2014-01-27

Automatic Fault Diagnosis for Rolling Element Bearings

Xu, Peng

UNIVERSITY OF CALGARY

Automatic Fault Diagnosis for Rolling Element Bearings

by

Peng Xu

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MECHANICAL AND MANUFACTURING ENGINEERING

CALGARY, ALBERTA

JANUARY, 2014

Abstract

Vibration-based condition monitoring plays an important role in rolling element bearing maintenance. Based on features in bearing vibration signals, envelope analysis is very popular because of its effectiveness in bearing fault diagnostics. However, its effectiveness heavily relies on selection of the frequency band which has been accomplished manually. In this research, we develop an automated signal analysis procedure including frequency band selection and fault signature identification. Band selection is based on wavelet packet transform and signal energy decomposition. Wavelet packet transform decomposes the spectrum of a bearing vibration signal into finite frequency bands. Then Root Mean Square is applied to locate the band with the highest energy suitable for envelope analysis. Further, cepstrum analysis is employed to identify repetitive nature in the enveloped signal which is associated to bearing fault signature. The techniques developed are verified using experimental data from Bearing Data Center of Case Western Reserve University.

Acknowledgements

I would like to thank Case Western Reserve University for providing bearing benchmark data.

I sincerely acknowledge my supervisor Dr. Qiao Sun for her guidance and help. Dr. Sun is a respectable professor and she impresses me with her knowledge, insight and diligence so much.

I greatly appreciate my family for their huge support and encouragement.

Dedicated to my daughter

Table of Contents

List of Tables

List of Figures and Illustrations

List of Symbols, Abbreviations and Nomenclature

Symbol	Definition
TSA	Time synchronous averaging
LP	Linear prediction filtering
SANC	Self-adaptive noise cancellation
DRS	Discrete/Random separation
SK	Spectral kurtosis
STFT	Short Time Fourier transform
CWT	Continuous wavelet transform
DWT	Discrete wavelet transform
WPT	Wavelet packet transform
RMS	Root mean square
α	Bearing contact angle
D_c	Bearing pitch circle diameter
D_b	Bearing rolling element diameter
Z	Rolling element number
f_i	Inner race rotation frequency
f_o	Outer race rotation frequency
f_c	Bearing cage rotation frequency
f_{bpfi}	Bearing ball pass frequency on the inner race
f_{bpfo}	Bearing ball pass frequency on the outer race
f_{bsf}	Bearing rolling element rotation frequency about their own axes of rotation
ξ	Damping ratio
ω_n	Natural frequency
ω_r	Damped natural frequency
f_d	Bearing fault frequency
T_d	$\frac{1}{f_d}$
T_e	Time constant of the decay
$c(t)$	Carrier signal
ω_c	Carrier frequency
$a_m(t)$	Amplitude modulating signal
$\varphi_m(t)$	Frequency modulating signal
$c_+(t)$	Analytic signal or pre-envelope

Chapter One: **INTRODUCTION**

1.1 Background and motivation

The motivation for this research is to develop an automatic diagnostic method for localized fault detection of rolling element bearings. This method aims at minimizing the dependence on vibration analysts for bearing fault diagnosis.

A rolling element bearing is used primarily to support rotating shafts in mechanical equipment. It can be found in a wide variety of industries and is an important asset in mechanical equipment. However, it is also the most vulnerable component of a machine because it is often under high load and high running speed conditions. Rolling element bearing failures represent a high percentage of breakdowns in rotating machinery and result in catastrophic failures in many situations. It is reported that 25 percent of machine shutdowns in US Naval aircraft are due to bearing failures [1]; turbine bearing system failures constitute one of the leading causes of unplanned outages of power plants and are responsible for loss amounting to 1.1 – 1.8% of theoretical value of power generation [2]. Hence bearing condition monitoring becomes an important and necessary task for the maintenance professionals.

Existing strategies for rolling element bearings include time-based maintenance and condition-based maintenance [3]. Time-based maintenance refers to inspections and repairs conducted at regular intervals. In this way, catastrophic failures can be avoided to large extent, but a small number of unforeseen failures can still occur. In addition, maintenance may be carried out too often and unnecessary component replacement may result. Condition-based maintenance is to schedule maintenance based on machine

condition monitoring in order to prevent production loss and serious accidents. As more and more reliable condition monitoring techniques have become available, this maintenance strategy is considered the best in most cases in recent years.

Condition monitoring is a method of ascertaining whether an asset is in good condition or not by the surveillance of its behaviour based on operational parameters or indicators. For rolling element bearings, condition monitoring techniques include oil analysis, thermography, vibration analysis, and acoustic analysis [4-6]. In the power generation industry, generator current and power can be analyzed for bearing condition monitoring also [7]. Vibration analysis can be used for continuous as well as intermittent condition monitoring with the best sensitivity and immediate response to structure change. Thereby, it has been the most popular method among all techniques in this area [8-13].

Vibration analysis refers to processing and analyzing machine vibration signals obtained with vibration transducers like proximity probes, velocity pickups, accelerometers, shaft encoders and laser vibrometers etc. Vibration analysis is by far the most prevalent method for machine condition monitoring, especially for rotating equipment. An operating machine generates vibration which can be directly linked to periodic events in machine operation, such as shafts rotating, gear tooth meshing, bearings running, and so on. Vibration signals contain a significant amount of information related to the machine components. Many diagnostic techniques have been developed to detect faults and to locate their sources.

Among a large quantity of vibration analysis methods, spectral analysis is the most fundamental and most common technique. Fourier transform is used to project vibration signals from time domain to frequency domain.

$$\mathrm{X}(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt, \tag{1.1}$$

where $x(t)$ is a time domain signal, f is frequency, and $\mathrm{X}(f)$ is the frequency spectrum of $x(t)$ [14]. Then this spectrum is used for vibration analysis. In general, different elements in a mechanical system produce vibration at different frequencies. For example, a shaft fault generates vibration at the fundamental and harmonics of the shaft rotating frequency; failures derived from a gearbox cause vibration at the gear tooth meshing frequencies which are usually different for each pair of gears in the gearbox. Thus such changes in the spectrum as magnitude increase at a particular frequency or the appearance of modulation sidebands around that frequency can often be traced to a certain element in a complex system [15].

However, a bearing fault is much more difficult to diagnose than shaft or gear faults. Spectral analysis is likely to fail in most situations. At early stage when a localized bearing fault such as fatigue spalling has developed, resonance of the bearing and its supporting structure is excited every time the fault is run over by another component. Such resonance is typically at frequencies higher than 1 kHz. In this case, no dominant signal would appear at bearing characteristic frequencies. When bearing faults continue to develop and spread to a larger area and become more serious, bearing vibration signal will manifest itself at the bearing characteristic frequencies in lower than 1 kHz band. Nevertheless, the low frequency band is also full of signals with much higher energy from shafts and gears. This makes detecting bearing fault signal difficult.

Envelope analysis (also called high-frequency resonance technique) was first developed in the 1970s [16-18]. The early version of envelope analysis used analog techniques with inherent limitations. As digital techniques, especially Hilbert transform,

are introduced, envelope analysis has become popular for bearing fault diagnosis recent years. Since resonance signals generated by bearing faults are usually concentrated in a frequency band, locating such band may lead to the identification of a bearing fault. Envelope analysis isolates a resonance frequency band, excluding most of the vibration generated by other parts of a machine, and then demodulates the isolated signal to get the envelope signal. The appearance of spectral lines at the bearing characteristic frequencies indicates the presence and the location of a fault in the bearing [10, 19-22]. Further details are presented in Chapter 3.

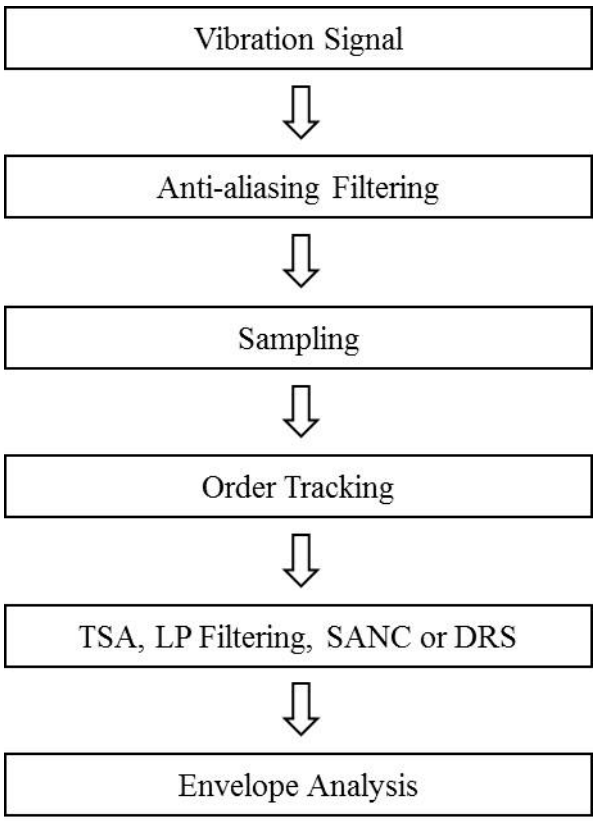


Figure 1.1 Procedure of bearing fault diagnosis

Another challenge with bearing fault diagnosis is when the bearing operates under variable speed and load condition. For envelope technique to be applicable, speed change must be compensated; otherwise the spectrum of the envelope signal will be smeared. Spectral lines related to shaft fundamental frequency are located in the low frequency band, not affecting envelope analysis; however gear meshing induced spectral lines are likely to extend into the resonance frequency band, interfering with envelope analysis as a result. Consequently, envelope analysis is preceded by pre-processing in order to improve the performance of bearing fault diagnosis. Fig. 1.1 shows the procedure of bearing fault diagnosis.

A vibration signal collected by a transducer is first anti-aliasing filtered in accordance with the frequency response range of the transducer, and is sampled to obtain a digital signal. Order tracking is a technique that compensates for changes in the operation speed. A tachometer or shaft encoder can be installed on the shaft to extract the speed information [23, 24]. When speed change is less than 30%, the instantaneous speed can be obtained by phase demodulation of the vibration signal itself [25-29]. Then the vibration signal is resampled according to equal angular increments of the shaft rotation [23, 30, 31]. By this means, the variable speed operation in the time domain is converted to a constant profile in the angular domain. Fourier transform projects the signal from the angular domain into the order domain which is actually a normalized frequency domain in reference to the shaft rotation frequency. Time Synchronous Averaging (TSA) [15, 32-34], Linear Prediction filtering (LP) [35], Self-Adaptive Noise Cancellation (SANC) [36] or Discrete/Random Separation (DRS) [37] are techniques used to remove signals generated by shafts, blades and gears. TSA, based on the phase information extracted by

the shaft encoder or keyphasor, divides the vibration signal into sections. Each section matches with a shaft revolution. Then ensemble averaging among all sections is taken. By this means, a comb filter is constructed with pass band containing the shaft rotation frequency and its harmonics. This comb filter extracts signals of shafts, blades and gears which are then subtracted from the original signal. The residual signal is dominated by bearing signals. LP, SANC and DRS apply different algorithms to train a Wiener filter which is used to remove other signals and to obtain the residual signal dominated by bearing signals. Subsequently, the residual signal is processed by envelope analysis. The focus of this research is on vibration signals dominated by bearing signals.

1.2 Problem definition

Correct selection of the frequency band that contains bearing fault induced resonance is crucial to the performance of envelope analysis [38]. Traditionally, this is done manually by a vibration analyst each time a signal is to be analyzed. It means a huge amount of repetitive work, making bearing fault diagnosis time consuming.

Spectral Kurtosis (SK) was proposed to select the frequency band for the envelope analysis automatically in 2006 [39]. A bearing signal is a train of impulses and an impulse has much higher kurtosis value than Gaussian type signals. Kurtosis is a statistical parameter, defined as [40]

$$\text{Kurtosis} = \frac{\frac{1}{N}\sum_{i=1}^{n}(x_i-\bar{x})^4}{(\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2)^2}, \tag{1.2}$$

where x is the sampled time signal, i is the sample index, N is the number of samples and $\bar{x}$ is sample mean. This normalized fourth moment is designed to reflect the “peakedness”

of the signal. This technique is designed to finding the frequency band by high kurtosis value. Short Time Fourier transform (STFT) coefficients $H(t,f)$ of each time window of the vibration signal is calculated for kurtosis individually, all of which are averaged to result in the spectral kurtosis, as shown in Fig. 1.2. This method is similar to the Welch method for power spectral estimation. The pivot of this method is that the time window must encompass only one impulse, otherwise the bridge between impulses will smooth kurtosis out. However, it is very difficult to determine the time window length.

