the TESPAR stream into the TESPAR codes vector. The duration and shape component of each epoch is converted into a TESPAR code by using the look-up table. The output of this step is a 2xN length vector comprising the magnitude and TESPAR code associated with each epoch.

  To classify the fault location:

  - Step L5: Generate TESPAR S and A-Matrices for each signal by using the corresponding vector codes output from Step 4.

- o Step L6: Generate a TESPAR archetype for each bearing condition.

- o Step L7: Correlate each matrix produced in step L5 with each matrix archetype produced in step L6. Whichever archetype produces the closest match with the matrix under test is determined to be the winner and thus the predicted fault location.

To classify the fault severity:

- o Step S5: For each signal, find all epochs that have durations greater than the average duration calculated from the signal as a whole. Generate the average absolute magnitude value from all of the epochs that are identified.

- o Step S6: The average absolute magnitude value generated from each signal under test is used to identify the fault severity.

6.4 Fault Location Classification

This section describes in more detail how TESPAR classifies the different fault locations by using the S and A-Matrices with the use of correlation. The S matrix is a 2 dimensional representation of TESPAR codes, as explained in Section 4.5.1 that have been generated from the TESPAR coding stream. In its raw form, the S-Matrix is simply a histogram of each TESPAR code on the X-axis and the number of occurrences of each code on the Y-axis.

The A-Matrix is a 3 dimensional (28×28) matrix, as explained in Section 4.5.2, that provides the number of occurrences of each combination of TESPAR codes (with the X and Y axes representing the combination of TESPAR codes and the Z axis representing the number of occurrences of each TESPAR code combinations). The discussion below shows that the correlations between the test S-Matrices and the S-Matrices archetypes for faults and the test

154

A-Matrices and the A matrices archetypes for faults works well and can be used as a good measure for detection of fault location in rolling element bearings.

6.4.1 Classification of Bearing Faults Using S-Matrix

This section presents the matrices and results when using TESPAR S-Matrices to classify the fault location. Figures 6-12 (a), (b), (c) and (d) show the three examples of bearing normal conditions and their archetype.
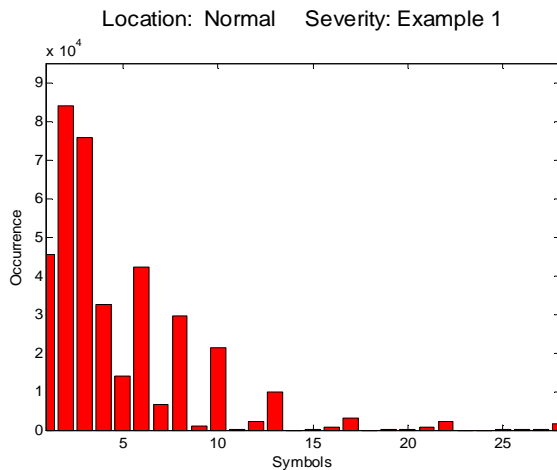

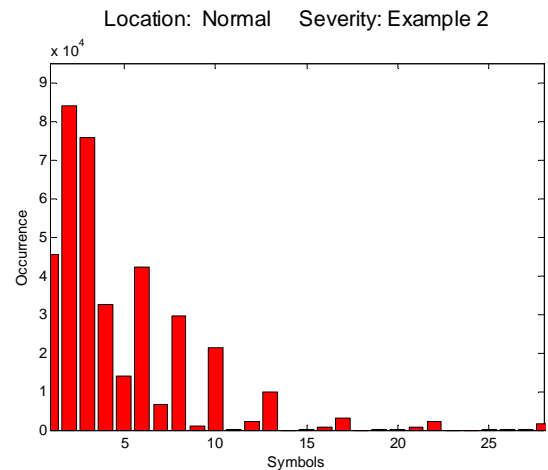**Figure 6-12 (a) S-Matrix for the first Normal condition**

Figure 6-12 (b) S-Matrix for the second Normal condition
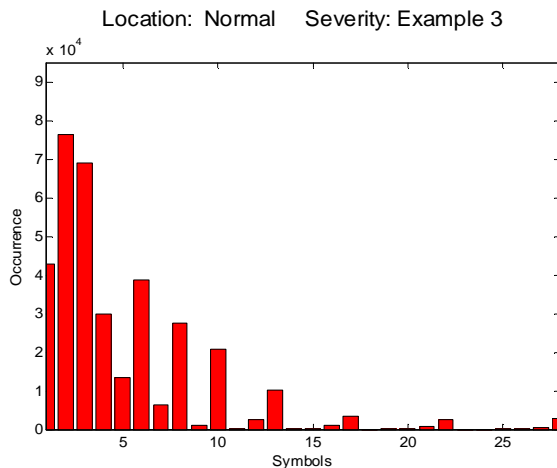

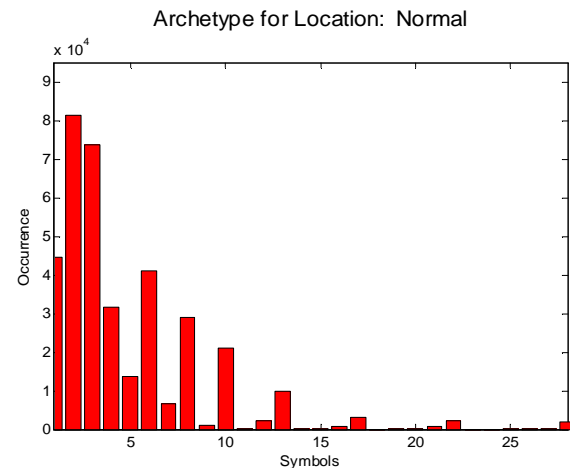**Figure 6-12 (c) S-Matrix for the third Normal condition**

Figure 6-12 (d) S-Matrix Archetype For the Normal condition

Also Figures 6-13 (a), (b), (c) and (d) show the three different degrees of outer race faults and their archetype. Figures 6-14 (a), (b), (c) and (d) show the three different degrees of roller faults and their archetype and Figures 6-15 (a), (b), (c) and (d) show the three different

155

degrees of inner race faults and their archetype respectively. It can be seen that most of S-Matrices have symbol values less than about 10 which mean the signals being presented contain a greater proportion of higher frequency epochs. S-Matrices populated by symbols in the range 20-30 would indicate a greater proportion of longer duration epochs and a lower frequency signal.
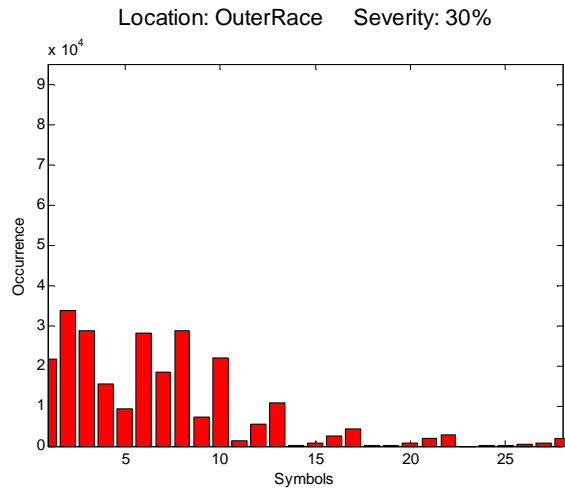

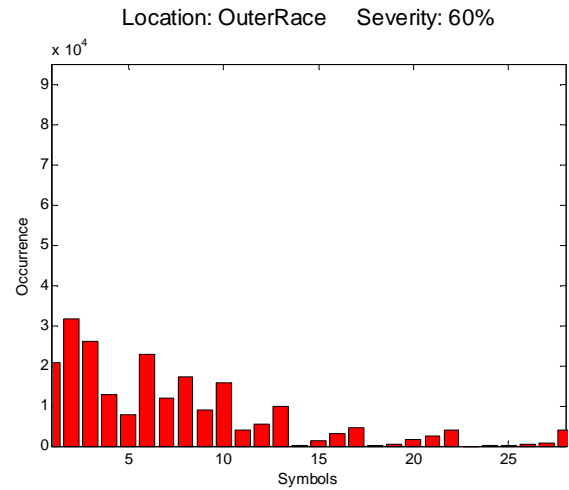**Figure 6-13 (a) S-Matrices for the 30% fault of the Outer Race**

Figure 6-13 (b) S-Matrices for the 60% fault of the Outer Race

Figure 6-13 (c) S-Matrices for the 100% fault of the Outer Race

Figure 6-13 (d) S-Matrix Archetype for the Outer Race Fault

156

Figure 6-14 (a) S-Matrices for the 30% fault of the Roller

Figure 6-14 (b) S-Matrices for the 60% fault of the Roller

Figure 6-14 (c) S-Matrices for the 100% fault of the Roller

Figure 6-14 (d) S-Matrix Archetype for the Roller

Figure 6-15 (a) S-Matrices for the 30% fault of the Inner Race
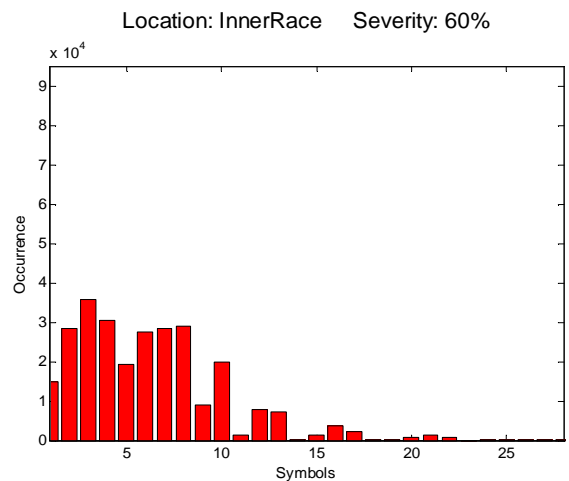
Figure 6-15 (b) S-Matrices for the 60% fault of the Inner Race