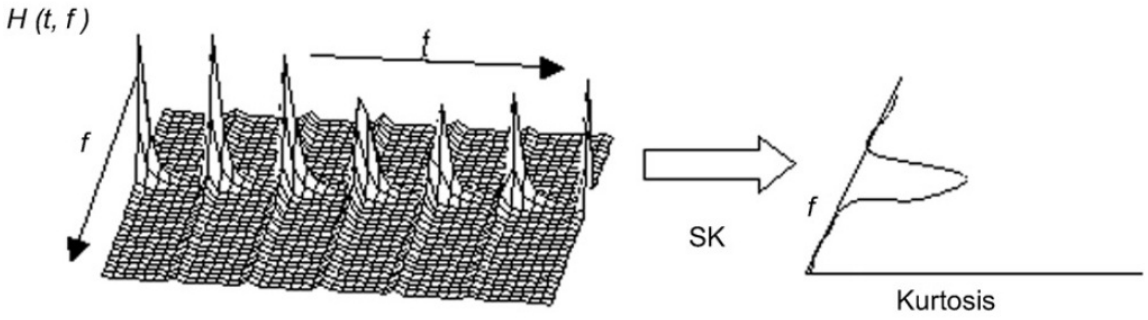


Figure 1.2 Spectral kurtosis (From [41])

Then kurtogram is developed to overcome this problem in 2007 [41, 42]. Multirate filter banks or the complex Morlet wavelets are used to decompose the vibration signal and kurtosis is calculated based on each decomposed frequency band to find the frequency band with large enough kurtosis value. Nevertheless, the kurtosis value depends on both central frequency and bandwidth of each frequency band, so it is hard to determine the decomposition mode. In practice, many combinations of different central frequency and bandwidth have to be tried in order to find the suitable frequency band for envelope analysis, which needs considerable computation. As shown in Fig. 1.3,

1, 2, 3, 4, 6, 12 and 24 levels of discrete wavelet decomposition are tried and level 3 is turned out to be the best in this case. The color scale in Fig. 1.3 denotes Kurtosis value.

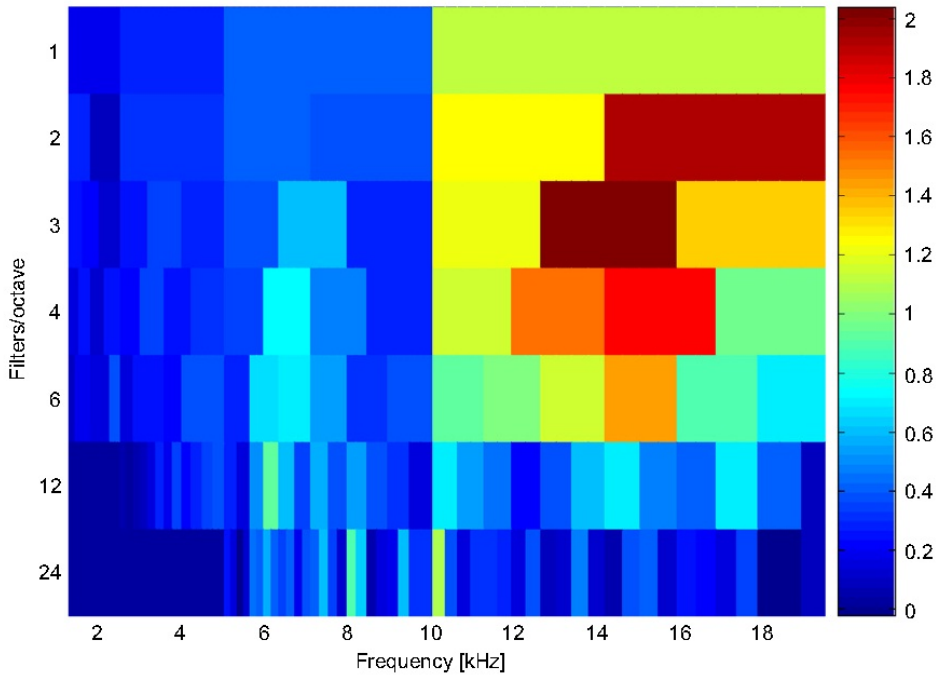


Figure 1.3 Kurtogram (From [42])

Consequently, it is beneficial to develop an algorithm to determine the frequency band for envelope analysis automatically with less cost.

1.3 Research objective

The main objective of this thesis is to develop a more efficient method to determine the frequency band for envelope analysis automatically; moreover, to propose an automated algorithm to extract bearing fault signature. These two methods, associated with

traditional techniques mentioned above, constitute a whole automatic fault diagnostics for rolling element bearings.

1.4 Organization of the thesis

The thesis is organized in seven chapters including this chapter.

- ✧ Chapter 2 presents the structure and failure modes of the rolling element bearing, characteristic frequencies of the rolling element bearing, and the bearing localized incipient fault signal and system response. This chapter is the basis of rolling element bearing fault diagnostics.
- ✧ Chapter 3 elaborates the mechanism of signal modulation and a very valuable demodulation tool – Hilbert transform. They are followed by the introduction of the most popular and powerful localized defect diagnostic method of the rolling element bearing – envelope analysis. Then the spectral patterns are demonstrated of different types of bearing fault signals.
- ✧ Chapter 4 deals with how to use wavelet packet transform to automatically determine the best frequency band for envelope analysis. Review of continuous wavelet transform and discrete wavelet transform precedes wavelet packet transform.
- ✧ Chapter 5 introduces cepstrum method which is used to automatically identify bearing fault signature.
- ✧ Chapter 6 applies automatic diagnostic methods developed by this thesis to experimental bearing vibration data. Effectiveness and validation of these methods are verified by experimental results.
- ✧ Chapter 7 presents conclusion, discussion and future work.

1.5 Scientific contribution

The scientific contribution of this thesis to condition monitoring and fault diagnostics of the rolling element bearing are two techniques for automatic diagnosis.

- An algorithm was developed that uses wavelet packet transform to decompose the bearing vibration signal into different frequency bands and calculates Root Mean Square of these bands in order to determine the best band for envelope analysis.
- Cepstrum is applied first and originally to process envelope signals in order to discover bearing fault signature automatically.

Chapter Two: **STRUCTURE, FAILURE MODE AND FAULT SIGNATURE OF ROLLING ELEMENT BEARINGS**

In this chapter, fundamental aspects about fault diagnosis of rolling element bearings are introduced. Bearing structure and failure modes are discussed, followed by the bearing fault signature and characteristic frequencies which serve as reference information for bearing fault diagnosis. Signal features generated by localized bearing faults embedded in system response are presented. Benchmark data are used to illustrate those features.

2.1 Structure of rolling element bearings

Rolling element bearing, antifriction bearing, and roller bearing belong to the class of bearings where load is transferred through elements in rolling contact rather than sliding contact. Bearings are manufactured to take pure radial loads, pure thrust loads, or a combination of the two kinds of loads [43]. This thesis focuses on rolling element bearings supporting radial loads. Fig. 2.1 shows the structure of a rolling element bearing. A bearing consists of four essential parts which are inner race, outer race, rolling element (roller or ball) and cage (separator).

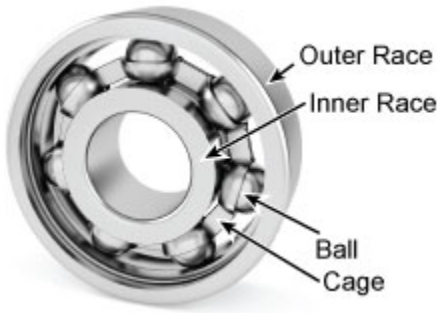


Figure 2.1 Structure of a rolling element bearing (Courtesy of Lion Precision)

Internal clearance is one of the most important factors affecting bearing performance. It allows relative movement between outer raceway and inner raceway when they are lightly pushed in opposite directions. Radial clearance allows diametrical direction movement and axial clearance allows movement in the shaft longitudinal direction. This research focuses on bearings under radial loads. The amount of clearance influences the load distribution in a bearing. Fig. 2.2 is a schematic diagram showing the load zone under a vertical load F_r. It spans about 120 degrees. $Q(\psi)$ denotes the load magnitude in Newton per unit length at location ψ. Bearing faults may occur at the

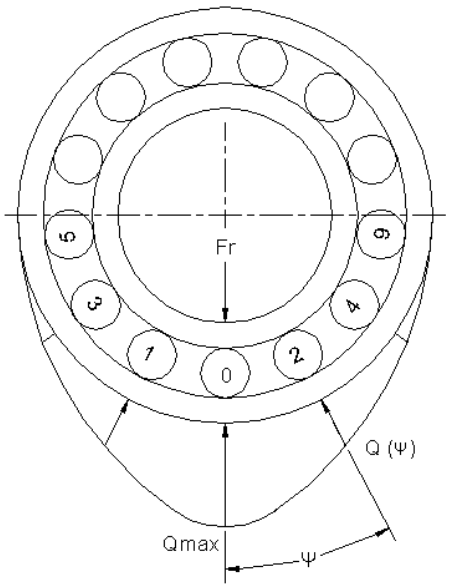


Figure 2.2 Bearing load zone

surface of outer raceway, inner raceway, cage or rollers. Bearing faults produce impact whenever they are in contact with another surface while inside the load zone. In Fig. 2.2,

rollers 0-4 are inside the load zone. If any one of these rollers comes in contact with faults on other parts such as inner or outer raceways, impact will be generated and visible in vibration signals. In most applications, the outer race is fixed and the inner race is rotating with the shaft. Hence, fatigue spalls on outer raceway are almost always found inside the load zone, near the six o'clock location. Faults situated on other parts of the bearing generate periodic impact associated to their appearance inside the load zone.

2.2 Failure modes of the rolling element bearing

In general, failure of a rolling element bearing operating under normal condition results from material fatigue and wear on running surfaces. Premature bearing failure can be caused by a large number of factors. Among them, such factors prevail as fatigue, wear, plastic deformation, corrosion, brinelling, poor lubrication, faulty installation and incorrect design [40, 44].

2.2.1 Fatigue

When a given alternating loading on a material repeats many times, the material will damage even if the stress caused by the loading remains in this material's elastic range. This is known as fatigue [45]. Fatigue damage begins with the formation of a minute crack below the bearing surface. As loading continues, the crack extends to the surface, manifesting itself as pitting, spalling or flaking of the bearing raceways or rolling elements [46, 47]. Fig. 2.3 shows pitting failure on a bearing inner raceway.

Figure 2.3 Bearing pitting failure (Courtesy of JAD Associates)

2.2.2 Wear

Wear is mainly caused by foreign particles entering the bearing through inadequate sealing or contaminated lubricant. The abrasive foreign particles roughen the contacting surface and severe wear changes raceway and rolling element profiles, causing changes in bearing clearance. Consequently, rolling friction increases considerably which can lead to high level of slip and skid, and eventually breakdown. This failure is shown in Fig. 2.4.

Figure 2.4 Bearing wear failure (Courtesy of Emerson Bearing)

2.2.3 Plastic deformation

Plastic deformation is generally caused by excessive loading. The symptom of this failure mode is indentation of the bearing raceway, causing the bearing to rotate very unevenly and producing large vibration.

2.2.4 Corrosion

Corrosion damage appears when water, acids or other contaminants in the oil enter the bearing assembly through damaged seal. As a result, rust on the raceway occurs. Rust particles affect bearing vibration similarly as wear while rust pits behave like fatigue, as shown in Fig. 2.5.

Figure 2.5 Bearing corrosion failure (Courtesy of Emerson Bearing)

2.2.5 Brinelling

Brinelling refers to distributed indentations with equal spacing over the entire raceway circumference, as shown in Fig. 2.6. It can be caused by static overloading or by vibration and shock loads to a rolling element bearing. Brinelling often leads to excessive

vibration which accelerates bearing wear. Each indentation acts like a small fatigue site and brinelling behaves like multiple point fatigue failure mode.

Figure 2.6 Bearing brinelling failure (Courtesy of Linear Motion Tips)

2.2.6 Poor lubrication

Poor lubrication in a bearing leads to skidding and slipping between the mating surfaces. It increases friction, and generates heat which can cause sticking. Without adequate lubrication, contacting surfaces will weld together at the high stress region of Hertzian contact, only to be torn apart when the rolling element moves on [48], as shown in Fig. 2.7. The increased temperature anneals the bearing parts, reducing material's hardness and fatigue life and aggravating wear.

Figure 2.7 Bearing seizure failure (Courtesy of Emerson Bearing)

2.2.7 Faulty installation

Faulty installation includes excessive preloading or misalignment due to excessive force used in mounting the bearing components. Excessive preloading or misalignment is very likely to result in indentation, scoring damage, premature fatigue, heave rolling element wear, overheating and raceway wear track.

2.2.8 Incorrect design

Incorrect design involves poor selection of bearing type and size for required operation which causes inadequate load carrying capacity or rating speed. Inadequate support can give rise to excessive clearance of the mating parts of the bearing resulting in slippage of the inner race on the shaft. Fretting follows, generating abrasive metal particles. Increased friction and temperature result in catastrophic failure.