Location: InnerRace   Severity: 100%

Archetype for Location: InnerRace

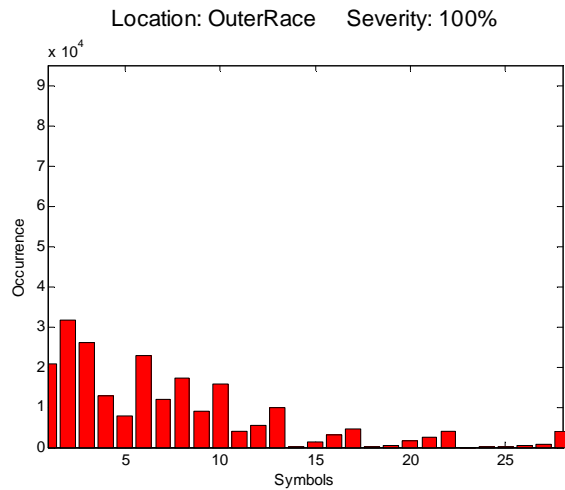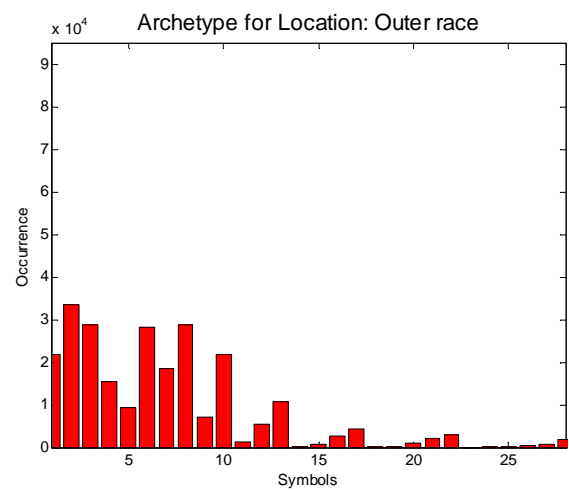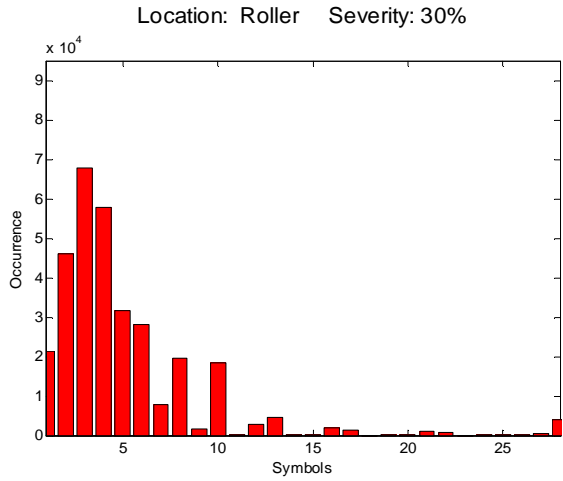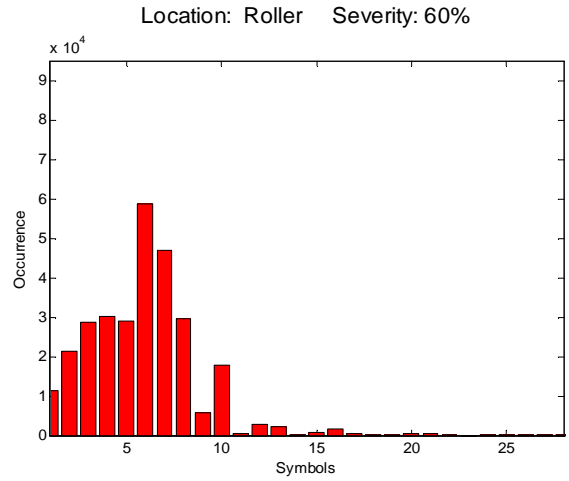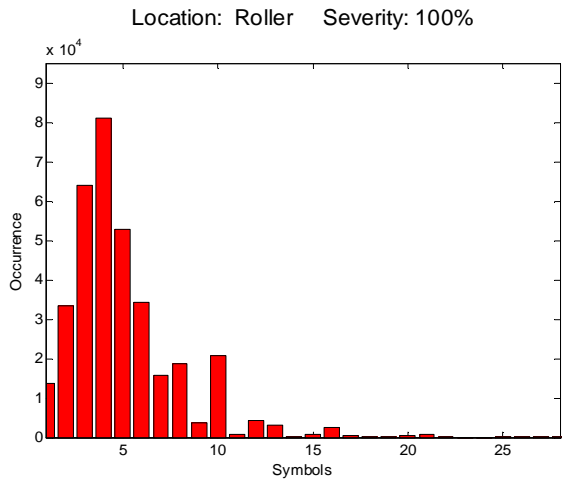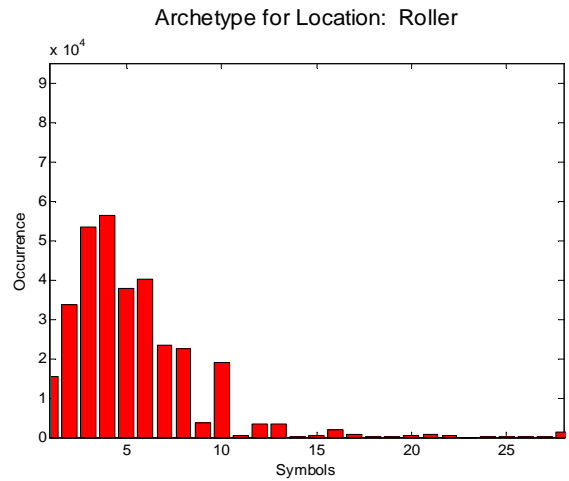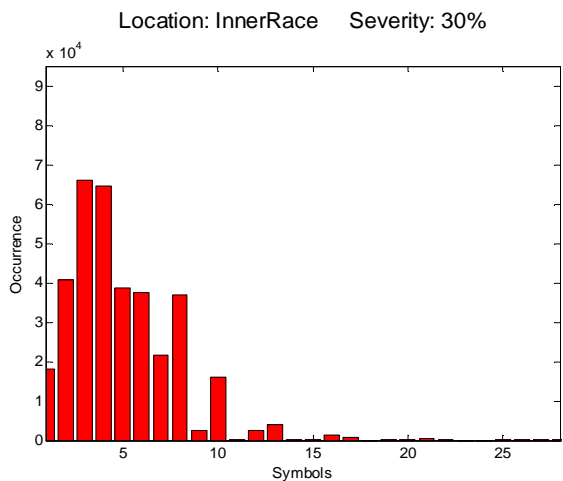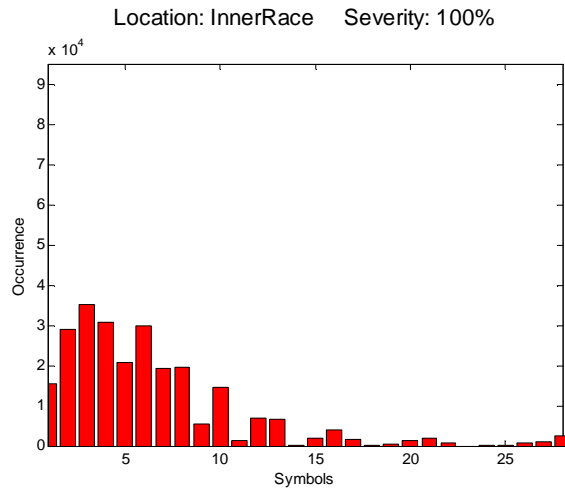**Figure 6-15 (c) S-Matrices for the 100% fault of the Inner Race**
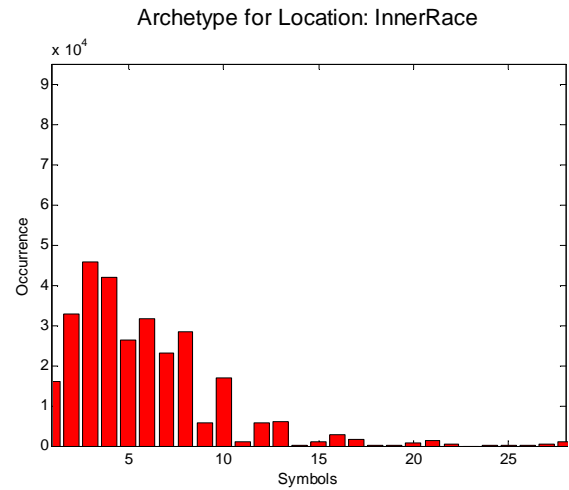
Figure 6-15 (d) S-Matrix Archetype for the Inner Race Fault

The S-Matrix archetypes associated with each fault location are simply the average of the S-matrix examples associated with each fault location.

The next step is to compare each of the twelve S-Matrices (three examples of healthy, three faulty outer races, three faulty rollers and three faulty inner races) with the four S-matrix archetypes. The "winning" bearing condition is the one producing the highest correlation score and thus the closest to one of the four archetype conditions. This test is to show that the different fault conditions are clearly distinguishable from each other using this technique. The correlation score produced is in the range 0-100%, with a score of 100% indicates that the profile of the two matrices are identical**.**

Table 6-1 shows the correlation scores produced by comparing the S-matrix generated from each signal under test with each S-matrix reference archetype associated with the healthy (normal) condition and three fault locations. The results presented on each row of the table are the percentage comparison scores produced when correlating each individual example against each of the four condition archetypes. The winning (highest) correlation score on each row is in bold. In a real-world scenario to identify the fault location, an S-Matrix will be

158

generated from the signal under test and correlated against each of the four S-Matrix reference archetypes with the highest score indicating the fault location. It can be seen from the table that the faults are clearly distinguishable from each other.

Table 6-1 Correlation scores for S-matrix archetypes

	Archetype: Normal	Archetype: Outer Race	Archetype: Roller	Archetype: Inner Race
Normal Example 1	**99.96%**	74.95%	44.99%	55.38%
Normal Example 2	**99.96%**	74.95%	44.99%	55.38%
Normal Example 3	**99.83%**	74.31%	44.12%	54.53%
Outer Race Small	74.62%	**93.60%**	66.60%	69.04%
Outer Race Medium	67.81%	**95.64%**	74.29%	71.92%
Outer Race Large	46.25%	**91.54%**	45.99%	57.71%
Roller Small	64.21%	70.66%	**90.72%**	74.67%
Roller Medium	23.67%	59.56%	**88.11%**	70.45%
Roller Large	40.49%	58.51%	**92.90%**	74.02%
Inner Race Small	55.57%	74.92%	73.91%	**93.96%**
Inner Race Medium	47.07%	74.42%	71.56%	**95.62%**
Inner Race Large	53.32%	68.92%	70.81%	**95.37%**

The correlation method used to generate the results in Table 6-1 was generated by calculating the angular correlation between the S-matrix under test and each of the four reference archetypes as explained in Chapter Four Section 4.9.

The next section discusses the results obtained using the TESPAR A-matrices to classify the fault location.

6.4.2 Classification of Bearing Faults Using A-Matrix

This section presents the results of A-Matrices when using TESPAR to categorize the location of faults.

Figures 6-16 (a), (b), (c) and (d) show the three examples of A-Matrices for the normal conditions and their archetype. Figures 6-17 (a), (b), (c) and (d) show the A-Matrices of the three different degrees of outer race faults and their archetype. Figures 6-18 (a), (b), (c) and (d) show the A-Matrices of the three different degrees of roller faults and their archetype and Figures 6-19 (a), (b), (c) and (d) show the A-Matrices of the three different degrees of inner race faults and their archetype respectively.

The matrices are primarily populated on the left-hand side signifying that the bearing data contains a greater proportion of TESPAR codes in the range 0-10 as demonstrated in Figures 6-12, 6-13, 6-14 and 6-15.
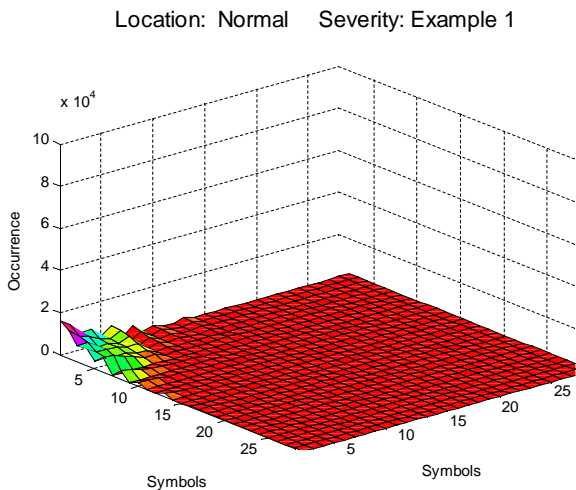
Figure 6-16 (a) A-Matrix for the first Normal condition
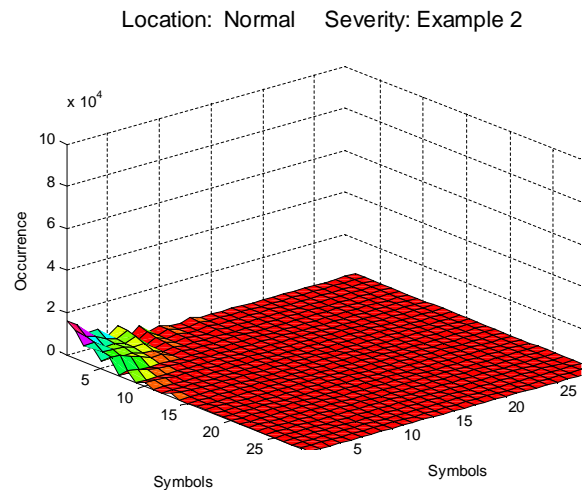
Figure 6-16 (b) A-Matrix for the second Normal condition
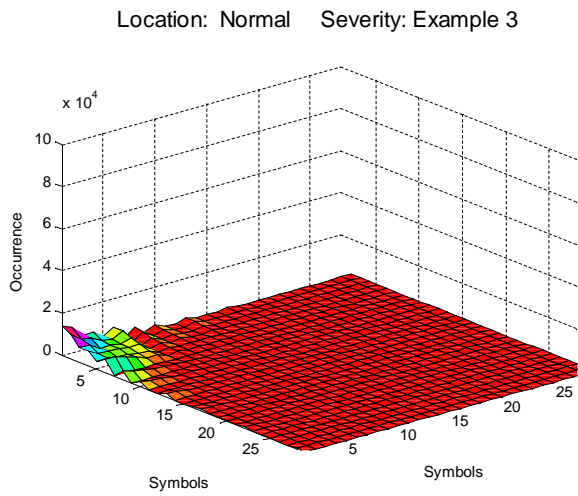
Figure 6-16 (c) A-Matrix for the third Normal condition
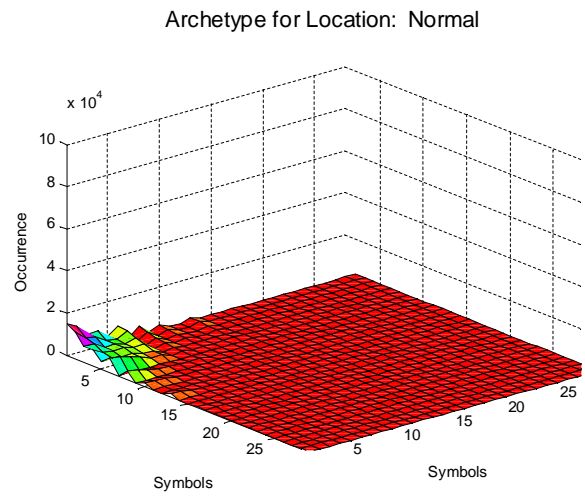
Figure 6-16 (d) A- Matrix Archetype for the Normal condition
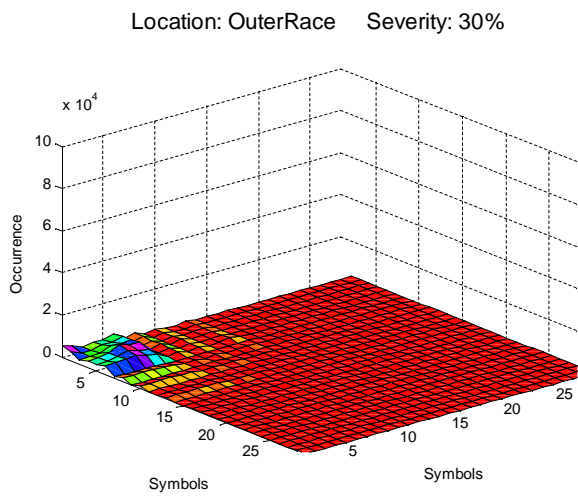
Figure 6-17(a) A- Matrices for the 30% fault of the Outer Race
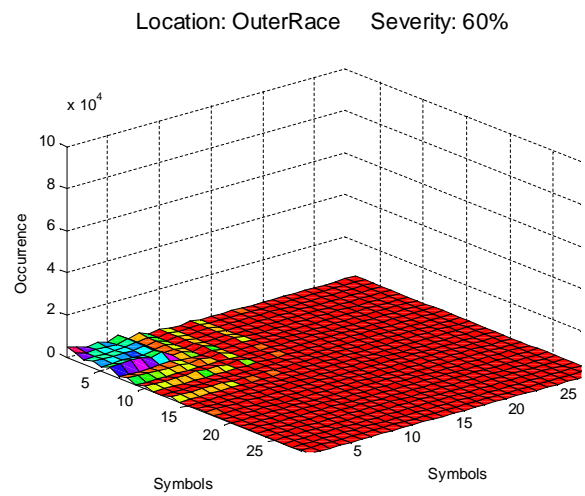
Figure 6-17 (b) A- Matrices for the 60% fault of the Outer Race
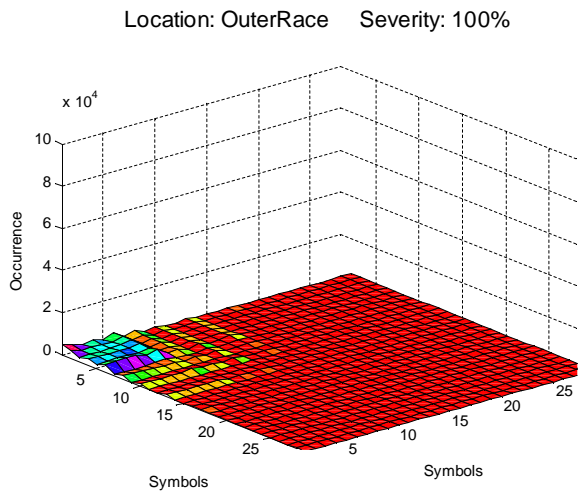
Figure 6-17 (c) A- Matrices for the 100% fault of the Outer Race
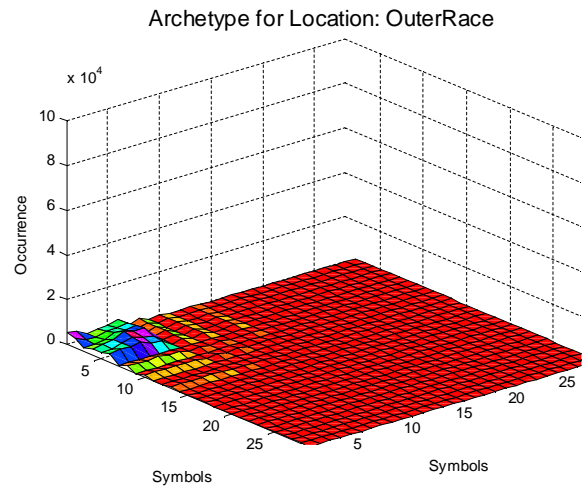
Figure 6-17 (d) A- Matrix Archetype for the Outer race fault
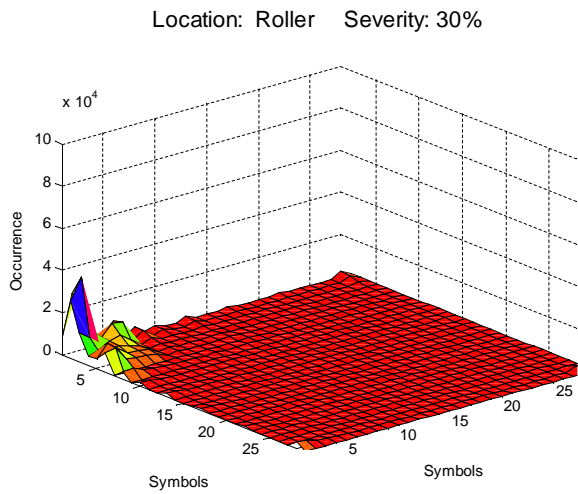
Figure 6-18 (a) A- Matrices for the 30% Roller fault
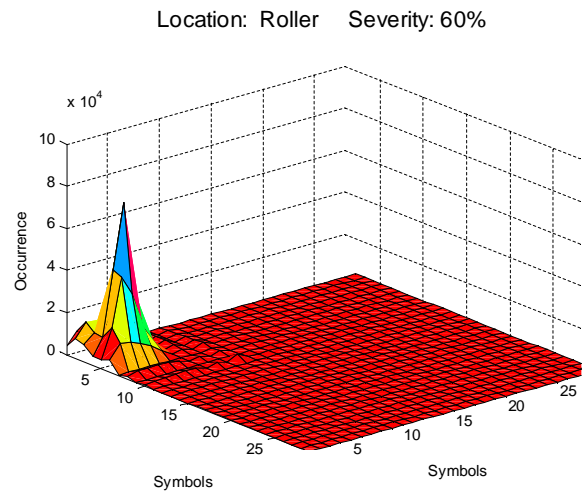
Figure 6-18 (b) A- Matrices for the 60% Roller fault

162
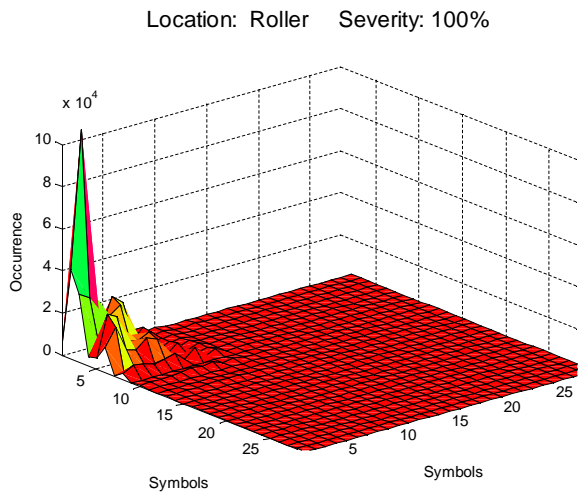
Figure 6-18 (c) A- Matrices for the 100% Roller fault
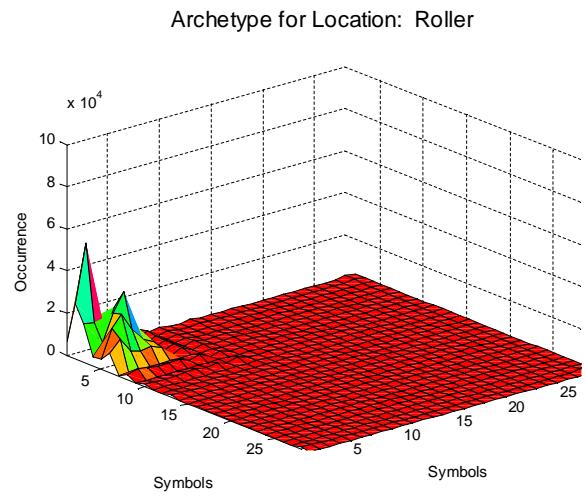
Figure 6-18 (d) A- Matrix Archetype for the Roller fault
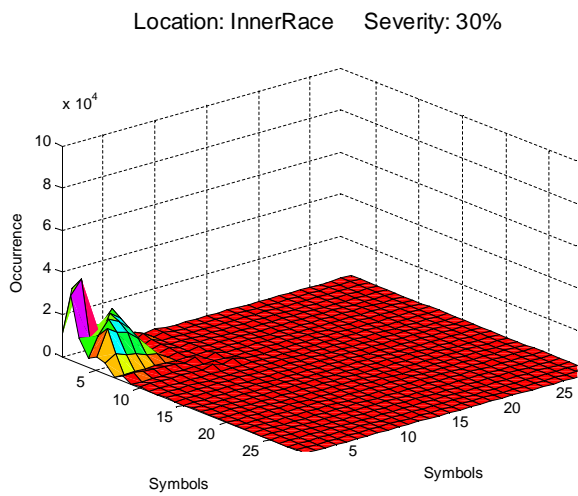
Figure 6-19 (a) A- Matrices for the 30% Inner Race fault
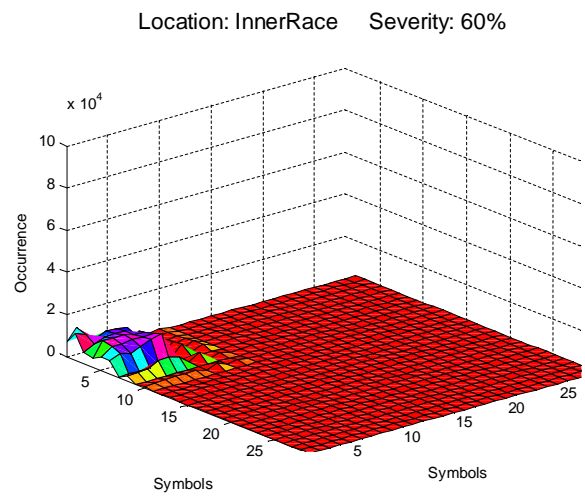
Figure 6-19 (b) A- Matrices for the 60% Inner Race fault

Location: InnerRace   Severity: 100%

Archetype for Location: InnerRace
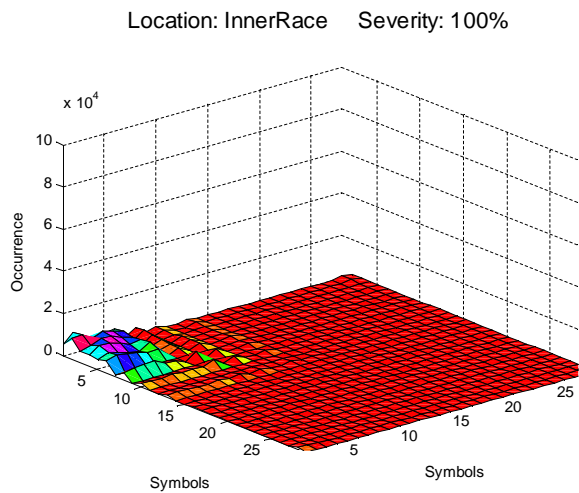
**Figure 6-19 (c) A- Matrices for the 100%
Inner Race fault**
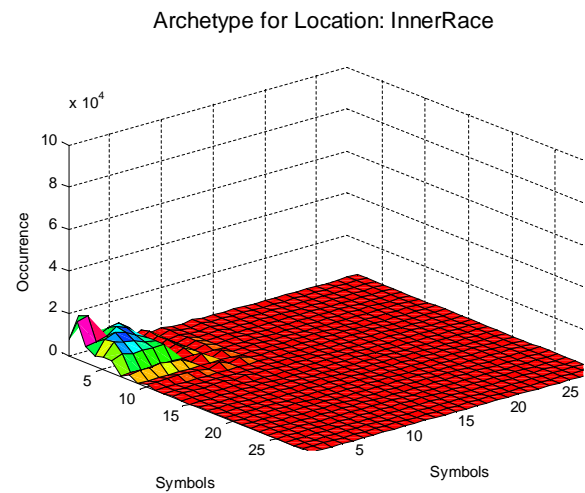
**Figure 6-19 (d) A- Matrix Archetype
for the Inner Race fault**

Figures 6-16 (d), 6-17 (d), 6-18 (d) and 6-19 (d) display the A-matrix archetypes associated with each fault location. As mentioned above in Section 6.2.1, each archetype is the average of the A-matrix examples associated with each of the three fault locations. As described above, it requires very little processing to generate the archetypes.