2.3 Bearing characteristic frequencies

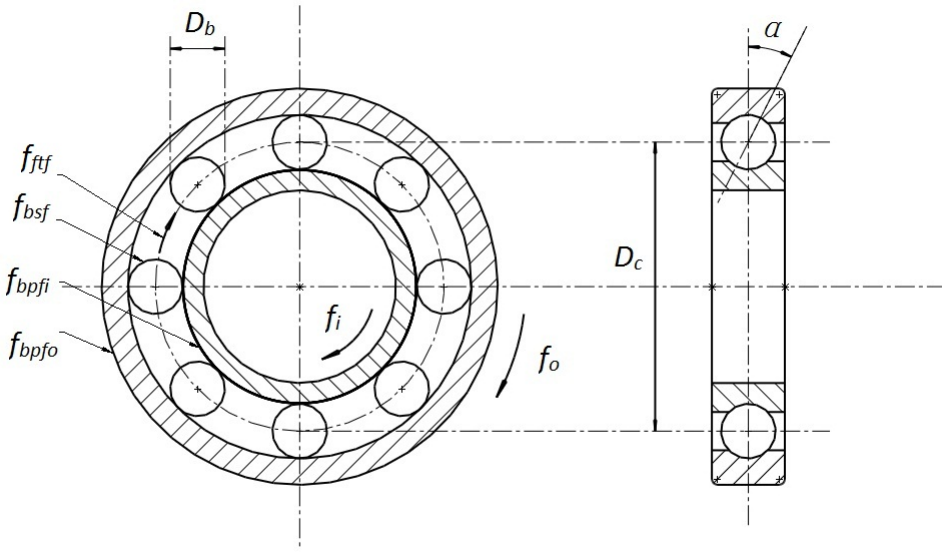


Figure 2.8 Bearing characteristic frequencies

A number of articles deal with geometry and kinematics of rolling element bearings and derive equations of motion of bearing elements [40, 49]. Fig. 2.8 shows a diagram of a typical angular contact ball bearing with rotating inner and outer races. The assumption for the kinematic analysis include pure rolling and constant contact angle α.[where is this defined?] Pitch circle and rolling element diameters are denoted as D_c and D_b respectively. Rolling element number is Z, inner race rotation frequency is f_i, and outer race rotation frequency is f_o. From [40, 49], the following equations hold.

Ball pass frequency on the inner race:

$$f_{bpfi} = \frac{Z(f_o - f_i)(1+\frac{D_b}{D_c}\cos\alpha)}{2}, \tag{2.1}$$

Ball pass frequency on the outer race

$$f_{bpfo} = \frac{Z(f_o - f_i)(1-\frac{D_b}{D_c}\cos\alpha)}{2}, \tag{2.2}$$

Fundamental train frequency (cage rotation frequency):

$$f_{ftf} = \frac{f_i(1-\frac{D_b}{D_c}\cos\alpha)}{2} + \frac{f_o(1+\frac{D_b}{D_c}\cos\alpha)}{2}, \tag{2.3}$$

Ball spin frequency (rolling element rotation frequency about their own axes of rotation):

$$f_{bsf} = (f_o - f_i)\frac{D_c}{D_b}(1 - (\frac{D_b}{D_c}\cos\alpha)^2), \tag{2.4}$$

In practice, outer race is used fixed like in our case. The above equations can be simplified:

$$f_{bpfi} = \frac{Zf_i(1+\frac{D_b}{D_c}\cos\alpha)}{2}, \tag{2.5}$$

$$f_{bpfo} = \frac{Zf_i(1-\frac{D_b}{D_c}\cos\alpha)}{2}, \tag{2.6}$$

$$f_{ftf} = \frac{f_i(1-\frac{D_b}{D_c}\cos\alpha)}{2}, \tag{2.7}$$

$$f_{bsf} = f_i\frac{D_c}{D_b}(1 - (\frac{D_b}{D_c}\cos\alpha)^2). \tag{2.8}$$

These frequencies are the bearing characteristic frequencies which are determined by bearing geometry and running speed. For localized incipient faults such as pitting and fatigue spalling, impact is generated when mating surface come in contact with the fault.

A train of impacts appears when the bearing stays in operation. The frequency of the impact can be considered as the fault frequency. Matching the fault frequency with bearing characteristic frequencies illuminates which bearing element is faulty. It should be noted that these theoretical bearing characteristic frequencies are approximate due to the inevitable slip, contact angle variation and manufacturing variation etc.

2.4 Bearing localized incipient fault signal and the system response

In order to monitor bearing condition during operation, a vibration transducer is often installed at a close proximity of the bearing, for instance, on the casing of the bearing housing. The transmission path from the bearing fault to the transducer can be regarded as a mechanical system. The impact produced by a rolling element striking the localized incipient fault can be referred to as the input to this mechanical system. Vibration signals detected by the transducer are the system response.

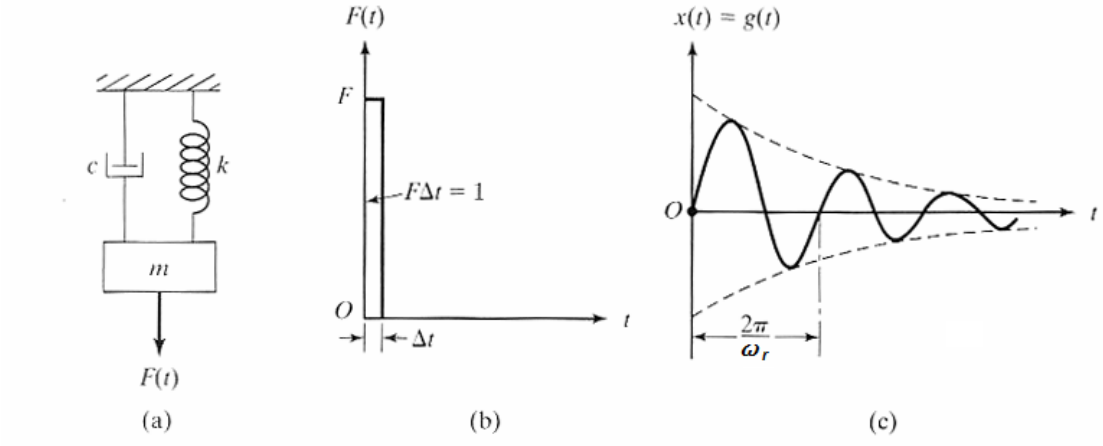


Figure 2.9 SDOF system shock vibration

As shown in Fig. 2.9 (a), we consider the mechanical filter as a single degree of freedom(SDOF) mass-spring-damper system whose equation of motion is

$$m\ddot{x} + c\dot{x} + kx = F(t), \tag{2.9}$$

or

$$\ddot{x} + 2\xi\omega_n\dot{x} + \omega_n^2 x = \frac{F(t)}{m}, \tag{2.10}$$

where $\xi = \frac{c}{2\sqrt{km}}$ is the damping ratio and $\omega_n = \sqrt{\frac{k}{m}}$ is the natural frequency. m, k, and c are mass, stiffness and damping coefficient, respectively.

The impact force generated by the bearing localized fault striking a mating surface is represented by the forcing function *F(t)* over a small time interval Δt , as shown in Fig. 2.9 (b). When Δt approaches zero, this forcing function approaches the Dirac delta function.

$$\mathrm{F(t)} = \begin{cases} 0, & t \le 0 \\ 1/\Delta t, & 0 < t < \Delta t, \\ 0, & t \ge \Delta t \end{cases} \tag{2.11}$$

Let *I* be the impulse of the force *F(t)*:

$$\mathrm{I}(\Delta t) = \int_0^{\Delta t} F(t)dt = F\Delta t = 1, \tag{2.12}$$

Consider the system is initially at rest. The instant of time just prior to the application of the impact force is denoted by 0^-. The initial conditions are $x(0^-) = \dot{x}(0^-) = 0$. Thus the change in momentum due to impact is

$$I(\Delta t) = m\dot{x}(0^+) - m\dot{x}(0^-) = 1, \tag{2.13}$$

so that $\dot{x}(0^+) = \frac{1}{m}$, while the initial displacement remains at zero. Therefore system response of a unit impulse (shock) is equivalent to free vibration with a non-zero initial

velocity. For an underdamped system ($0 < \xi < 1$), the response to the initial conditions $x(0^+) = 0,\ \dot{x}(0^+) = \frac{1}{m}$ is

$$x(t) = \frac{e^{-\xi\omega_n t}}{m\omega_r} \sin \omega_r t, \tag{2.14}$$

where $\omega_r = \omega_n\sqrt{1-\xi^2}$ is the damped natural frequency which is also the resonance frequency. As shown in Fig. 2.9 (c), the impulse response of the system is an exponentially decaying sinusoid. A general form of the impulse response of the system can be described as

$$x(t) = Ae^{-\frac{t}{T_e}} \sin(\omega_r t + \theta_0), \tag{2.15}$$

where T_e is the time constant of the decay, and θ_0 is the initial phase.

For a continuous operation, a train of impacts occurs and the general form of the system response can be represented by

$$x(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT_d)\, A(t) * [e^{-\frac{t}{T_e}} \sin(\omega_r t + \theta_0)], \tag{2.16}$$

where $T_d = \frac{1}{f_d}$, f_d is the bearing fault frequency, and $A(t)$ reflects the effect of the bearing load distribution. $\delta(t)$ is the Dirac delta function.

For a multiple degree of freedom system, modal analysis can be used to decompose system variables into a set of orthogonal eigen-modes. Similar expressions as eq. (2.16) are used to describe the eigen-modes. Theoretically, a continuous mechanical system possesses infinite number of eigen-modes. Different operation and fault conditions may excite different modes. The task of bearing fault diagnosis is to isolate the response of particular modes containing fault information from the system response. It should be noted that accelerometers are the most popular vibration transducers which

collect acceleration signals. Acceleration is the second order derivative of displacement, so the system response in acceleration carries the same frequency information as displacement above.

2.5 Benchmark data

The bearing data center of Case Western Reserve University published bearing signal data on-line for researchers to validate new theories and techniques. All data are annotated with bearing geometric, operating condition and fault information. Fig. 2.10 shows the test stand from which the test data are collected.

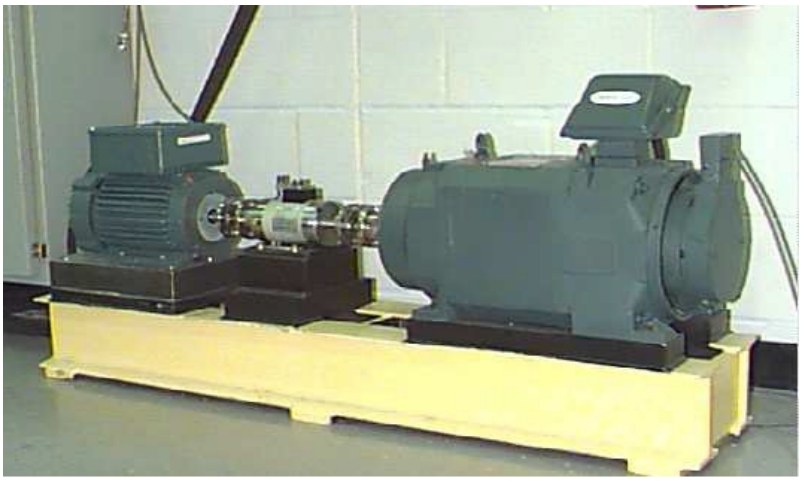


Figure 2.10 Test stand (Courtesy of Case Western Reserve University)

According to the description given by the provider of the test data, the test stand consists of a 3 hp motor (left), a torque transducer/encoder (center), and a dynamometer (right). The test bearings, including drive end and fan end bearings, support the motor shaft. Single point faults were introduced to the test bearings using electro-discharge

machining. In this thesis, we analyze signals of drive end bearing only. Its vibration data were collected using 3 accelerometers, which were attached to the housing with magnetic bases at 3, 6 and 12 o'clock positions. 6 o'clock position is located in the load zone, 3 o'clock is orthogonal to load zone, and 12 o'clock is in the bearing clearance zone. For better Signal-Noise-Ratio, we use signals from 6 o'clock for analyses in this thesis. Vibration signals were collected using a 16 channel DAT recorder at 12,000 samples per second. Speed and power data were collected using the torque transducer/encoder and were recorded by hand. Vibration signals generated by this test stand are dominated by bearing signals, since there are no gears and shafts are healthy.

Fig. 2.11 shows the spectrum of the vibration signal taken from the drive end bearing with a single pit on the outer raceway. The bearing is a deep groove ball bearing and the model is 6205-2RS JEM SKF. The diameter and depth of the pit are 0.18 mm and 0.28 mm respectively. In machine fault diagnosis, one way of fault detection is to identify the spectrum pattern change. This can be done by detecting the location of particular spectra lines with increased magnitude. As such the absolute value of the spectral magnitude is insignificant. In all spectral diagrams, magnitude is relative and we omit the units. As shown in Fig. 2.11, two vibration modes (resonances) are captured by the accelerometer. It is typical that resonance signals caused by bearing faults are visible in the kilohertz range rather than at the bearing characteristics frequencies. This makes bearing fault diagnosis by searching for spectrum peaks less effective.