The same steps as described in Section 6.2.1 are repeated here for the A-Matrices to compare each of the twelve A-matrices (three examples of healthy, three faulty outer races, three faulty rollers and three faulty inner races) with the four A-matrix archetypes. The "winning" bearing condition is the one generating the highest correlation score and thus the nearest to one of the four archetype conditions.

Table 6-2 presents the correlation scores generated by comparing the A-Matrix produced from each signal under test with each A-matrix reference archetype associated with the normal and three fault locations.

The results presented on each row of the tables are the percentage comparison scores produced when correlating each individual example against each of the four condition

164

archetypes. The winning (highest) correlation score on each row is in bold. In a real-world scenario to identify the fault location, an A-matrix will be generated from the signal under test and correlated against each of the four A-matrix reference archetypes with the highest score indicating the fault location.

Table 6-2 correlation scores for A-matrix archetypes

	Archetype: Normal	Archetype: Outer Race	Archetype: Roller	Archetype: Inner Race
Normal Example 1	**99.98%**	74.51%	62.67%	70.95%
Normal Example 2	**99.98%**	74.51%	62.67%	70.95%
Normal Example 3	**99.89%**	72.32%	60.76%	69.31%
Outer Race Small	67.17%	**93.90%**	59.38%	70.65%
Outer Race Medium	59.43%	**94.16%**	65.11%	69.11%
Outer Race Large	38.35%	**92.56%**	31.93%	49.00%
Roller Small	73.48%	57.88%	**91.84%**	74.70%
Roller Medium	56.26%	52.30%	**85.90%**	60.93%
Roller Large	30.53%	47.49%	**88.80%**	72.55%
Inner Race Small	68.89%	70.47%	74.85%	**90.16%**
Inner Race Medium	65.07%	71.51%	71.77%	**95.49%**
Inner Race Large	60.94%	70.52%	69.99%	**93.35%**

The reason that the correlation method works when separating the different fault locations is predominantly due to the TESPAR matrices being able to extract sufficient information from the signals of interest to be able to adequately represent the various fault locations.

The success of the solution proposed in this thesis is due to the combination of the TESPAR processing specification with a powerful correlation technique. During these investigations into the use of TESPAR both the S-Matrices and A-Matrices were able to correctly separate the fault locations without generating any errors when using correlation techniques, a general rule is that confidence in any solution is higher if a classification method can be used.

6.5   Fault Severities - Results and Discussion

TESPAR approach has various features which may be employed to different applications to investigate different practices. In other words, the previous section described how TESPAR approach was used to detect and classify various bearing fault locations using S and A-Matrices. However, it is not possible to use either S or A matrices to give a severity assessment for each type of the fault. Therefore, it is important to note that it is not necessary for TESPAR matrices to be used within every TESPAR generated solution. The TESPAR solutions may be generated by employing both TESPAR matrices and TESPAR statistics in isolation and in combination.

In this section we will see how a TESPAR feature called the mean absolute magnitude value used to determine the bearing faults severities for the conditions tested in this thesis.

Once a fault type has been identified using the S or A-Matrices, its corresponding fault severity can be estimated based on the information from TESTPAR stream. In particular the amplitude information, which has not been included in S or A-Matrices, will be utilised for fault severity classification.

As it shown in Chapter 4, the signal under test has been filtered and the TESPAR stream is generated, which produces a 3-by-N matrix, each representing the number of epochs, and for each epoch recording the magnitude (row 1), duration (row 2) and shape (row 3). The MATLAB code to generate the fault severity is two simple lines of code (N.B. TStream used in the code is the 3xN TESPAR stream):

```
Idx = find(TStream(2,:) >= mean(TStream(2,:)));
```

The line of code above finds all epochs that have durations greater than the average duration associated with the whole signal. The indices of all the epochs matching the search criteria are stored in the vector Idx.

```
SeverityPrediction = mean(abs(TStream(1,Idx)));
```

The line of code above generates the mean absolute magnitude value from all of the epochs matching the previous search criteria. The mean absolute magnitude value produced is then used as a direct indictor of fault severity.

The mean absolute magnitude values (fault severity values) associated with the bearing data are given in Table 6-3.

**Table 6-3 Severity values based on epoch
mean absolute magnitude**

Bearing conditions	Severity index
Normal_Example1	9.89
Normal_Example2	9.89
Normal_Example3	10.57
Small outer race	23.11
Medium outer Race	42.05
Large outer Race	55.90
Small roller	48.61
Medium roller	84.54
Large roller	92.12
Small inner race	49.56
Medium inner race	53.44
Large inner race	60.09

The reason why the mean absolute magnitude feature provides an indicator of fault severity is that the feature is essentially examining the magnitude of the most significant lower frequency events within the signal under investigation. In the time domain, longer duration

167

epochs are associated with lower frequency events (conversely the shorter the epoch, the more high frequency the sound). It is anticipated that for as the fault severity increases, the magnitude of the lower frequency epochs increases.

Figures 6-20, 6-21 and 6-22 display a graphical representation of the mean absolute magnitude values produced for each of the 12 signals under test.

Severity values for location: OuterRace
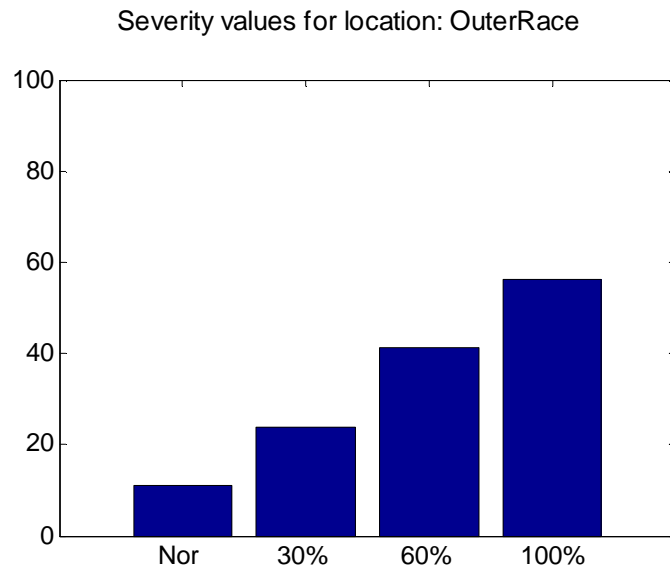
Figure 6-20 Mean absolute magnitude values for healthy and outer race location

Severity values for location: Roller
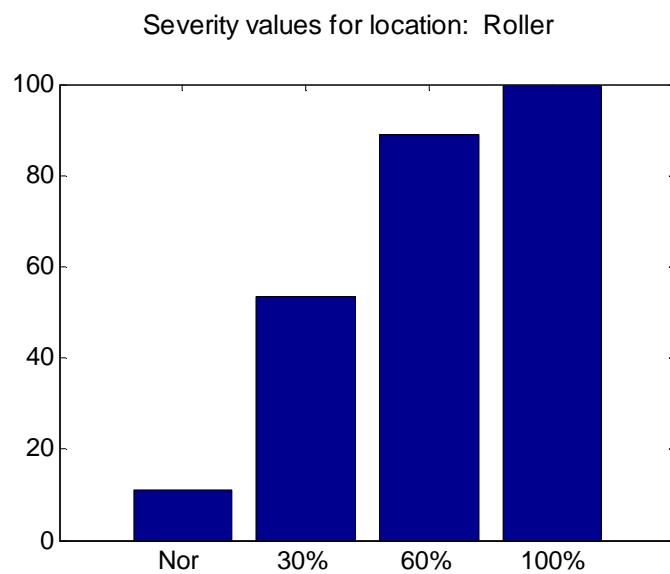
Figure 6-21 Mean absolute magnitude values for healthy and Roller Location

Severity values for location: InnerRace
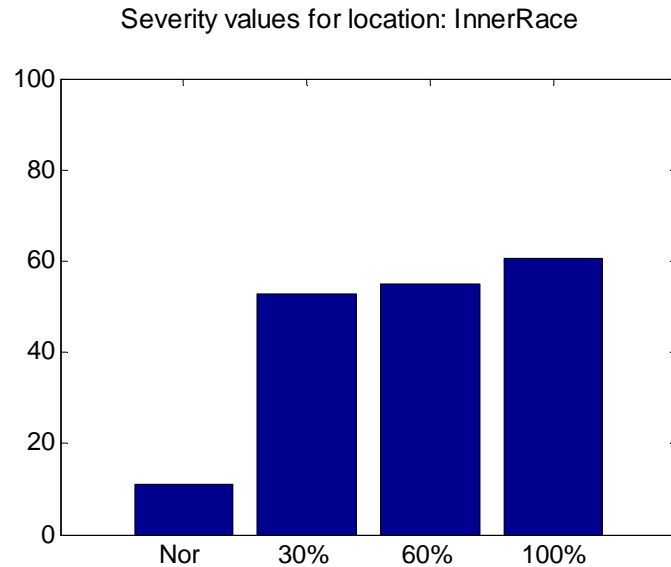
**Figure 6-22 Mean absolute magnitude values for healthy
and Inner Race Location**

The above results indicate that the mean absolute magnitude value could be used to determine the bearing faults severities. On its own severity assessment will certainly indicate the presence of even a small fault, but there is some overlapping of values which means it is not yet suitable for determining which fault is present. Once a fault is detected it is suggested that fault location will be investigated using the matrix correlation method described above, and once the fault location is known, its severity will be assessed by comparing the mean absolute magnitude with values listed in Table 6-3.

Using both methods (TESPAR matrices and the mean absolute magnitude values feature) in conjunction offers a simple technique to detect and diagnose rolling element bearing faults.

6.6   TESPAR Comparison with Conventional Methods Used

Conventional methods, such as envelope spectrum and wavelet transform, that depict a waveform in terms of its spectral content are notoriously complicated and power-hungry. Conversely the TESPAR solution presented in this thesis requires significantly less processing power and memory to achieve better results. In practice, almost all present

169

applications can be satisfied with S or A-Matrices - two or three dimensional matrices - making subsequent processing very much simpler, quicker and cheaper. It is a further benefit that the TESPAR matrices are of fixed size.

To demonstrate the lower processing power requirements of the TESPAR solutions a 10 word recognition system operating in real-time on a very low power 8-bit 8051-derivative micro-controller processing platform has been implements [173]. Others have demonstrated that TESPAR based signal processing solutions typically require at least two orders of magnitude less processing power than traditional Fourier based solutions [174].

An important distinction between the TESPAR and envelope solutions is that the TESPAR solution is able to identify both the condition and severity of the fault using only one frequency range of 250Hz-17000Hz. The traditional techniques solutions need to specifically target different frequency ranges and is therefore one of the reasons why the conventional methods solutions requires significantly greater processing overheads. Another important aspect of the TESPAR processing solution worth raising is that in addition to producing perfect classification performance of the fault locations and fault severities when using the full 10 seconds associated with each of the available bearing signals; the TESPAR solution also produces perfect classification performance of the fault locations and fault severities when using any random 5 second section from each of the available bearing signals. Both envelope and wavelet analysis were able to discriminate between the faults and severities based on characteristic frequencies and the level of their peaks, however, there was some difficulty with small outer race fault.

An alternative approach was used, manipulating the TESPAR codes produced from the Tescode function in order to discriminate between different operating conditions. The TESPAR solution described here uses the same stream of TESPAR symbols to identify both the fault location and determine the fault severity. Fault location is determined by comparing

170

archetype S and A-Matrices for the healthy and three given fault conditions with test data, and fault severity is obtained from the magnitude of the average ordinate over the range of symbols of interest. It is important to understand that exactly the same stream of TESPAR codes is used to identify both the fault locations and severity and thus minimal processing overheads are required.