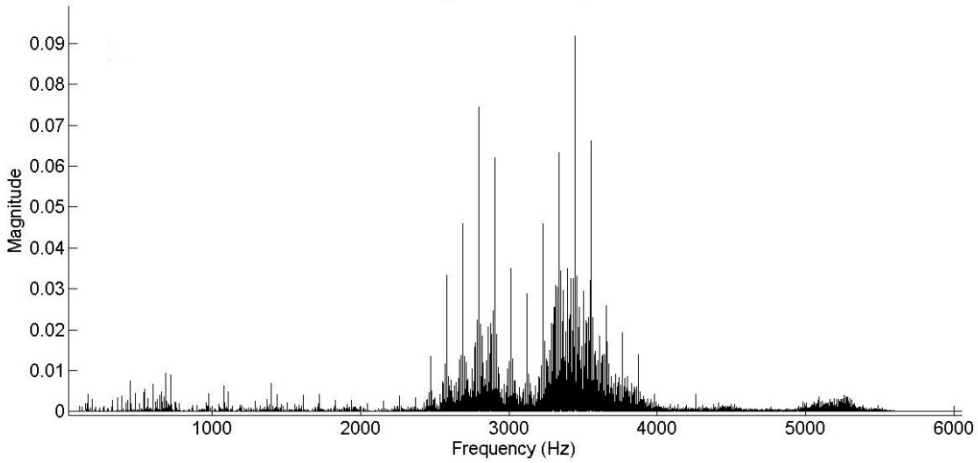


Figure 2.11 Spectrum of the vibration signal of a bearing with a single pit

As comparison, Fig. 2.12 shows the spectrum of the vibration signal generated by the drive end bearing without any fault. There are 4 high order harmonics of shaft frequency near 4 kHz. No resonance band as shown in Fig. 2.11 is found in this spectrum.

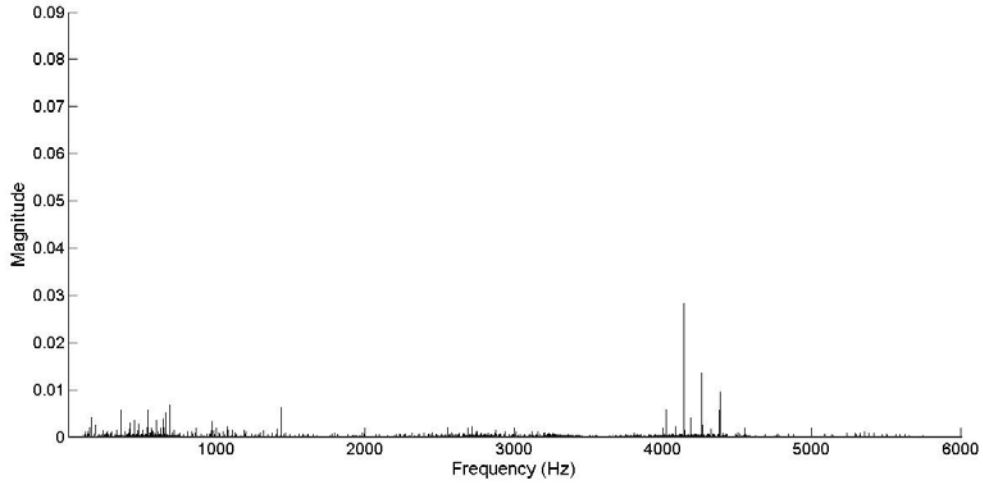


Figure 2.12 Spectrum of the vibration signal generated by a healthy bearing

Chapter Three: **ENVELOPE ANALYSIS AND SPECTRA OF ENVELOPE SIGNALS**

From the previous discussion, we know that bearing fault signature is embedded in the resonance signals commonly in high frequency bands. However, the fault induced signal is often too weak to be detected directly from the resonance signal in its measurement form. Thus, analyzing resonance signals allows us to find the key to bearing fault diagnosis. In this chapter, we discuss envelope analysis in detail.

3.1 Signal modulation

Machine vibration signals are the outcome of machine response to many sources. In normal conditions, vibration signals are dominated by periodic components in multiples of machine running speed. When a fault occurs, periodic impact is generated which excites machine natural modes of vibration and causes signal modulation of fault induced response with resonance. Understanding how signals are modulated to form an overall response will help to understand how to separate them from each other.

3.1.1 Amplitude modulation

Amplitude modulation is simply the product of two functions in time. Vibration signals produced by rotating machines are usually of sinusoidal form. A general amplitude modulated signal can be represented by

$$y(t) = a_m(t)c(t), \quad \text{with } c(t) = \sin(\omega_c t + \theta_c) \tag{3.1}$$

where $a_m(t)$ is the modulating signal and $c(t)$ is the carrier signal with unity amplitude and frequency ω_c. For simplicity and without losing generality, we set the carrier initial phase to zero: $\theta_c = 0$.

Let $A_m(j\omega)$, $C(j\omega)$ and $Y(j\omega)$ be the Fourier transform of $a_m(t)$, $c(t)$ and $y(t)$ respectively. From the properties of Fourier transforms [50],

$$Y(j\omega) = \frac{1}{2\pi} A_m(j\omega) * C(j\omega). \tag{3.2}$$

where "$*$" denotes convolution. The spectrum of the carrier signal is

$$C(j\omega) = \pi j\{\delta(\omega + \omega_c) - \delta(\omega - \omega_c)\}, \tag{3.3}$$

thus,

$$Y(j\omega) = \frac{1}{2} j\{A_m[j(\omega + \omega_c)] - A_m[j(\omega - \omega_c)]\}. \tag{3.4}$$

The above equation indicates that the spectrum of the modulated signal is simply that of the modulating signal, shifted in frequency by an amount equal to the carrier frequency $\pm\omega_c$. As shown in Fig. 3.1, the modulating signal modifies the amplitude of the carrier signal in time domain. It is important to observe that the modulated signal amplitude carries the shape of the modulating signal. In other words, it is "enveloped" by the modulating signal. It is also worth noting that the carrier is the one with higher frequency. In the figure, $c(t) = \sin(2\pi * 100t)$ is used as the carrier signal and the modulating signal is $a_m(t) = 0.5\sin(2\pi * 10t)$.

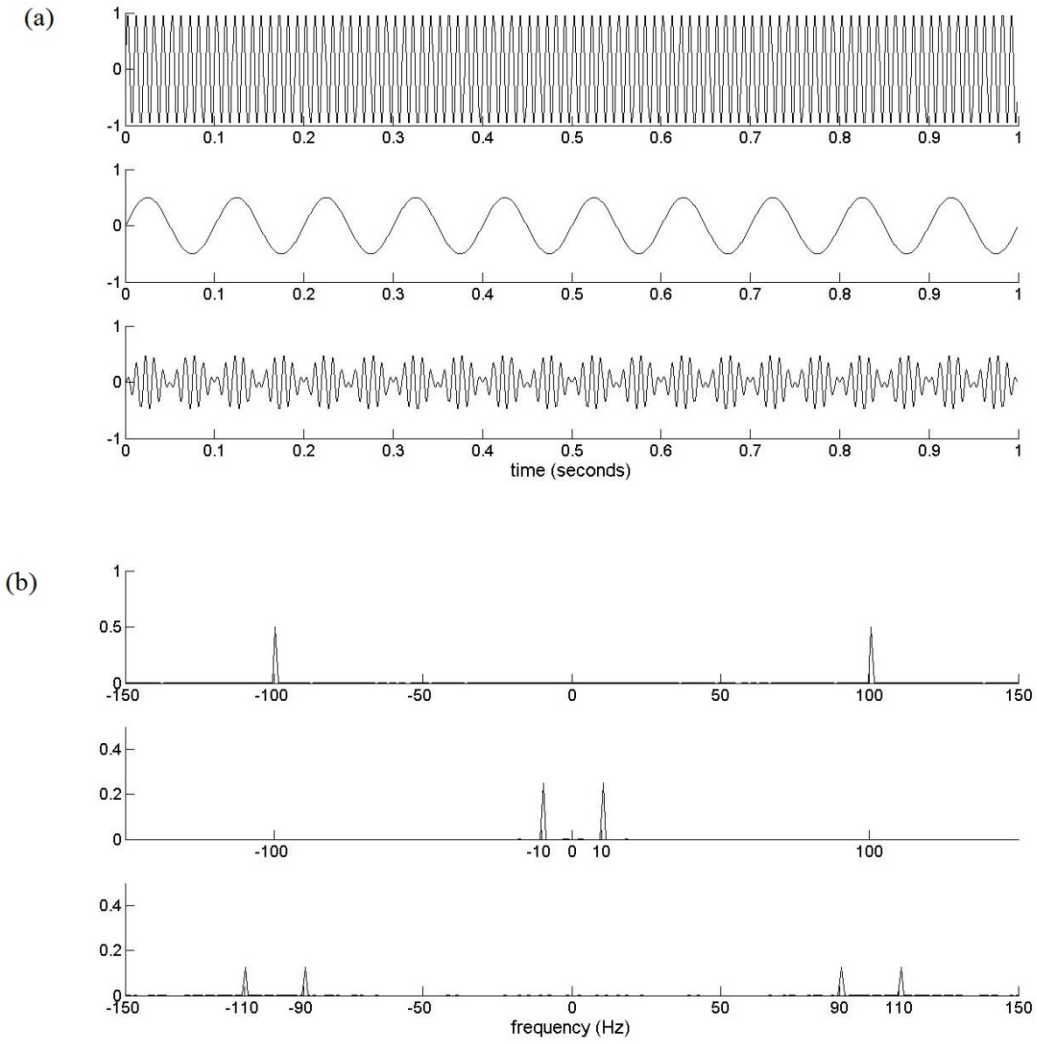


Figure 3.1 Amplitude modulation (a) waveform (b) spectrum

3.1.2 Frequency modulation

Again consider a sinusoidal carrier

$$c(t) = A_C \, sin(\omega_C t + \theta_c), \tag{3.5}$$

whereA_C, ω_C and θ_c are the amplitude, frequency and phase of the carrier respectively. We have seen in the above section that in amplitude modulation, the amplitude of the

carrier signal is varied by the modulating signal. Similarly, in frequency and phase modulation, the frequency and phase of the carrier are altered by the modulating signal. Although frequency modulation and phase modulation are different modulation forms, phase modulation with a modulating signal $\varphi_m(t)$ is identical to frequency modulation with the derivative of $\varphi_m(t)$. Likewise, frequency modulation with $\varphi_m(t)$ is identical to phase modulation with the integral of $\varphi_m(t)$ [50].

Consider a frequency modulating signal

$$\varphi_m(t) = A_m \sin(\omega_m t). \tag{3.6}$$

and a carrier

$$c(t) = \sin \omega_c t. \tag{3.7}$$

The frequency modulated signal is

$$\begin{aligned} y(t) = & \sin[\omega_c t + \int \varphi_m(t) dt] \\ = & \sin[\omega_c t - m \cos(\omega_m t) + \theta_0], \end{aligned} \tag{3.8}$$

where $m = \frac{A_m}{\omega_m}$ and θ_0 is a constant of integration. For convenience, we choose $\theta_0 = 0$, so that

$$y(t) = \sin[\omega_c t - m \cos(\omega_m t)] \tag{3.9}$$

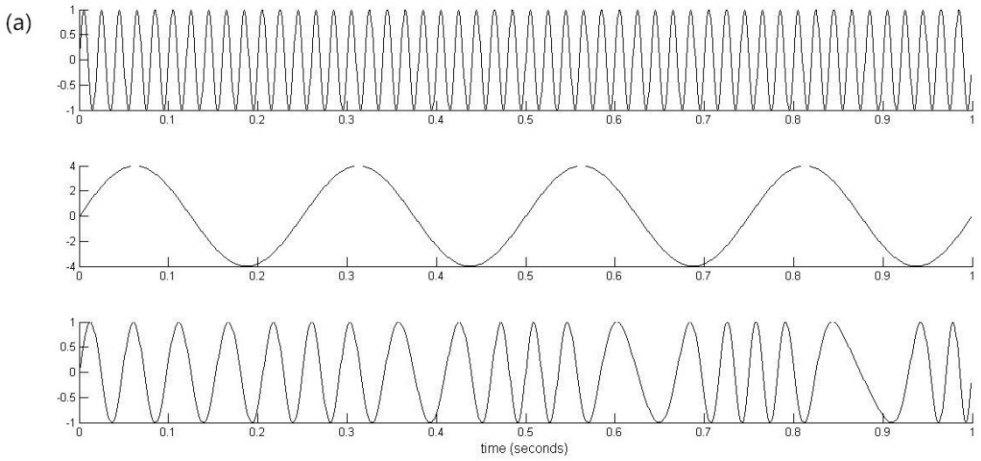
or

$$y(t) = \sin(\omega_c t) \cos[m \cos(\omega_m t)] - \cos(\omega_c t) \sin[m \cos(\omega_m t)] \tag{3.10}$$

In the above, $\cos[m \cos(\omega_m t)]$ and $\sin[m \cos(\omega_m t)]$ are periodic functions with fundamental frequency ω_m. The Fourier transform of each of these signals is an impulse train with impulses at the integer multiples of ω_m. The first term in eq. (3.10) can be considered as a sinusoidal carrier of the form $\sin(\omega_c t)$ amplitude modulated by a signal

$\cos[m\cos(\omega_m t)]$. Similarly, the second term in eq. (3.10) is a sinusoidal carrier $\cos(\omega_c t)$ amplitude modulated by signal $\sin[m\cos(\omega_m t)]$. From the previous discussions, the effect of amplitude modulation in the frequency domain is the translation of the impulse trains to the carrier frequency, so that they are centred at plus and minus ω_c [50]. Fig. 3.2 demonstrates the effect of the frequency modulation in the time domain and the frequency domain. In the figure, the carrier signal is $\sin(2\pi * 50t)$ and the modulating signal is $4\sin(2\pi * 4t)$.

In machinery vibration signals, frequency modulation is often the result of running speed variation. Hence, frequency demodulation has been used to determine the instantaneous running speed of a machine for order tracking [25]. On the other hand, amplitude demodulation is the core of envelope analysis to determine machine fault signature by isolating the fault related modulating signals.



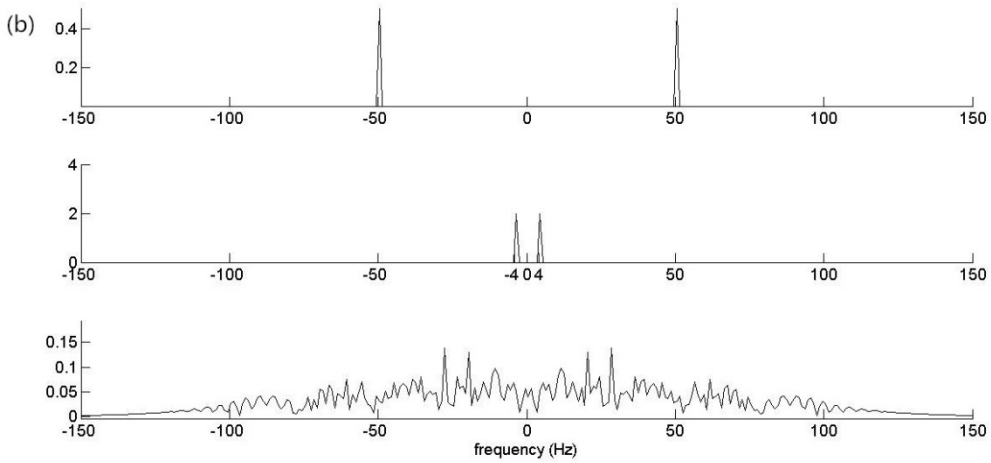


Figure 3.2 Frequency modulation (a) waveform (b) spectrum

3.2 Hilbert transform and envelope analysis

Modulated signals, as demonstrated in the preceding section, are not linear additions of two signals but non-linear distortions of each other. In addition, the signal strength of the fault-induced modulating signal is often weak compared with the overall machine vibration. It makes the detection of modulating signal difficult for fault diagnosis. Take the rolling element bearing as an example. When an incipient fault starts to develop, distinctive impulses are generated every time the fault is encountered by a mating surface. From equation (2.16), the fault induced impulse train signal is

$$a_m(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT_d)\, A(t) * e^{-\frac{t}{Te}} \tag{3.11}$$

and the carrier signal is due to the resonance frequency ω_r

$$c(t) \;=\; \sin(\omega_r t + \theta_0). \tag{3.12}$$

The frequency of the bearing fault signal, $f_d = \frac{1}{T_d}$, is the detection target. It is also the frequency of the amplitude modulating signal. Separation of the modulating signal from the modulated signal is demodulation. As indicated in Fig. 3.1 (a), the envelope of the modulated signal – that is, a smooth curve connecting the peaks of the modulated signal – would appear to be a reasonable approximation to the modulating signal. In the case of the rolling element bearing signal, the envelope signal is a train of impulse response with fast decaying amplitude. Fig. 3.3 is an example of the envelope signal of an outer race fault. Envelope analysis is an amplitude demodulation technique to extract the modulating signal based on the Hilbert transform.

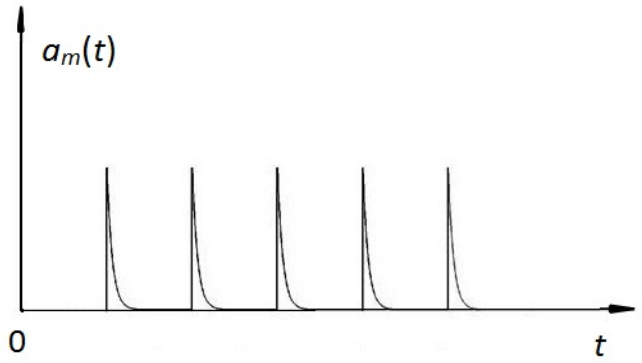


Figure 3.3 Envelope signal of an outer race fault

Envelope analysis was devised more than 30 years ago [51], and analog techniques were used in the early days. An envelope detector is essentially a band pass filter applied to the analog vibration signal centered at a structural resonance frequency. Location of the filter is crucial to the collection of fault induced signals. Bandwidth of the

filter is equally important to the inclusion of fault signals while rejecting all other non-fault related signals. After band-pass filtering, rectification is applied followed by a smoothing circuit to recover the envelope signal [11]. Due to limitations of hardware performance, the envelope detector is susceptible to noise and cannot respond to high frequency signals. Recently, envelope analysis based on Hilbert transform has been commonly implemented. It uses digital techniques and outperforms the analog counterpart. It also opens the door of this work by allowing the convenience of automatic tuning and maximizing its performance by the best selection of center location and bandwidth.

3.2.1 Hilbert transform

Hilbert transform of a function $g(t)$ is the convolution of $g(t)$ with function $h(t) = \frac{1}{\pi t}$. Function $h(t)$ is the impulse response of a linear time-invariant filter also called the Hilbert transformer. Hence, Hilbert transform of $g(t)$ is the output of the Hilbert transformer responding to input $g(t)$ [52].

$$\mathcal{H}[g(t)] = g(t) * h(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{g(\tau)}{t-\tau} d\tau \tag{3.13}$$

Obviously, Hilbert transform is a linear operator, i.e.

$$\mathcal{H}[a_1 g_1(t) + a_2 g_2(t)] = a_1 \mathcal{H}[g_1(t)] + a_2 \mathcal{H}[g_2(t)], \tag{3.14}$$

where a_1 and a_2 are arbitrary constants.What makes the Hilbert transform extremely useful is the property [52]:

$$\mathcal{H}[f_1(t) f_2(t)] = f_1(t) \mathcal{H}[f_2(t)], \tag{3.15}$$

where $f_1(t)$ and $f_2(t)$ are signals with distinct non-zero frequency contents. $f_1(t)$ contains low frequency contents only while $f_2(t)$ contains only high frequency contents. Low or high frequency is defined with respect to location *W*.

$$\begin{cases} \mathcal{F}[f_1(t)] = 0, & |\omega| \geq W \\ \mathcal{F}[f_2(t)] = 0, & |\omega| < W \end{cases}. \tag{3.16}$$

In other words, to compute the Hilbert transform of the two signals with distinct frequency concentrations, only the high frequency signal needs to be transformed. Furthermore, the low frequency signal appears unaltered as factors after the Hilbert transform. Property (3.15) is the key to successful demodulation using the Hilbert transform.

In the frequency domain, the Hilbert Transformer has the following Fourier transform

$$\mathcal{F}\left[\frac{1}{\pi t}\right] = -j\text{sgn}(\omega) = \begin{cases} -j, & \omega > 0 \\ 0, & \omega = 0 \\ j, & \omega < 0 \end{cases}. \tag{3.17}$$

Consider a sinusoidal carrier signal $c(t) = \sin(\omega_c t + \theta_c)$, its Fourier transform is

$$\mathcal{F}[c(t)] = j\pi\left\{e^{-j\theta_c}\delta(\omega + \omega_c) - e^{j\theta_c}\delta(\omega - \omega_c)\right\}. \tag{3.18}$$

Denote its Hilbert transform as $\hat{c}(t)$: $\hat{c}(t) = \mathcal{H}[c(t)]$. From property of Fourier Transform, we have

$$\mathcal{F}[\hat{c}(t)] = \mathcal{F}\left[\frac{1}{\pi t}\right]\mathcal{F}[c(t)].$$

Substitute eq. (3.17) and (3.18) in the above, we see that

$$\mathcal{F}[\hat{c}(t)] = -\pi\left\{e^{-j\theta_c}\delta(\omega + \omega_c) + e^{j\theta_c}\delta(\omega - \omega_c)\right\}$$

Taking an inverse Fourier transform on the above expression yields:

$$\hat{c}(t) = -\cos(\omega_c t + \theta_c).$$

Hence,

$$\mathcal{H}[\sin(\omega_c t + \theta_c)] = -\cos(\omega_c t + \theta_c). \tag{3.19}$$

This means that the Hilbert transform of a sinusoid signal is different from the original sinusoid only in a constant phase delay. More significantly, we can use the Hilbert property (3.15) to isolate the modulating signal as long as the frequency of the carrier signal is higher than that of the modulating signal.

$$\begin{aligned} \hat{y}(t) = \mathcal{H}[y(t)] &= \mathcal{H}[a_m(t)\sin(\omega_c t + \theta_c)] \\ &= a_m(t)\mathcal{H}[\sin(\omega_c t + \theta_c)] \\ &= -a_m(t)\cos(\omega_c t + \theta_c) \end{aligned} \tag{3.20}$$

Hence, the Hilbert transform of the amplitude modulated signal is a phase delayed version of the original signal. At any time instant, manipulate the signal and its Hilbert transformed version by

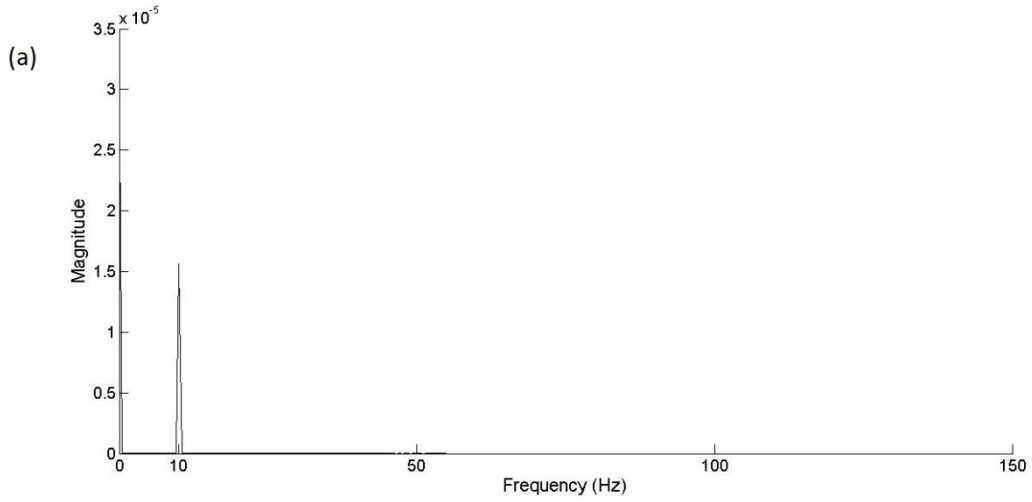
$$\begin{aligned} |z(t)| &= \sqrt{y(t)^2 + \hat{y}(t)^2} \\ &= \sqrt{(a_m(t)\sin(\omega_c t + \theta_c))^2 + (-a_m(t)\cos(\omega_c t + \theta_c))^2} \\ &= a_m(t) \end{aligned} \tag{3.21}$$

one can extract the modulating signal. Here $a_m(t)$ is also the envelope signal.

Practically, for Property (3.15) to hold true, one may anticipate the condition on the two signals $f_1(t)$ and $f_2(t)$ to be stricter than (3.16). This is because digital signal processing has limited frequency resolution plus sidebands due to finite number of processing samples, signal spectrums may appear to be overlapping in and around location *W*. We propose a modified version of condition (3.16) by the following to accommodate achievable reality:

$$\begin{cases} \mathcal{F}[f_1(t)] = 0, & |\omega| \geq W \\ \mathcal{F}[f_2(t)] = 0, & |\omega| < 2W \end{cases} \tag{3.22}$$

We demodulate the modulated signal in Fig. 3.1 by Hilbert transform and recover the modulating signal as shown in Fig. 3.4 (a). In this case, carrier frequency is 100 Hz and the frequency of the modulating signal is 10 Hz. Then, keeping the carrier signal unchanged, we change the frequency of the modulating signal into 50 Hz and 51 Hz and the demodulated signals are shown in Fig. 3.4 (b) and (c) respectively. Demodulated signals are all same as shown in Fig. 3.4 (c) when the frequency of the modulating signal is higher than 50 Hz.