CHAPTER SEVEN

Conclusions and Recommendation of Future Work

The chapter summarises the TESPAR finding results on bearing detection and diagnosis. Conclusions are drawn on the progress made on developing a method using the TESPAR approach to discriminate bearing fault locations and severities. Contributions to knowledge made by the researcher are presented as are suggestions for possible future work.

7.1   Summaries and Conclusions

This research has been focused on detecting and diagnosing bearing faults in three different locations: outer race, inner race and roller element based on TESPAR approaches. Vibration was chosen as the measurement parameter because it has long been used for CM of machinery, and many analytical techniques based on vibration have been introduced to detect and diagnose machine faults.

The need for an improved technique to recognize faults in bearings always remains important. The TESPAR has been in use for over thirty years, but hardly has the literature contained reports of it being successfully applied to raw time-domain data for condition monitoring of rotating machinery. The reason for this is most likely that TESPAR was originally developed as tool for speech recognition and was based on the observed fact that humans gain virtually no information from the amplitude of the sound (the same information can be conveyed in a whisper or a shout); instead the frequency content and temporal pattern is much more important.

However, the TESPAR technique came as a result of looking at the approximation of the complex zeros location alongside duration and shape of the signal waveform.

The method proposed here has little in common with envelope spectrum which has been discussed in Chapter Three alongside other conventional methods. The simulation of TESPAR for the detection of bearing defects discussed in section 4.7 in Chapter Four shows that if two signals are to be compared they must be at the same sampling rate. Also a maximum value for the Duration of 37 imposes a lower limit on the frequency analysis. Thus while analysis of raw data using TESPAR could detect the appearance of BCFs it may not be

able to discriminate between them, therefore, trained S and A-Matrices were used in this thesis for discrimination.

It is also found that TESPAR can be used to detect amplitude modulated signals where the BCF modulates a machine resonance. In addition, It is confirmed, exactly as found by previous researchers investigating amplitude modulation as a means of fault detection, it is necessary for the machine resonance to be of much greater magnitude than the background noise. Also the TESPAR handbook recommends a sampling rate of three times the Nyquist criterion [172] but this work shows that where the equipment allows for bigger rates they should be used.

Consequently, it could be expected that using TESPAR on the raw time domain bearing signal might provide a new insight into fault detection and diagnosis, and it was found that TESPAR approach applied on time-domain data successfully discriminated different bearing conditions. The correlations of healthy and faulty S-Matrices and the A-Matrices proved capable of discerning between fault locations.

In addition, the approach adopted here which appears most promising is to simply look directly at the absolute magnitude of the signal and compare healthy and faulty conditions using a fault severity index. This indicates the presence of a fault but does not diagnose the type of fault. In other words, bearing faults severities have been classified using different TESPAR features. The mean absolute magnitude value was capable of differing between the three different severities for each bearing fault tested.

The TESPAR technique was compared with very widely used bearing CM technique: the envelope spectrum analysis in an attempt to develop a powerful method that can detect faults in different bearing components. To evaluate the effectiveness and efficiency of this proposed

technique, its results were compared with those obtained using the envelope spectrum. It is recognised that while conventional techniques for rolling element bearing are in widespread use they are not always very useful and effective for fault detection.

7.2   Review of Aim and Objectives

This section reviews the objectives and achievements of the research, by comparing them one by one to the objectives for this study presented in Chapter 1. Substantial progress has been made towards meeting the aim of the research: TESPAR application has produced an efficient technique for reliable and precise detection and diagnosis of bearing faults. The main objectives and achievements of this research were:

Objective one: To present and discuss machine condition monitoring and its application to rolling element bearings.

Achievement one: The concept of condition monitoring for machinery has been defined and discussed in Chapter One in Sections 1.3 and 1.4 in terms of fault detection and diagnosis, the importance of timely maintenance to minimise overall product cost and increase process profitability. Chapter one emphasises vibration measurement as this is both the most commonly used method in industry and the technique selected for monitoring of the rolling element bearings used in this project setting the motivation, the aim and objectives also providing the structure for the research.

Objective two: To describe the fundamentals of rolling bearing transmission, including bearing types and components, their failure modes and methods of monitoring.

Achievement two: The fundamentals of roller element bearing have been dealt with comprehensively in Chapter Two in Section 2.2 which describes bearing types and

components as well as explaining modes of failure in Section 2.4. The dynamic response of roller bearings to local faults was also reviewed in Section 2.5 which shows that typical symptoms produce characteristic peaks in the vibration spectrum. The calculated characteristic frequencies for the test bearing are compared with measured results in Section 2.6, and good agreement was found.

Objective three: To review signal processing conventional methods and their parameters of bearing fault detection and diagnosis using vibration signal.

Achievement three: The time domain, frequency domain, envelope spectrum and wavelet analysis are the most common CM techniques for bearing faults detection and diagnosis, and have been introduced primarily in Chapter One in Section 1.5 and reviewed in details in Chapter Three in Section 3.3 with experimental examples and explanations.

Objective four: To describe and explore the fundamentals of the TESPAR approach and its capability as being the focus of this research.

Achievement four: The TESPAR concept has been illustrated by giving a detailed background of TESPAR origin, alphabets and TESPAR coding and its S and A matrices. How TESPAR works was the focus of Chapter Four which gave examples of TESPAR in action. A study of TESPAR simulation illustrating S and A-Matrices in details was implemented in Section 4.7 including the idea that TESPAR can be used to detect bearing faults was proposed.

Objective five: To discuss and analyse the data outputs from an accelerometer using TESPAR methods applied to the time domain data collected for ten different bearing conditions.

176

Achievement five: All data signals from the accelerometer were successfully TESPAR encoded using the code written for the MATLAB environment by the researcher. Chapter Six reported the successful extraction of TESPAR streams and codes for the vibration signals from both healthy and faulty bearings. S-Matrices and A-Matrices have been created for the healthy condition and for each fault. Archetype or typical S and A-Matrices were created for each fault, and then the measured S or A-Matrix correlated with the different fault Archetypes to determine the highest correlation coefficient. This was two tests in one, the closer the measured S or A-Matrix was to one of the fault Archetypes (the higher the correlation coefficient) the more likely it was the fault that was developing, and the higher the score the more developed the fault. Table 6-1 shows the results for the S-Matrices and Table 6.2 for the A-Matrices. The results are very promising, particularly as a method of fault identification.

This chapter has also successfully developed a fault severity classification in Section 6.4 based on the TESPAR system. All epochs that had durations greater than the average duration associated with the whole signals were identified. The code was further adapted to generate the mean absolute magnitude value from all of these longer than average epochs. The mean absolute magnitude value produced is then used as a direct indicator of fault severity as it has been shown in Figures 6-20 to 6-22. This measure of fault severity appears directly related to the severity of the seeded fault but did depend on the type of fault. The results suggest a very promising method of fault detection.

The results obtained indicate that TESPAR is a suitable method for detection and diagnosis if faults in rolling element bearings.

Objective six: To perform a relative evaluation of the TESPAR application and the envelope spectrum as the most commonly used method for bearing detection in order to ascertain the TESPAR efficiency and effectiveness for bearing fault detection and diagnosis in the given system.

Achievement six: A detailed systematic and theoretical study with justifications has been carried out with different examples of comparing the performance of the S and A Matrices of the TESPAR approach against the standard bearing fault detection method, in particular the envelope spectrum in Section 6.5.

7.3   Contributions to Knowledge

First contribution: A new technique has been produced that can detect faults at three different bearing locations with three levels of severity. This research will provide future researchers with a basis of building on TESPAR results for monitoring bearing conditions.

Second contribution: A variety of TESPAR features have been employed and tested on different bearing vibration signals in order to select the most suitable and promising for each kind of signal waveform.

Third contribution: Both S matrices and A matrices were constructed and used in this thesis and which eventually reach the same conclusion for bearing detection.

Fourth contribution: Statistical classification has been implemented on TESPAR results for both S matrices and A matrices to categorise different bearing waveforms.

Fifth contribution: A complete Matlab code has been written to implement the TESPAR coding process and produce different TESPAR features for different bearing vibration signal waveforms for bearing condition classification in both terms the location and severity.

Sixth contribution: The author has suggested a new method for rolling element bearing fault detection. An underpinning technique has been developed and successfully tested.

Seventh contribution: The TESPAR solution proposed uses significantly less processing time and computer memory than traditional Fourier based techniques making TESPAR a more cost efficient solution, and one that requires less energy to operate and thus is more environmentally friendly.

Eighth contribution: Highlighting the use of non-traditional signal processing techniques could encourage others to investigate the use of TESPAR when investigating other condition monitoring problems.

7.4   Suggestions for Further Work

Based on this research a number of suggestions are made for further researches in this field:

1- The S matrices and A matrices normally consists of a total of 28 symbols in all, and a small percentage of observations occur for symbols whose value is > 28. Future work could examine whether including these higher values might improve fault classifications when using other TESPAR features and specifications.

2- TESPAR features and specifications could be used to test other different kinds of bearings types and design as well as detect and classify other rolling element bearing defects such as misalignment, unbalance, looseness, and lack of lubrication, and also to investigate and explore bearing signal with other types of signal waveforms such as acoustic emission, etc.

3- Future researchers should develop an experimental rig that can cope with harsh conditions such as high torsion and radial loads as well as high speeds, as this work shows that where the equipment allows for faster rates they should be used.

4- In addition to using correlation as a technique to detect and classify a bearing fault, future research could focus on applying other methods such as an artificial neural network for bearing faults classification using both TESPAR S-matrices and A matrices.

5- Modify and improve the TESPAR coding table (i.e. the combination of epoch duration and shape information that defines each TESPAR code) to determine whether further performance improvements can be achieved by including fewer frequency range codes.

6- Investigate how TESPAR can be used to assist in the determination of time to failure of equipment components.

7- Determine whether any single TESPAR technique can be used to determine both fault condition and severity.

8- Combine TESPAR with other signal processing techniques to determine whether a combined solution could provide the best of both techniques.

9- Test TESPAR and other traditional techniques on real-world bearing rigs to identify whether a solution works better in an operational scenario.

10-   Possible improvements in the method could be to either to (i) simply determine the severity index for all codes, or to (ii) simply determine the severity index for all codes and subtract the value for the normal archetype.

11-   TESPAR ought to generate interest leading to the development of a wireless-based condition monitoring solution.

Appendix

Photograph of the simulated bearing faults used in this research

Photograph of NSK type N406 cylindrical roller bearing with of 100% roller fault (R_1)

Photograph of NSK type N406 cylindrical roller bearing with of 60% roller fault (R_2)

Photograph of NSK type N406 cylindrical roller bearing with 30% roller fault (R$_3$)

Photograph of NSK type N406 cylindrical roller bearing with of 100% inner race fault (I$_1$)

Photograph of NSK type N406 cylindrical roller bearing with of 60% inner race fault (I$_2$)

Photograph of NSK type N406 cylindrical roller bearing with of 30% inner race fault (I$_3$)

Photograph of NSK type N406 cylindrical roller bearing with of 100% outer race fault (O1)

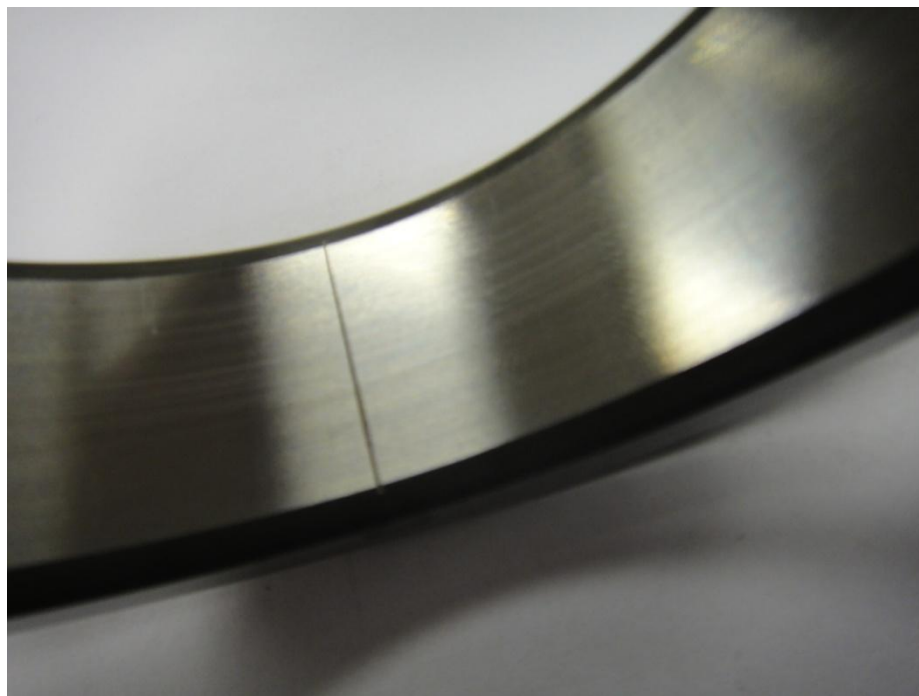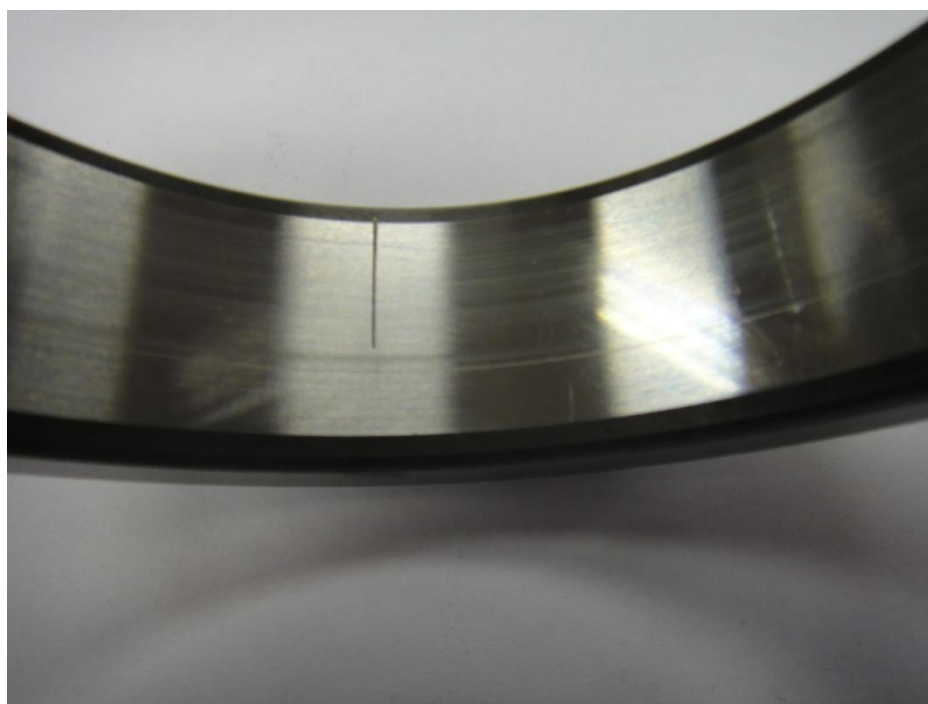**Photograph of NSK type N406 cylindrical roller bearing with of 60% outer race fault (O2)**
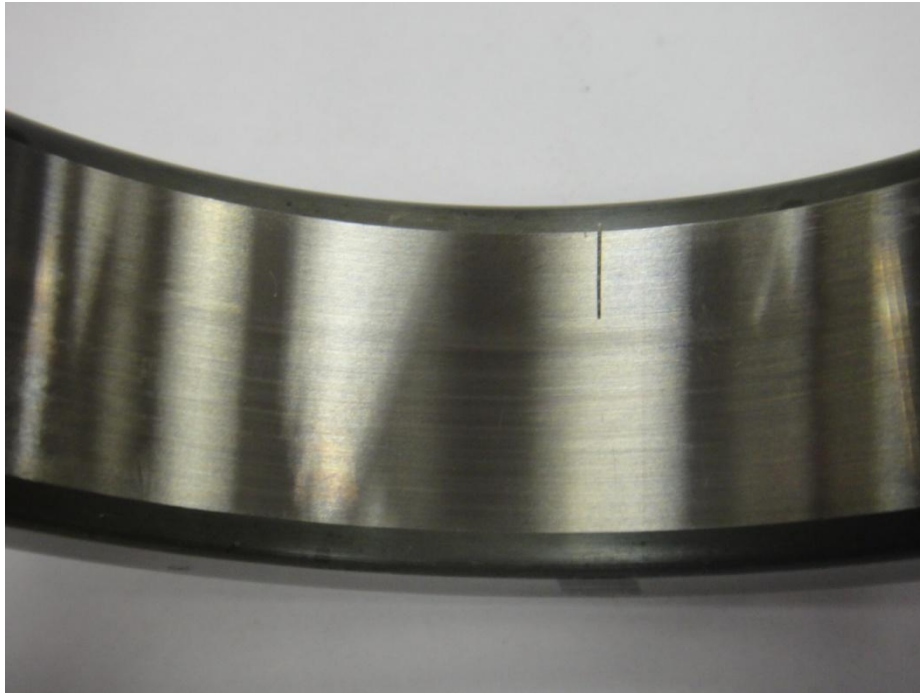
Photograph of NSK type N406 cylindrical roller bearing with of 30% outer race fault (O$_3$)

Bibliography

[1] Rao, B.K.N. (1996) *Handbook of condition monitoring*. Oxford: Elsevier.

[2] Hunt, T.M., ed. (1998) *Concise encyclopaedia of condition monitoring*. Coxmoor Publishing Company.

[3] Starr, A. & Rao, B.K.N. (2001) Condition monitoring and diagnostic engineering management. *Proc. 14th Int. Congress COMADEM*. Elsevier.

[4] Various Authors (1987 – 2011) *Proceedings of COMADEM International Conferences.*

[5] Al Thobiani, F. (2011) The non-intrusive detection of incipient cavitation in centrifugal pumps. Unpublished PhD Thesis, *Huddersfield University*.

[6] Kelly, A. (1997) *Maintenance strategy*. Oxford: Butterworth-Heinemann.

[7] Randall, R. B. (2010) *Vibration based condition monitoring*. Chichester: John Wiley & Sons, Ltd.

[8] Randall, R. B. (2004a) State of the art in monitoring rotating machinery Part 1. *Sound and Vibration*, (March), pp. 14-20.

[9] Randall, R. B. (2004b) State of the art in monitoring rotating machinery Part 2. *Sound and Vibration* (May), pp. 10-16.

[10] Patton, R. J. & Chen, J. (eds) (1998) *Proceedings of IFAC Symposium on fault Detection*, Supervision and Safety for Technical Processes-SAFE-PROSS'97, Pergamon.

[11] Latino, J. (2009) *The power of failure analysis to eliminate process interruptions*. Reliability Center Inc., Hopewell, Virginia.

[12]   Moore, R., Pardue, F., Pride, A. & Wilson, J. (1993) *The Reliability-based maintenance strategy: a vision for improving industrial productivity* (CSI Industry Report No. AN-GM-00I-09.09.93).

[13]   Yung, C. & Bonnett, A. (2004) Repair or replace? *IEEE Industry Applications Magazine*, 10 (5), pp. 48-58.

[14]   Huang, H. B. &Wang, H. P. (1996) An integrated monitoring and diagnostic system for roller bearings. *Int. Journal of Advanced Manufacturing Technology*, 12(1), pp. 37-46.

[15]   Bardou, O. (1994) Early detection of leakages in the exhaust and discharge systems of reciprocating machines by vibration analysis. *Mechanical Systems and Signal Processing*, 8(5), pp. 551-570.

[16]   Ball. A. D. & Gu, F. (1995) A basis for the vibration monitoring of diesel fuel injectors. *Journal of Maintenance*, 10(2), pp. 24- 30.

[17]   Moore, R., Pardue, F., Pride, A. & Wilson, J. (1993) The Reliability-based maintenance strategy: a vision for improving industrial productivity. *(CSI Industry Report No.* AN-GM-00I-09.09.93).

[18]   Botton, G., Ben-Ari, J. & Sher, E. (2000) Vibration monitoring as a predictive maintenance tool for reciprocating engines. *Journal of Automobile Engineering*, 214(8), pp. 895-903.

[19]   McFadden, P. & Smith, J. (1985) A signals processing technique for detecting local defects in a gear from the signal average of the vibration. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 199(4), pp. 287-292.

[20]   Latino, C. J. (1999) Hidden treasure: eliminating chronic failures can cut maintenance costs up to 60%. Report, *Reliability Center, Inc*., Hopewell, Virginia.

[21]   Bently, D.E. (1982) Monitoring rolling element bearings, *Orbit*. (November), pp.2-3.

[22]   Stack, J.R., Habetler, T.G. & Harley, R.G. (2004) Fault Classification and Fault Signature Production for Rolling Element Bearing in Electric Machines. *Proc. 4th IEEE Symp. on Diagnostics for Electric Machines, Power Electronics and Drives*, pp 172-176.

[23]   Shreve, D. H. (1995) Vibration. Analysis: Fundamentals. *P/PM Technology*. pp. 36-40.

[24]   FAG (1999) Rolling Bearing Fundamentals. *Technical Information* TI No. WL 43-1190 EA. D-97419 Schweinfurt Germany.

[25]   Springer, T. E. (2000) Bearing knowledge is power. *Spexx Bearings, Timken Company, 1983 North Hunt Street, Terre Haute*, IN 47805, USA.

[26]   Calvin, J. (1995) Maintenance and troubleshooting are keys to longer bearing life. *Plant Engineer*, October 9.

[27]   Phillips, M. and Kilbey, B. (1999) Bearing Defect Detection at Very Low Frequencies [ONLINE].   Available: httpy/www.compsys.com/DRKNOW/APLPapr.nsf/84clb852566b9004G354?OpenDocument.

[28]   Widner, R. L. & Littmann, W. E. (1976) Bearing Damage Analysis. *National Bureau of Standards Special Publication*, No. 423.

[29]   Norton, M. & Karczub, D. (2003) *Fundamentals of noise and vibration analysis for engineers*. UK: CUP.

[30]   Norton, M. & Karczub, D. (2003) *Fundamentals of noise and vibration analysis for engineers*. UK: CUP.

[31]   Martin, H. R. & Honarvar, F. (1995) Application of Statistical Moments to Bearing Failure Detection. *Applied Acoustics*, Vol. 44, pp. 67-77.

[32]   Heng, R. B. W. & Nor, M. J. M. (1998) Statistical Analysis of Sound and Vibration Signals for Monitoring Rolling Element Bearing Condition. *Applied Acoustics*, Vol. 53, No. 1-3, pp. 211-226.

[33]   Su, Y. T. & Lin, S. J. (1992) On Initial Fault Detection of a Tapered Roller Bearing: Frequency Domain Analysis. *Journal of Sound and Vibration*, 155(1), pp. 75-84.

[34]   McFadden, P. D. & Smith, J. D. (1984) Vibration Monitoring of Rolling Element Bearings by High Frequency Resonance Technique - A Review. *Tribology International*, Vol. 77, pp. 3-10.

[35]   Howard, I. (1994) A Review of Rolling Element Bearing Vibration: Detection, Diagnosis and Prognosis. *Aeronautical and Maritime Research Laboratory Airframes and Engines Division.*

[36]   Li, C. J. & Ma, J. (1997) Wavelet Decomposition of Vibrations for Detection of Bearing-Localized Defects. *NDT & E International*, Vol. 30, No. 3, pp. 143-149.

[37]   Lee J.H., Kim J. & Kim H. J. (2001) Development of enhanced Wigner-Ville distribution function. *Mechanical Systems and Signal Processing*, 15(2), pp.367-398.

[38]   Yen, G. Y. & Lin, K. C. (1999) Wavelet Packet Feature Extraction for Vibration Monitoring. *Proceedings of the IEEE International Conference on Control Applications*, pp. 1573-1578.

[39]   Li, C. J., Ma, J. & Hwang, B. (1995) Bearing Localized Defect Detection by Bicoherence Analysis of Vibrations. *Journal of Engineering for Industry*, Vol. 117, pp. 625-629.

[40]   McCormick, A. C. & Nandi, A. K. (1999) Bispectral and Trispectral Features for Machine Condition Diagnosis. *IEEE Proc.-Vis. Image Signal Process*, Vol. 146, No. 5, pp. 229-234.

[41]   Schwabacher, M. & Goebel, K. (2007) A Survey of Artificial Intelligence for Prognostics. Working Notes of 2007 *AAAI Fall Symposium: AI for Prognostics*.