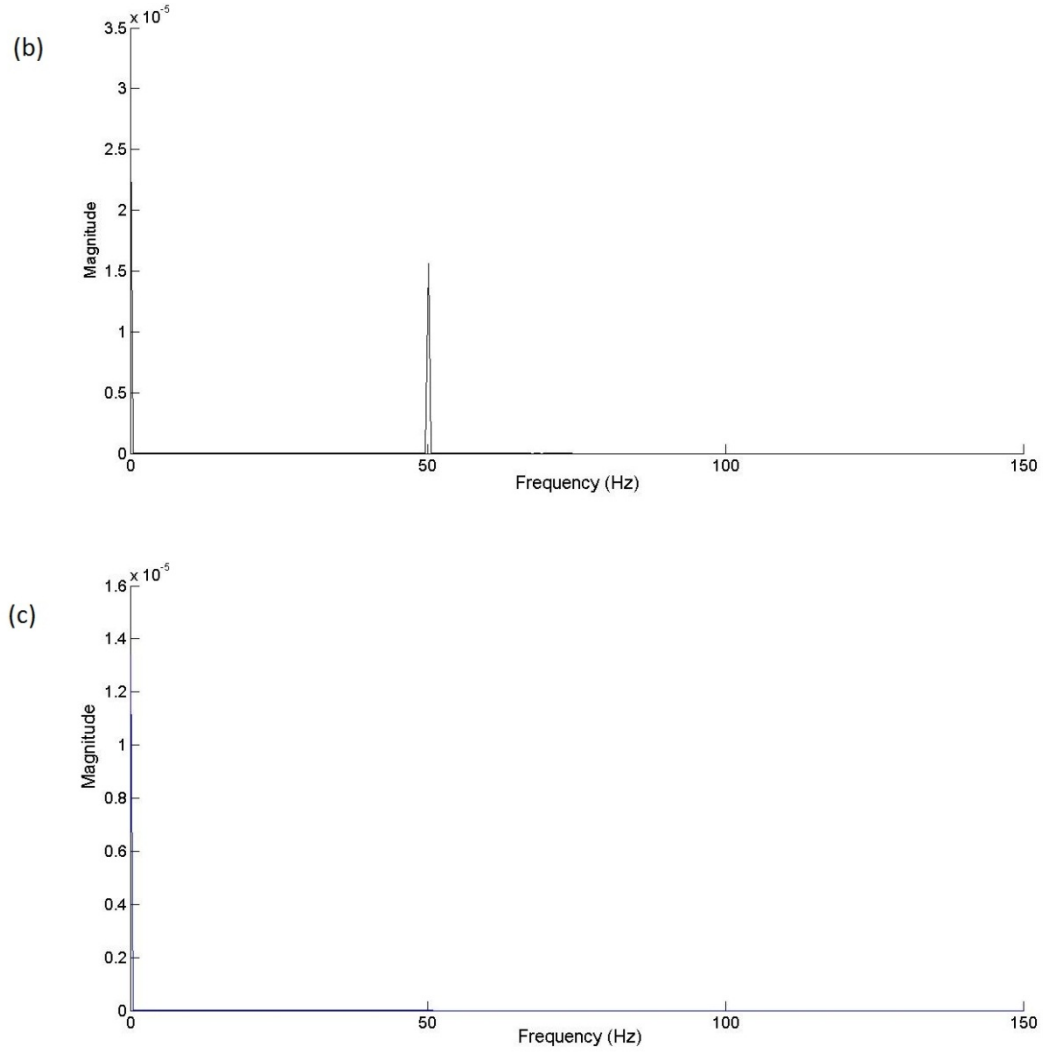


Figure 3.4 Spectra of demodulated signals with different modulating frequencies (a) 10 Hz (b) 50 Hz (c) 51 Hz

Fig. 3.4 shows that Hilbert transform can be used for demodulation only if the frequency of the modulating signal is equal to or lower than half of the frequency of the carrier signal.

For fault diagnosis of rolling element bearings, the carrier frequency is the resonance frequency ω_r as appears in eq. (3.12). Fortunately, it is much higher than the frequency of the bearing fault signal f_d in eq. (3.11) i.e. the frequency of the modulating signal. Hence, Hilbert transform can be very effective in separating the modulating signal from the overall bearing vibration signal.

3.2.2 Envelope analysis

Envelope analysis consists of the following three steps.

1. Selecting a frequency band for envelope analysis
2. Building the analytic signal or pre-envelope
3. Analyzing the spectrum of the envelope signal

It is very important to select a frequency band to effectively include the fault induced response for the envelope analysis. Sensitivity of the envelope analysis is improved by this means [11, 38]. Since the envelope signal is an approximation of the modulating signal, from eq. (3.11) and (3.12), we know that the spectrum of the envelope signal is shifted and centred at plus and minus the resonance frequency ω_r based on discussion about amplitude modulation. The transmission path from the bearing fault to the vibration transducer is a continuous mechanical system. Theoretically, there exist infinite vibration modes in this system. Depending on the modal damping properties and excitation condition, there may be several vibration modes excited in a specific machine under specific operating conditions. It is unlikely that a suitable centre frequency can be determined beforehand. This fact also speaks for the need of a reliable automated band

selection process. Vibration signals from an operating bearing indicate whether or not bearing is undergoing resonance. Any frequency band encompassing one or more resonance can be used for envelope analysis. Obviously, the frequency band with the highest resonance energy would be the best. In Fig. 3.5, we circled two vibration modes. The spectral magnitude indicates that they have similar energy. Thus, either one can be used for the envelope analysis with no appreciable difference.

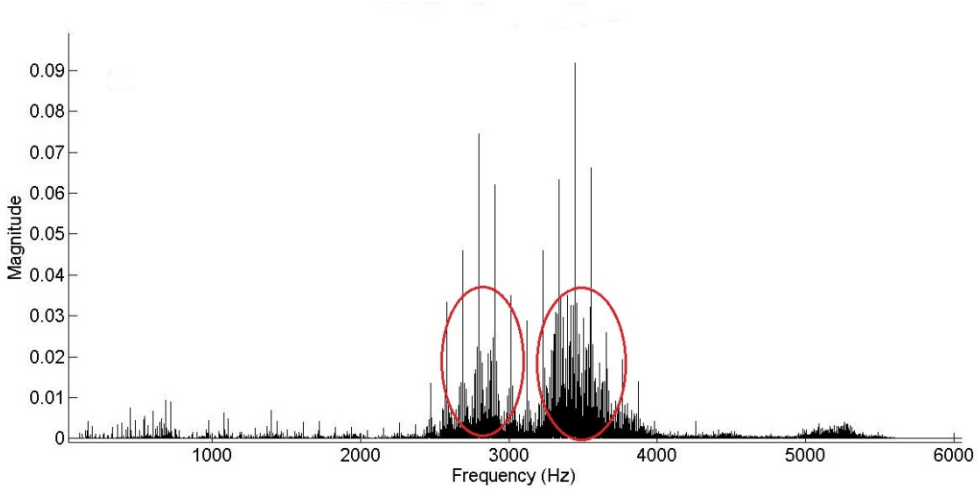


Figure 3.5 Frequency band for envelope analysis

Another important parameter for the band selection is the bandwidth. A rule of thumb is that roughly ten spikes be enclosed in the frequency band. Very few spikes in a spectrum mean a sinusoidal signal rather than a train of impact is involved in modulation. On the other hand, a wider than necessary bandwidth will include unwanted noise. In a rolling element bearing with a fixed outer race, ball passing inner race frequency f_{bpfi} is

the highest among all bearing characteristic frequencies. Therefore the band width should be about ten times f_{bpfi}.

After selecting the frequency band and band pass filtering the vibration signal, one can calculate the envelope signal through Hilbert transform. However, a more efficient method is preferred.

Given a modulated signal *y*(t), an analytic signal can be constructed:

$$y_+(t) = y(t) + j\hat{y}(t), \tag{3.23}$$

where $\hat{y}(t) = \mathcal{H}[y(t)]$. This analytic signal is also called pre-envelope. Its imaginary part is the Hilbert transform of its real part. Let $Y_+(j\omega)$, $Y(j\omega)$ and $\hat{Y}(j\omega)$ denote the Fourier transform of $y_+(t)$, $y(t)$ and $\hat{y}(t)$ respectively. From eq. (3.13) and (3.17), we have

$$\begin{aligned} Y_+(j\omega) &= Y(j\omega) + j\hat{Y}(j\omega) \\ &= Y(j\omega) + jY(j\omega)[-j\mathrm{sgn}(\omega)] \\ &= Y(j\omega) + Y(j\omega)\mathrm{sgn}(\omega) \\ &= \begin{cases} 2Y(j\omega), & \omega > 0 \\ 0, & \omega = 0. \\ 0, & \omega < 0 \end{cases} \end{aligned} \tag{3.24}$$

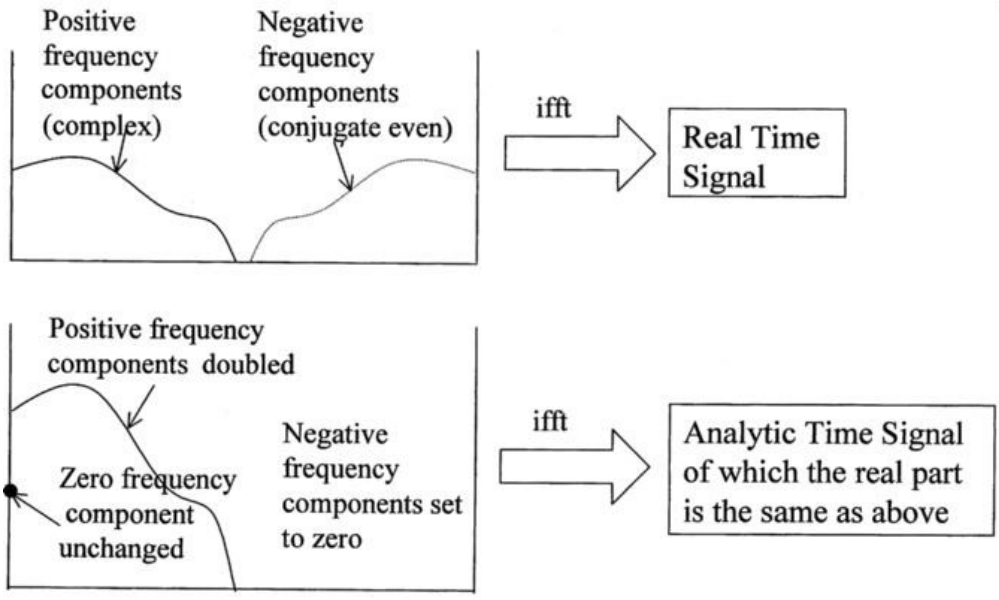


Figure 3.6 The procedure of constructing the analytic signal (from [53])

The modulus of the analytic signal is the envelope signal according to eq. (3.21). The spectrum of the analytic signal can be acquired simply by setting the negative part of the spectrum of $y(t)$ to be zeros and doubling the positive part, as shown in Fig. 3.6. The benefit of building the analytic signal to get the envelope signal is that we avoid calculating the Hilbert transform of the vibration signal $y(t)$, making the fault diagnosis more efficient.

Once the envelope signal is available, its spectrum will be analyzed to determine the frequency of the bearing fault signal $f_d = \frac{1}{T_d}$. The envelope signal obtained by the preceding procedure is represented by eq. (3.11). The fault frequency is just associated with the first term; there are still two more terms $A(t)$ and $e^{-\frac{t}{T_e}}$ affecting the spectrum of

the envelope signal. This will be discussed in the following section. The procedure of envelope analysis is shown in Fig. 3.7.

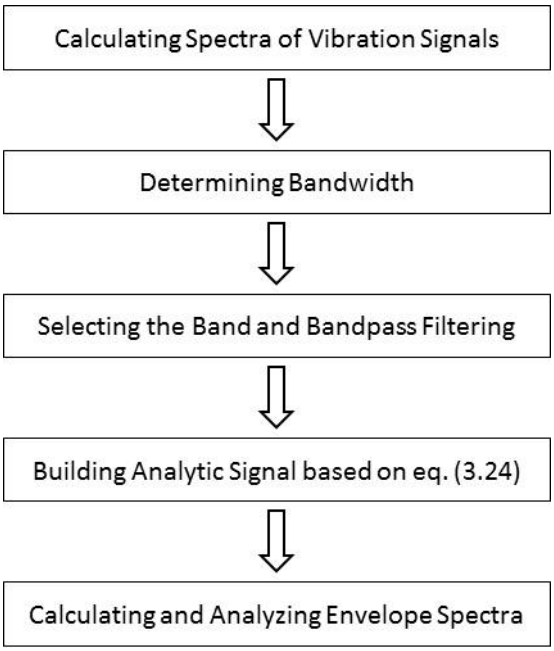


Figure 3.7 Procedure of envelope analysis

3.3 The spectra of envelope signals

McFadden and Smith developed a single point fault vibration model [54] and a multiple point fault vibration model [55] for rolling element bearings. In these models, the features of the spectrum of the envelope signal are explained in terms of the bearing inner race fault. Spectral features from the bearing outer race fault and the rolling element fault was obtained by analogy.

3.3.1 Spectra of the envelope signals from a single point fault

When bearing has a single point inner race fault, a train of weighted impacts acts as an input to excite the bearing assembly and its housing structure. It can be regarded as a combination of load distribution effect and a train of impacts with uniform amplitude. The distribution of the load around the circumference of a rolling element bearing under radial load is shown in Fig. 3.8.

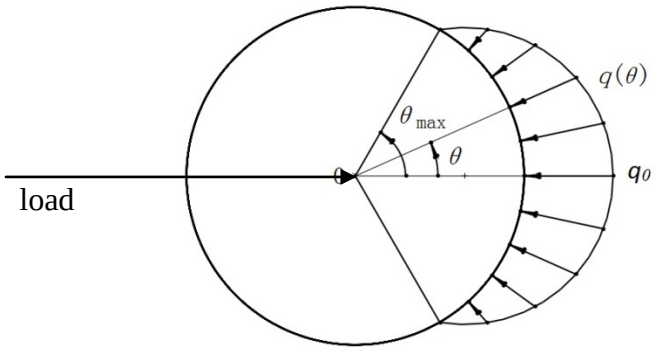


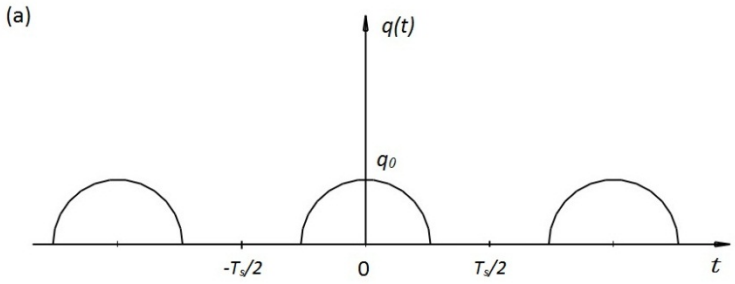
Figure 3.8 The bearing load distribution under a radial load

Consider that the outer race is fixed and the inner race is rotating at a constant shaft frequency f_s. The instantaneous load at a point on the inner race as a function of time can be represented by [54]

$$q(t) = \begin{cases} q_0 \left[1 - \left(\frac{1}{2\varepsilon}\right)(1 - \cos\theta)\right]^n & |\theta| < \theta_{max} \\ 0 & |\theta| \geq \theta_{max} \end{cases}, \tag{3.25}$$

where q_0 is the maximum load intensity, ε is the load distribution factor, $\theta = 2\pi f_s t$, θ_{max} is the angular extent of the load zone, and $n = \frac{3}{2}$ for ball bearings and $\frac{10}{9}$ for roller

bearings [7]. It is assumed that the point on the inner race is located at $\theta = 0$ at $t = 0$. For any other point, a non-zero phase at time 0 can be used. Function $q(t)$ is periodic with a period of $T_s = \frac{1}{f_s}$, since a given point on the inner race passes through the load zone at the shaft frequency. The load function $q(t)$ is shown graphically in Fig. 3.9 (a) as a series of load distribution curves extending to infinity in both negative and positive time directions. Fig. 3.9 (b) shows the spectrum $Q(f)$ of $q(t)$. Because $q(t)$ is periodic with the period T_s, $Q(f)$ is discrete with spacing f_s between adjacent spectral lines. Further, $q(t)$ is defined as real and even, so $Q(f)$ is real and even too and the phase of $Q(f)$ is zero for all f.



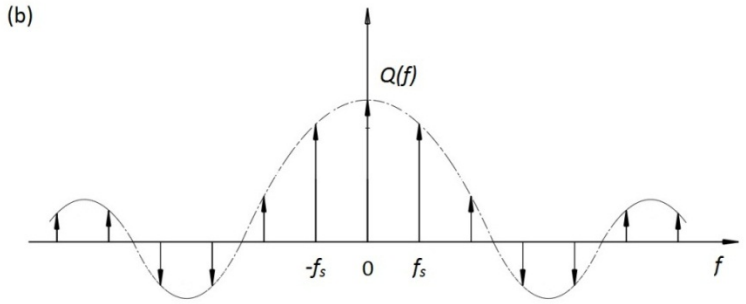


Figure 3.9 The instantaneous load at a point on the inner race (a) waveform (b) spectrum

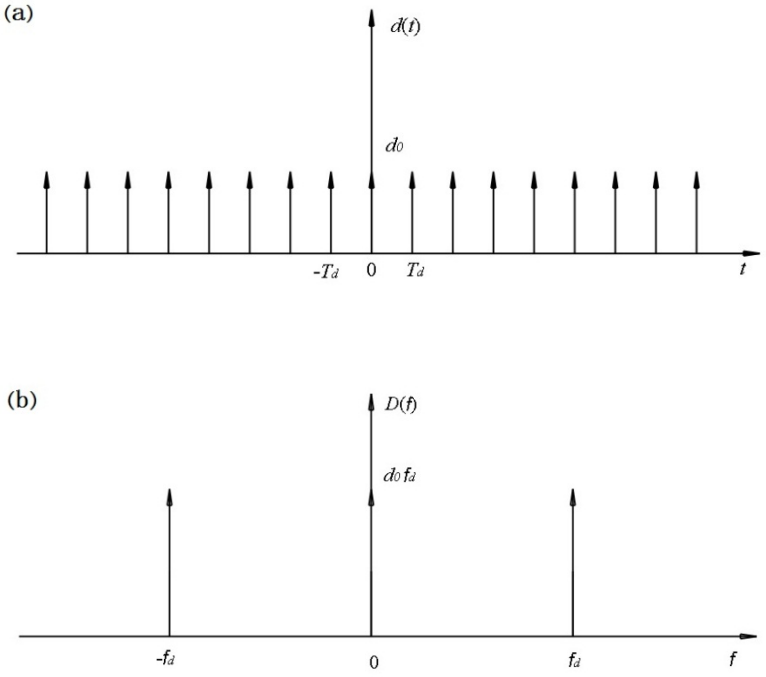


Figure 3.10 Impacts produced by a bearing fault (a) waveform (b) spectrum

Impacts produced by an inner race fault can be described by

$$d(t) = d_0 \sum_{k=-\infty}^{\infty} \delta(t - kT_d), \tag{3.26}$$

where d_0 is the impact amplitude.

Take the Fourier transform of $d(t)$ to yield:

$$D(f) = d_0 f_d \sum_{k=-\infty}^{\infty} \delta(f - kf_d), \tag{3.27}$$

where $T_d = \frac{1}{f_d}$, f_d is ball-passing-inner-race frequency, that is, $f_d = f_{bpfi}$. Referring to Fig. 3.8, let the moment when an impact strike position $\theta = 0$ be $t = 0$. Fig. 3.10 shows function $d(t)$ in time domain and $D(f)$ in frequency domain. Because $d(t)$ is discrete and periodic with period T_d, $D(f)$ is discrete and periodic too with period f_d, extending to infinity in both negative and positive frequency directions. $d(t)$ as defined here is real and even, therefore $D(f)$ is real and even with zero phase for all f.

The impacts produced by the inner race fault is the multiplication of a uniform amplitude impact train function $d(t)$ with the load function $q(t)$. Hence, the input that excites bearing response is $u(t) = q(t)\, d(t)$. From eq. (3.27), we can calculate the Fourier Transform of the fault induced input function $u(t)$ as:

$$\begin{aligned} U(f) &= D(f) * Q(f) \\ &= Q(f) * d_0 f_d \sum_{k=-\infty}^{\infty} \delta(f - kf_d) \\ &= d_0 f_d \sum_{k=-\infty}^{\infty} Q(f - kf_d) \end{aligned} \tag{3.28}$$

As shown in Fig. 3.11 (a), $u(t)$ is an amplitude modulated function. Since fault frequency is always higher than the shaft rotational frequency, $d(t)$ represents the carrier while $q(t)$ is the modulating function. Fig. 3.11 (b) shows the spectrum of $u(t)$. We can see that the modulating signal $Q(f)$ is translated and centered at the multiples of f_d in

frequency domain. Again, $u(t)$ and $U(f)$ are real and even, and $U(f)$ has zero phase for all f.

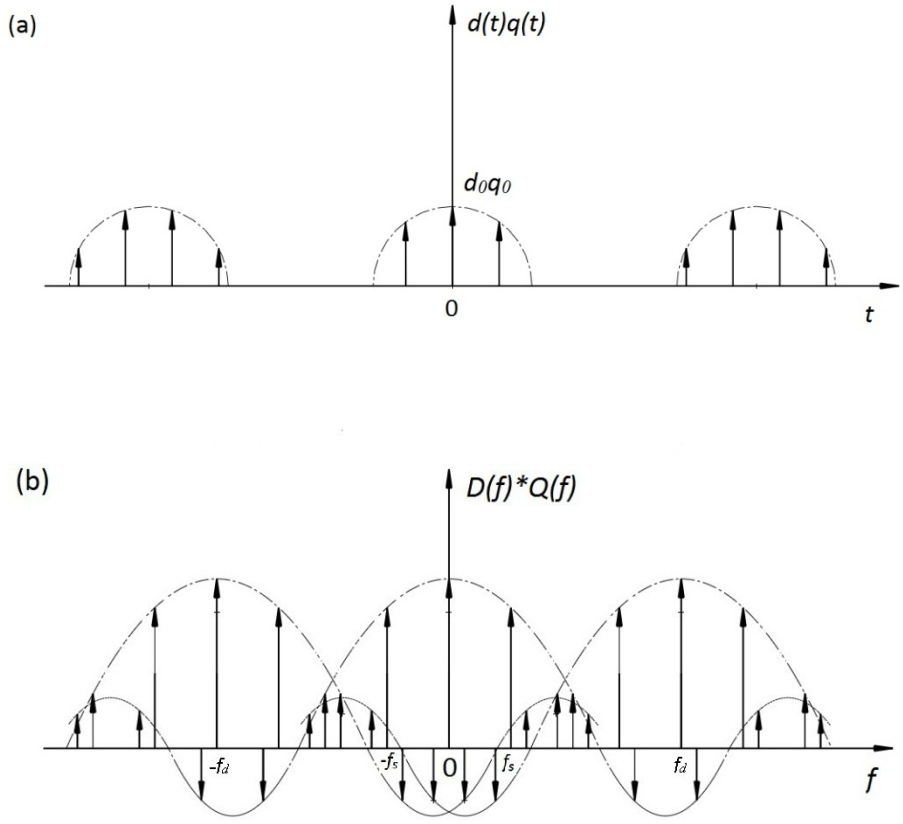


Figure 3.11 Fault induced impact function as an amplitude modulation of the impact train and load function (a) waveform (b) spectrum

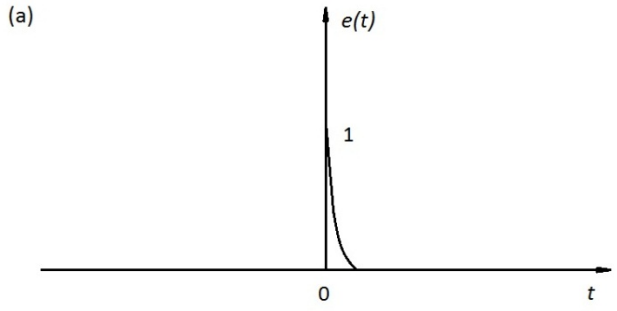
Up to this point, we have discussed the input of the mechanical system - the transmission path from the bearing fault to the vibration transducer. For clarity, let us repeat equation (2.15) here

$$x(t) = Ae^{-\frac{t}{T_e}} \sin(\omega_r t + \theta_0),$$

which is the impulse response function of a mechanical system with a single degree of freedom. The characteristics of the impulse response function can be specified for a given resonance frequency by the initial amplitude A of the response and the time constant T_e of the exponential decay. The envelope analysis removes the effect of the term $\sin(\omega_r t + \theta_0)$ in the above equation. Thereby, the envelope signal is the convolution of the input $u(t)$ with the impulse response $x(t) = Ae^{-\frac{t}{T_e}}$. For convenience, we set $A = 1$ and define

$$e(t) = \begin{cases} e^{-\frac{t}{T_e}}, & t > 0, \\ 0, & t \leq 0 \end{cases} \tag{3.29}$$

It is set to zero for negative time because we are dealing with a causal system. As shown in Fig. 3.12 (a), $e(t)$ is real but not even. Hence the amplitude and phase of its spectrum $E(f)$ are even and odd respectively, as shown in Fig. 3.12 (b). The amplitude of $E(f)$ is maximum at the origin and decreases as frequency increases.