[42]   Subrahmanyam, M. & Sujatha, C. (1997) Using Neural Networks for the Diagnosis of Localized Defects in Ball Bearings. *Tribology International*, Vol. 30, No. 10, pp. 739-752.

[43]   Li, B., Chow, M. Y., Tipsuwan, Y. & Hung, J. C. (2000) Neural-network based Motor Rolling Bearing Fault Diagnosis. *IEEE Transactions on Industrial Electronics*, Vol. 47, No. 5, pp. 1060-1069.

[44]   Godi, G., Li, B., Chow, M. Y. & Hung, J. C. (1998) Motor Bearing Diagnosis by a Fundamental Frequency Amplitude Based Fuzzy Decision System. *Proceedings of the 24l Annual Conference of the IEEE Industrial Electronics Society*, Vol. 4, pp. 1961-1965.

[45]   Zacksenhouse, M., Braun, S. & Feldman, M. (2000) Toward Helicopter Gearbox Diagnostics from a Small Number of Examples. *Mechanical systems and Signal Processing,* Vol. 14(4), pp. 523-43.

[46]   Gunn, R. (1998) Support vector machines for classification and regression, *technical report, University of Southampton.*

[47]   Goudong, G., Li, S. & Chan, L. (2001) Support vector machines for face recognition. *Image and Vision Computing*, Vol. 19, pp. 631-8.

[48]   Barzilay, O. & Brailovsky, L. (1999) On domain knowledge and feature selection using a support vector machine. *Pattern Recognition Letters*, Vol. 20(5), PP. 475-84.

[49]   Ge, M., Du, R., Zhang, C. & Xu, S. (2004) Fault diagnosis using a support vector machine with application in sheet metal stamping operations. *Mech. Systems and Signal Processing*, Vol. 18 pp. 143-159.

[50] Hu, Q., He, Z., Zang, Z. & Zi, Y. (2007) Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble. *Mech. Systems and Signal Processing,* Vol. 21, pp. 688-705.

[51] Yang, Y., Yu, D. & Cheng, J. (2007) A fault diagnosis approach for roller bearing based on envelop spectrum and SVM. *Measurement*, Vol. 40, pp. 943-50.

[52] Zadeh, A. (1965) Fuzzy sets. *Information and Control,* Vol. 8, pp. 338–353.

[53] Kosko, B. (1997) Fuzzy engineering. *Prentice-Hall: Upper Saddle River.*

[54] Jang, S., Sun, T., Mizutani, E. (1997) Neuro-Fuzzy and Soft Computing. *Prentice Hall.*

[55] Takagi, T., Sugeno, M. (1985) Fuzzy identification of systems and its applications to modelling and control. *IEEE Transactions on Systems*, Man, and Cybernetic SMC-15, pp. 116–132.

[56] Sugeno, M., Kang, T. (1988) Structure identification of fuzzy model. *Fuzzy Sets and Systems,* Vol. 28(1), pp. 15–33.

[57] McFadden, P. D., Smith, J. D. (1984) Model for the Vibration Produced by a Single Point Defect in a Rolling Element Bearing. *Journal of Sound and Vibration*, 96(1), pp. 69-82.

[58] Wang, Y. F. & Kootsookos, P. J. (1998) Modeling of Low Shaft Speed Bearing Faults for Condition Monitoring. *Mechanical Systems and Signal Processing*, 12(3), pp. 415-426.

[59] Afshari, N. & Loparo, K. A. (1998) A Model Based Technique for the Fault Detection of Rolling Element Bearings Using Filter Design and Sliding Mode Techniques. *Proceedings of the 37' IEEE Conference on Decision & Control*, pp. 2593-2598.

[60] Lundberg, G. & Palmgren, A. (1952) Dynamic Capacity of Rolling Bearing. *Ada Polytechnica, Mechanical Engineering Series* 2, No. 4, 52.

[61] Khan, A. F. (1991) Condition Monitoring of Rolling Element Bearings: A Comparative Study of Vibration based Techniques. Ph.D. *Thesis, University of Nottingham.*

[62] Li, Y., Billington, S., Zhang, C., Kurfess, T., Danyluk, S. & Liang, S. (1999) Adaptive Prognostics for Rolling Element Bearing Condition. *Mechanical Systems and Signal Processing*, 13(1), pp. 103-113.

[63] Mba, D., & Rao, R. (2006) Development of Acoustic Emission Technology for Condition Monitoring and Diagnosis of Rotating Machines; Bearings, Pumps, Gearboxes, Engines and Rotating Structures. *The Shock and Vibration Digest* 38(1) pp 3-16).

[64] Dornfeld, D. A. (1984) Acoustic emission in monitoring and analysis and analysis in manufacturing. *Proc. AE Moint. Nal. Manuf.*, 4, pp. 24.

[65] Huang, M., Jiang, L., Liaw, P. K., Brooks, C. R., Seeley, R. & Klarstrom D. L. (1998) Using Acoustic Emission in Fatigue and Fracture Material Research. *Journal of the Minerals, Metals and Materials Society*, 11, pp. 1047-4838.

[66] Toenshoff, H. K., Jun, M., MANNEL, S. & Rietz, W. (2000) Using acoustic emission signal for monitoring of production processes. *Ultrasonics*, 37, pp. 671-676.

[67] Inasaki, I. (1998) Application of acoustic emission sensor for monitoring machining processes. *Ultrasonics*, 36, pp. 273-281.

[68] Yoshioka, T. & Fuhjiwara, T. (1984) Application of acoustic emission technique to detection of rolling element-bearing failure. *ASME D.A. Dornfield(Ed), Acoustic Emission monitoring and analysis in Manufacturing,* New York, pp.54-75.

[69] Yoshioka, T. (1992) Detection of rolling contact subsurface fatigue cracks using acoustic emission technique. *Lubrication Engineering*, 49 (4), pp. 303-308.

[70]   Tandon, N. & Naka, B.C. (1990) Defect detection in rolling element bearing by acoustic emission method. *Journal of Acoustic Emission*, 9 (1), pp. 25-28.

[71]   Tan, C. C. (1990) Application of acoustic emission to the detection of bearing failures*, Proceeding Tribology Conference*, Brisbane, pp. 110-114.

[72]   Yan, A., El-Wardany, T. I. & Elbestawi, M.A. (1995) A Multi-ssensor Strategy for Tool Failure Detection in Milling. *Int. J. Mach. Tools Manufact*., 35(3), pp. 383-398.

[73]   King, R.A. & Phipps, T.C. (2003) Shannon, TESPAR and Approximation Strategies. *Hydralogica*. Paper 517, 4th December 2003. 17pp.

[74]   Licklidder, J.C.R. & Pollack, I. (1948) Effects of differentiation, integration and infinite peak clipping upon the intelligibility of speech. *JASA*, 20(1), pp. 42-51.

[75]   Berry, J.E. (1991) How to track rolling element bearing health with vibration signature analysis. *Sound and Vibration*, 25(11), pp.24-35.

[76]   Harris, T. & Kotzalas, M. (2007) Rolling bearing analysis 5th edition*, Essential Concepts of Bearing Technology*, Taylor & Francis.

[77]   Guo, H., Jack, L.B. & Nandi, A.K. (2005) Feature generation using genetic programming with application to fault classification. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 35(1), pp.89-99.

[78]   Mathew, J. & Alfredson, R.J. (1984) The condition monitoring of rolling element bearings using vibration analysis. *Transactions of the ASME, Journal of Vibration, Acoustics, Stress, and Reliability in Design*. 106, pp.447-453.

[79]   Morando, L.E. (1998) Bearing condition monitoring by the shock pulse method. *Iron and Steel Engineer*, 75(12), pp.40-43.

[80]   Burgess, P.J.F. (1988) Antifriction bearing fault detection using envelope detection. *Transactions of the Institute of Professional Engineers New Zealand - Electrical/Mechanical Chemical Engineering Section*, 15(2), pp.77-82.

[81]   Li, Y., Kurfess, T.R. & Liang, S.Y. (2000) Stochastic prognostics for rolling element bearings. *Mechanical Systems and Signal Processing*, 14(5), pp.747-762.

[82]   Brie, D. (2000) Modelling of the spalled rolling element bearing vibration signal: an overview and some new results. *Mechanical Systems and Signal Processing*, 14(3), pp.353-369.

[83]   Bently, D.E. (1999) Acceleration enveloping, *Orbit*, First Quarter, Nevada.

[84]   Randall, R.B. & Gao, Y. (1996) Masking effect in digital envelope analysis of faulty bearing signals, IMechE, pp.351-359.

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

[86]   Shiroishi, J., et al. (1997) Bearing condition diagnostics via vibration and acoustic emission measurements. *Mechanical Systems and Signal Processing*, 11(5), pp.693-705.

[87]   Williams, T., Ribadeneira, X., Billington, S. & Kurfess, T. (2001) Rolling element bearing diagnostics in run-to-failure lifetime testing. *Mechanical Systems and Signal Processing*, 15(5), pp.979-993.

[88]   Tao Meng, Ming-Fu Liao & Hui Li. (2003) Detection and diagnosis of the gear fault by the delayed correlation-envelope technique, Hangkong Dongli Xuebao. *J. Aerospace Power*, 18(1), pp. 109-113.

[89]   White, G. (1991) Amplitude demodulation - a new tool for predictive maintenance. *Sound and Vibration*, pp. 14-19.

[90]   Lee, S.K. & White, P.R. (1998) The enhancement of impulsive noise and vibration signals for fault detection in rotating and reciprocating machinery. *Sound and Vibration*, 217(3), pp.485-505.

[91]   Eren, L. & Devaney, M.J. (2003) Motor bearing damage detection via wavelet analysis of the starting current transient. *Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference*, Vol 3, pp.1797-1800.

[92]   Altmann , J. & Mathew, J. (2001) Multiple band-pass autoregressive demodulation for rolling-element bearing fault diagnosis. *Mechanical Systems and Signal Processing*, 15(5), pp.963-977.

[93]   Xia Limin. (2002) Wavelet packets analysis of rolling bearing vibration signal and fault testing, *Proc. 4th World Congress on Intelligent Control and Automation,* June 10-14, pp. 3005-3008.

[94]   Eren, L. & Devaney, M.J. (2004) Bearing damage detection via wavelet packet decomposition of the stator current. *IEEE Transactions on Instrumentation and Measurement*, 53(2), pp.431-436.

[95]   Zhang, H., Wang, S., Zhang, Q. & Zhai, G. (2003) The research on rolling element bearing fault diagnosis based on wavelet packets transform, *IEEE*, pp. 1745-1749.

[96]   Shao, Y. & Nezu, K. (2004) Extracting symptom of bearing faults in the wavelet domain, *Proc. Inst Mech. Engrs*, (218), Part I: J. Systems and Control Engineering, pp. 39-51.

[97]   Li, C., Song, Z. & Li, P. (2004) Bearing fault detection via wavelet packet transform and rough set theory. *Proc. 5th World Congress on Intelligent Control and Automation*, June 15-19, pp. 1663-1666.

[98]   Hindhede, U. (1983) *Machine design fundamentals: a practical approach*. New Jersey: Prentice-Hall.

[99]   Dynaroll Technical Literature www.dynaroll.com. Accessed Jan 2011.

195

[100] JTEKT Corporation (1992) Koyo. *Ball and bearings: failures, causes and countermeasures*. CAT No B3001E. Available on http://www.koyousa.com/brochures/pdfs/catb3001e.pdf.

[101] FAG Bearings Corporation (1998) *Rolling bearing damage: Recognition of damage and bearing inspection*. Publ. No. WL 82 102/2 ED.

[102] Miller, J. & Kitaljevich, D. (2000) In-line Oil Debris Monitor for Aircraft Engine Condition Assessment, *Standard IEEE copyright* 0-7803-5846-5/00.

[103] Goddard, K. & MacIsaac, B. (1995) Use of oil borne debris as a failure criterion for rolling element bearings. *Lubrication Engineering*, 51(6), pp. 481–487.

[104] ISO 281:2007, Rolling bearings - Dynamic load ratings and rating life. *ISO*.

[105] Williams, J. & Hyncica, A. (1992) Mechanisms of abrasive wear in lubricated contacts, *Wear* Vol 151 pp57-74.

[106] SKF (1994) *Bearing failures and their causes*. Product information 401.

[107] Nilsson, R., Dwyer-Joyce, T. & Olofsson (2006) Abrasive Wear of Rolling Bearings by Lubricant Borne Particles, *Proc. IMech E, Part J: Journal of Engineering Tribology*, pp 429-439.

[108] Tandon, N. & Choudhury, A. (1997) An analytical model for the prediction of the vibration response of ball bearings due to a localized defect, *Sound and Vibration*, Vol 205, No 3, pp. 275-292.

[109] Choudhury, A. & Tandon, N. (2006) Vibration response of ball bearings in a rotor bearing system to a local defect under radial load, *Tribology*, Vol 128, No 2, pp. 252-261.

[110] Emerson bearing, Bearing failure analysis, http://www.emersonbearing.com/technical-toolbox.html.

[111] Liu, T. I. & Iyer, N. R. (1992) On-line recognition of roller bearing states. *Proceedings of the 1992 Japan - USA Symposium on Flexible Automation,* Part 1 (of 2), Jul 13-15, San Francisco, CA, USA, pp. 257-262.

[112] Lindh, T. (2003) *On predictive bearing condition monitoring of induction motors.* PhD. Dissertation, Lappeenranta University of Technology. ISBN 951-764-841-3.

[113] International Standards Office (2002) ISO 13373-*1Condition monitoring and diagnostics of machines - Vibration condition monitoring*. London: HMSO.

[114] Ghafari, S. H. (2007) A Fault Diagnosis System for Rotary Machinery Supported by Rolling Element Bearings, *PhD Thesis University of Waterloo*, Canada.

[115] Ho, D. & Randall R.B. (2000) Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals. *Mechanical Systems and Signal Processing*, 14(5), pp. 763-788.

[116] McFadden, P.D. & Toozhy, M.M. (2000) Application of synchronous averaging to vibration monitoring of rolling element bearings, *Mechanical Systems and Signal Processing,* Vol 14, No 6, pp. 891-906.

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

[118] Batchelor, B.G. (1978) *Pattern Recognition*. New York: Plenum Press.

[119] Unal, A. (1994) Intelligent Diagnostics of Ball Bearings. *The Shock and Vibration Digest*, Nov./Dec., pp. 9-12.

[120] International Standards Office (2005) ISO 13373-2 *Condition monitoring and diagnostics of machines -- Vibration condition monitoring -- Part 2: Processing, analysis and presentation of vibration data*. London: HMSO.

[121] Sinocera (2009) Technical specification sheets, Shanghai 201103 China (also available on the internet).

[122] Norton, M. & Karczub, D. (2003) *Fundamentals of noise and vibration analysis for engineers*. UK: CUP.

[123] McFadden, P.D. (1990) Condition monitoring of rolling element bearings by vibration analysis. *Proc. Institution of Mechanical Engineers,* London pp 49-54.

[124] Sun, Q., Xi, F., Chen, P. & Krishnappa, G. (1999) Bearing condition monitoring through pattern recognition analysis. *6th International Conference on Sound and Vibration*, Denmark.

[125] Sun, Q., Xi, F. & Krishnappa, G. (1998) Signature analysis of rolling element bearing defects. *Proceedings of CSME Forum*, Toronto, pp. 423-429.

[126] Gustafsson, O. & Tallian, T. (1962) Detection of damage in assembled rolling bearings, *Transaction of the American Society of Lubrication Engineers*, Vol. 5, pp. 197-209.

[127] Al Kazzaz, S.A.S. & Singh, G.K. (2003) Experimental investigations on induction machine condition monitoring and fault diagnosis using digital signal processing techniques. *Electric Power Systems Research* 65 pp. 197-221.

[128] Alfredson, R.J. & Mathew, J. (1985) Time-domain methods for monitoring the condition of rolling element bearings. *Mechanical Engineering Transactions*, ME 10(2), pp.l02-l 17.

[129] Howard, I.M. (1994) *A review of rolling element bearing vibration - detection, diagnosis and prognosis*. Melbourne, Australia: Defence Science and Technology Organization.

[130] Dyer, D. & Stewart, R. (1978) Detection of rolling element bearing damage by statistical vibration analysis. *Journal of Mechanical Design*, 100(2), pp.229-235.

[131] Alguindigue, I.E., Loskiewicz-Buczak, A. & Uhrig, R. (1993) Monitoring and diagnosis of rolling element bearings using artificial neural networks. *IEEE Transactions on Industrial Electronics,* 40(2), pp. 209-217.

[132] Li, C. & Pickering, C. (1992) Robustness and sensitivity of non-dimensional amplitude parameters for diagnosis of fatigue spalling. *Condition Monitoring and Diagnostic Technology*, 2(3), pp.81-84.

[133] Konstantin-Hansen, H. (2003) *Envelope analysis for diagnostics of local faults in rolling element bearings*. B&Q Application Note BO 0501 02/03. Available on http://www.bksv.com/doc/bo0501.pdf

[134] Lyons, R.G. (2009) *Understanding digital signal processing*. Pearson Education.

[135] Halim, E.B., Shah, S.L., Choudhury, S. & Kadali, R. (2008) Application of bicoherence analysis on vibration data for condition based monitoring of rotating machinery. *Proceedings of the 17th World Congress of the International Federation of Automatic Control*. Seoul, Korea, July 6-11, 2008.

[136] Sawahli, N. & Randall, R. (2007) Simulation of vibrations produced by localized faults in rolling elements of bearings in gearboxes *5th Australasian Congress on Applied Mechanics ACAM* 10-12 December Brisbane Australia.

[137] Konstantin-Hansen, H. & Herlufsen, H. (2010) B&K SOUND & VIBRATION, *Brüel and Kjær*, May, Denmark.

[138] Tyagi, S. (2007) Wavelet Analysis And Envelope Detection For Rolling Element Bearing Fault Diagnosis, A Comparative Study, *Centre of Marine Engineering Technology*, Lonavla, India 410 402.

[139] Spatenka, P., Lindh, T., Ahola, J. & Partanen, J. (2003) Embedded DSP based system for bearing condition monitoring of electric motors, *Research report 17, Department of Electrical Engineering Lappeenranta University of Technology*, Finland.

[140] Courrech, J. (2000) Envelop Analysis for Effective Rolling Element Fault Detection-Facts or Fiction? *Up Time Magazine*.

[141] Fidler, P. (2007) Theoretical verification of envelope analysis of rolling element bearings using Matlab, *PhD Thesis*, *Fakulta elektrotechniky a komunikačních technologií* VUT v Brně, Cz.

[142] Friedman, A. (1997) Demodulation an essential tool for vibration analysis, *Shock and Vibration Digest* Vol 32 No 1 pp 49-50.

[143] Toersen, H. (1998) Application of an envelope technique in the detection of ball bearing defects in a laboratory experiment TriboTest Vol 4 No 3 pp 297-308.

[144] Mendel,E., Rauber, T., Varejao,F., & Batista, J. (2009) Rolling element bearing fault diagnosis in rotating machines of oil rig extraction rigs, *17th European Signal Processing Conference* (EUSIPCO 2009).

[145] Proakis, J.G. & Salehi, M. (2000) Contemporary Communication Systems, *BookWare Companion Series*.

[146] Randall, R. B. & Luo, D. (1990) Hilbert Transform Techniques for Torsional Vibration Analysis, *Australian Vibration and Noise Conference*, Melbourne, Australia.

[147] Gabor, D. (1946) Theory of Communication. *Proc IEE 93*, pp 429-457.

[148] Lee J.H., Kim J. & Kim H. J. (2001) Development of enhanced Wigner-Ville distribution function. *Mechanical Systems and Signal Processing*, 15(2), pp.367-398.

[149] Kaiser, G. (1994) *A Friendly guide to wavelets*. Boston: Birkhauser.

[150] Daubechies, I. (1992) *Ten lectures on wavelets*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

[151] Vetterli, M. & Herley, C. (1993) Wavelets and recursive filter banks. *IEEE Transactions on Signal Processing*. IT-41, pp. 2536-2556.

[152] Mori, K., Kasashima, N., Yoshioka, T. & Ueno, Y. (1996) Prediction of spalling on a ball bearing by applying the discrete wavelet transform to vibration signals. *Wear*, 195(1-2), pp. 162-168.

[153] Ueno, Y., Mori, K. & Kasashima, N. (1996) An on-line diagnostic method to predict the spalling on a ball bearing. *Symposium on Flexible Automation*, ASME, Vol. 1, pp. 449-452.

[154] Staszewski, W., Ruotolo, R. & Storer, D. (1999) Fault detection in ball bearings using wavelet variance. *IMAC Proceedings of the 1999 17th Int. Modal Analysis Conference*, 2, pp. 1335-1339.

[155] Gao, R. & Yan, R. (2011) Wavelets: Theory and Applications for Manufacturing, *Springer Science and Business Media LLC,* London*,* England.

[156] George, M. & King, R. (1995) A Robust Speaker Verification Biometric, *Proc IEEE 29th Int. Conf. on Security Technology*.

[157] King, R.A.& Phipps, T.C. (1999) Shannon, TESPAR and Approximation Strategies. *ICSPAT 98*, 18, pp. 445-453.

[158] Freebody, N. & Watton. J. (1999) A time-encoded signal-processing approach to fault classification of an electro-hydraulic pressure control system and its application to a hot steel strip rolling mill. *Proc. Inst. Mech. Engrs.* 213(1), pp. 407-426.

[159] Bond, F.E. & Cahn, C.R. (1958) A relationship between zero crossings And Fourier coefficients for bandwidth-limited functions. *IRE Trans. Information Theory*, IT-4, pp. 110-113.

[160] Wall, R. (2003) Simple methods for detecting zero crossing. *IEEE 29th Conference on Industrial Electronics,* Vol 3 pp2477-2481.

[161] Yang, S., Er, M. & Gao, Y. (2001) A high performance neural-network-based speech recognition system. *Proceedings of the International Joint Conference on Neural Networks*, Vol 2, p.1527.

[162] Rabiner, L. & Juang, B. (1993) Fundamentals of speech recognition. *Englewood Cliffs*, N.J.: PTR Prentice Hall.

[163] Junqua, J. C. & Haton, J. P. (1996) Robustness in automatic speech recognition: fundamentals and applications. *Kluwer Academic Publishers,* Boston.

[164] Satya Prasad, K., Anitha Sheela, K. & Sridevi, M. (2007) Optimization of TESPAR Features using Robust F-Ratio for Speaker Recognition. *Int. Conf. on Signal Processing*, *Journal of Multimedia*, Vol. 2, No. 6. Pp34-43.

[165] Sheela, K. & Prasad, K. (2007) Linear Discriminant Analysis F-Ratio for Optimization of TESPAR & MFCC Features for Speaker Recognition, *Journal of Multimedia*, Vol. 2, No. 6. Pp34-43.

[166] Adar, P. & Odhiambo, M. (2009) TESPAR coded speech quality evaluation. *IEEE Africon.*

[167] Geoge, M. H. (2007) TESPAR Paves the Way to Smart Sensor. *Sensor Review*, 17(2), pp.131-137.

[168] King, R.A. & colleagues (1999) TESPAR TOOLBOX 2.0. *Domain Dynamics Limited*.

[169] Lupu, E., Pop, P. & Pătraş, M. (2004) Low Complexity Speaker Recognition System Developed on the DSP TMS320C541 Board. *9th Conference Speech and Computers*, St. Petersburg, Russia September 20-22.

[170] King, R.A. & Gosling, W. (1978) Time-encoded speech. *Electronics Letters*, 14, pp.456-457.

[171] Scheffer, C. & Girdhar, P. (2004) Practical Machinery Vibration Analysis and Predictive Maintenance, *Elsevier*.

[172] TESPAR Toolbox 2.0.5 for use with Matlab (2012) *TESPAR DSP Ltd*, p2-20.

[173] Phipps, T. C. & King, R. A. (2003) A Low-Power, Low-Complexity, Low-Cost TESPAR-Based Architecture For The Real-Time Classification Of Speech And Other Band-Limited Signals, *Domain Dynamics Limited.*

[174] Phipps, T. C. & King, R. A. (1998) A Speaker Verification Biometric In 40 Bytes, *Domain Dynamics Limited.*

[175] Wang, X.F., Shi, X.Z. & Xu, M. (1998) The fault diagnosis and quality evaluation of ball bearing by vibration signal processing. *Proc. 1st Int. Machinery Monitoring and Diagnostics Conference*, Nov, pp 318-321.

[176] Mechefske, C.K. & Liu, L. (2001) Detecting and diagnosing faults in variable speed machines. In: A. Star & B. Rao, eds. *Condition Monitoring and Diagnostic Engineering Management*. Oxford Elsevier Science Ltd., pp.709-716.