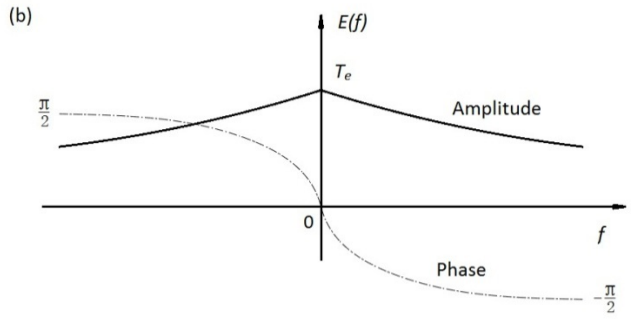
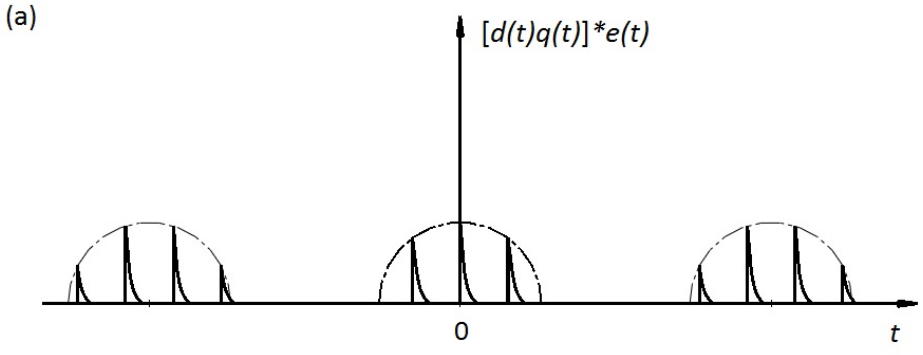


Figure 3.12 Envelope of the impulse response function (a) waveform (b) spectrum

The resultant envelope signal and its spectrum are demonstrated in Fig. 3.13. As shown in Fig. 3.13 (b), the spectrum contains a number of overlapping groups of spectral lines, centered at multiples of $f_d = f_{bpfi}$. Each group is composed of spectral lines separated by shaft rotation frequency f_s. The amplitude of the spectrum decreases with increasing frequency.



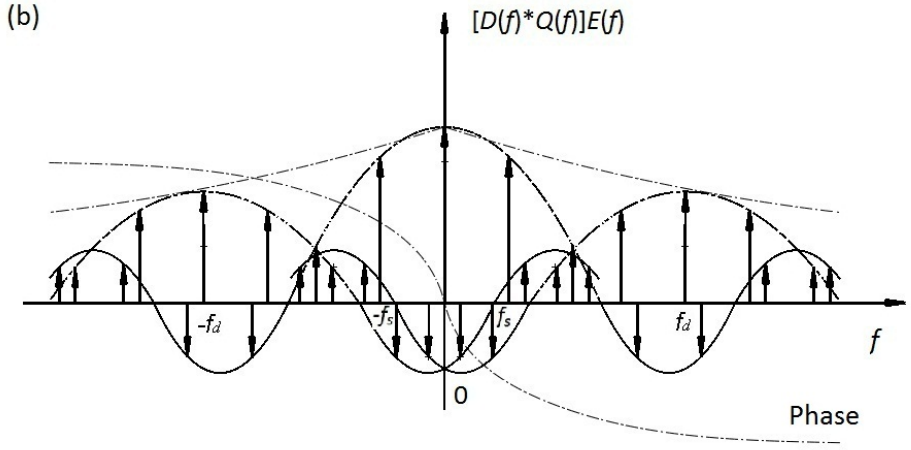


Figure 3.13 The enveloped response to the inner race fault excitation (a) waveform (b) spectrum

For a rolling element fault, the frequency in the load function $q(t)$ is the bearing cage rotation frequency f_{ftf}; the impact frequency f_d in $d(t)$ is the ball spinning frequency f_{bsf}. The spectrum of the envelope signal from the rolling element fault has the same pattern as the inner race fault, but these groups are centred at multiples of f_{bsf} and spectral lines in each group are separated by f_{ftf}. The amplitude of the spectrum decreases with increasing frequency.

Since an outer race fault is stationary and does not move during bearing operation, there is no amplitude modulation in the input. The fault induced impulse response should repeat in ball-passing-outer-race frequency $f_d = f_{bpfo}$. The spectrum of the envelope signal from an outer race fault is a train of spikes without sideband. These spikes appear at integer multiples of $f_d = f_{bpfo}$. The amplitude of the spectrum decreases with

increasing frequency. Fig. 3.14 shows typical signals and their corresponding envelope signals for faults on outer race (left); rolling element (center); and on inner race (right).

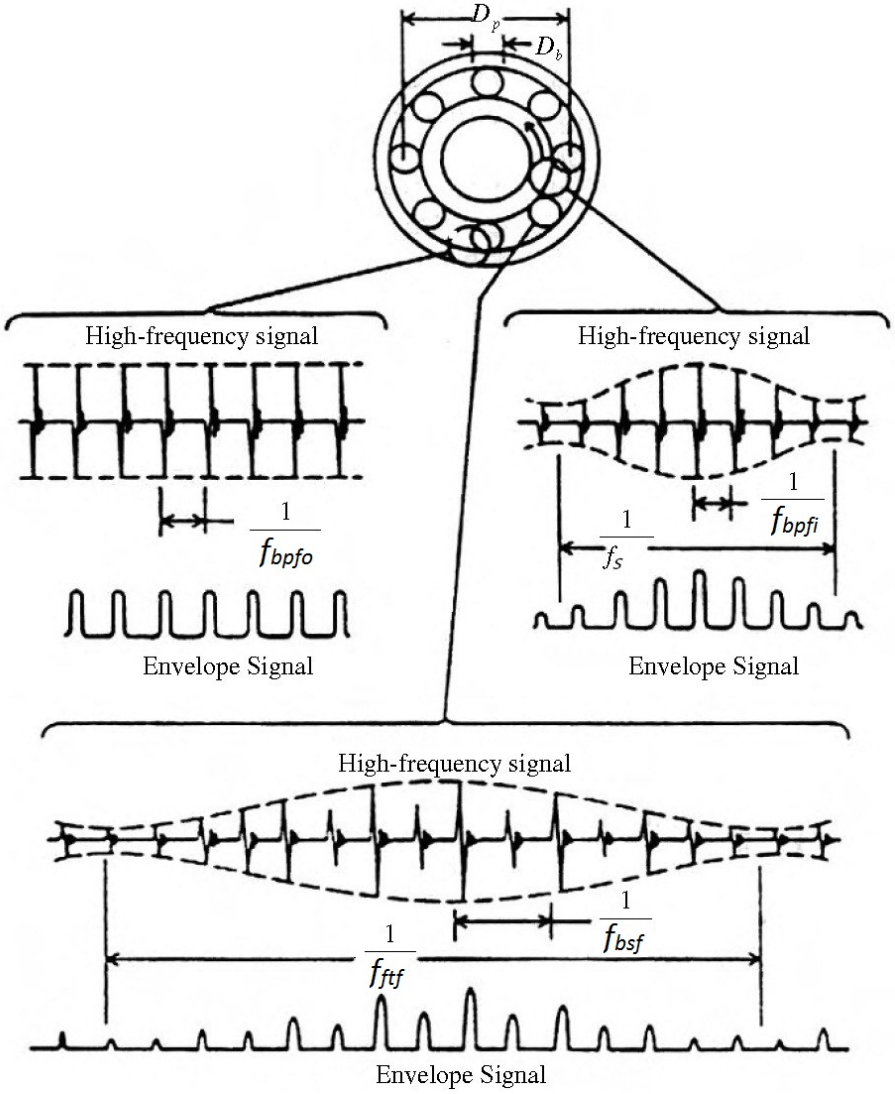


Figure 3.14 Typical signals and envelope signals from bearing faults (From [56])

3.3.2 Multiple point faults

According to [55], if a rolling element bearing has a multiple point fault on the inner race, the spectrum produced by each point fault will have spectral lines at the same frequency locations corresponding to the ball passing inner race frequency. The spectral amplitude depends on the severity of these faults. Phase is determined by the relative position of the faults on the inner raceway. The spectrum of the envelope signal is a linear superposition of the individual spectra. The spectrum contains components at the same frequencies as in the individual spectra with possibly different magnitudes, depending on the differing phases. By analogy, the same effect can be extended cases of outer race and rolling element faults.

Amplitude spectra of envelope signals with a single point and two point faults were compared in [55]. Fig. 3.15 (a) shows the amplitude spectrum of the envelope signal with a single point fault; Fig. 3.15 (b) shows the amplitude spectrum of the envelope signal with two point faults. It can be seen that the total amplitude of the spectrum with two faults is higher than that with one fault for the sake of additive effect. The spectrum with two faults has spectral lines at the same frequencies as that of a single fault, but relative magnitudes of the components have changed.

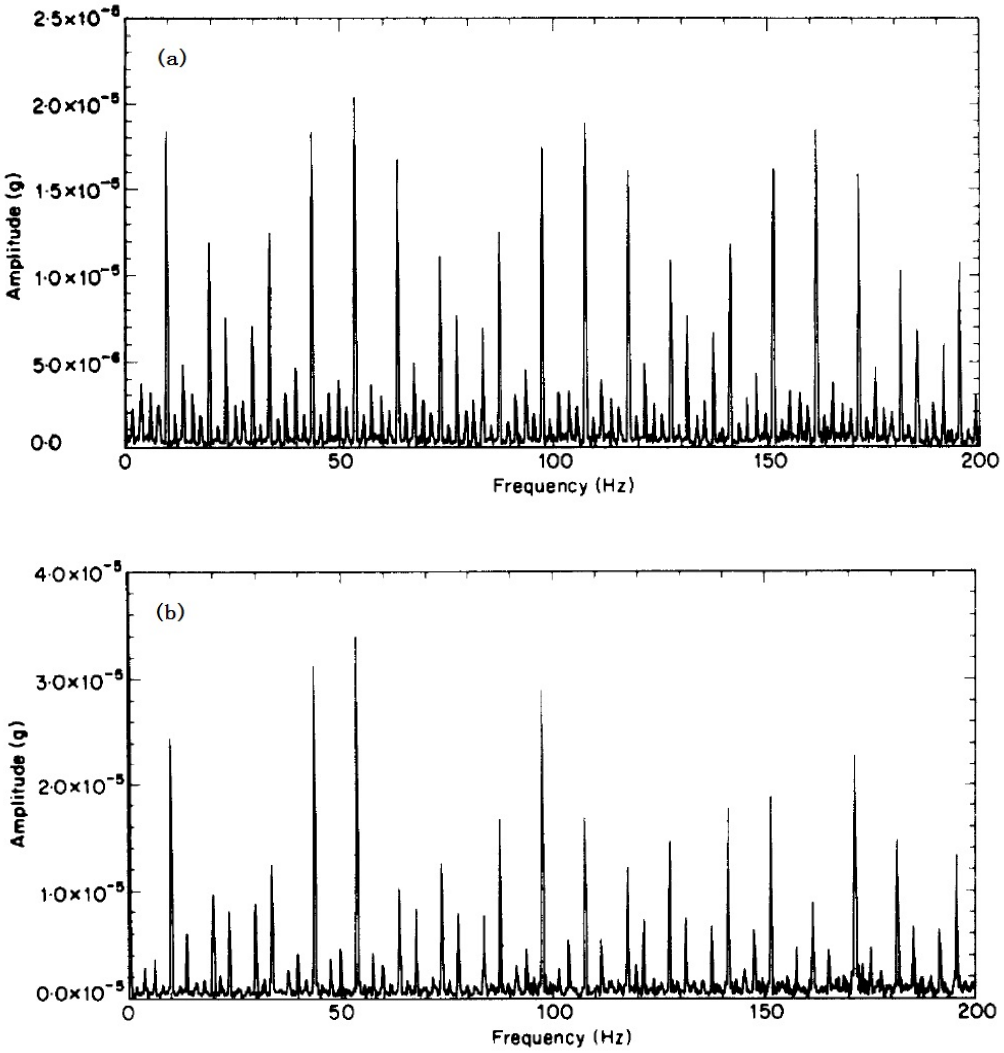


Figure 3.15 Comparison of the spectra (a) one fault (b) two faults (From [55])

3.4 Case studies

Two cases are studied here to demonstrate the procedure and effect of envelope analysis.

3.4.1 Case one: outer race fault

In this case, the shaft frequency is 29.95 Hz (1797 rpm). The motor load is 0. The sampling rate is 12 kHz. The fault is located on the outer race, so the fault frequency is the ball-passing-outer-race frequency f_{bpfo}= 107 Hz. The diameter and depth of the fault are 0.18 mm and 0.28 mm respectively. The spectrum of this signal is shown in Fig. 2.11.

As mentioned in Section 3.2.2, the resonance frequency band with the highest energy should be used for envelope analysis, and the bandwidth of this frequency band should be about 10 times the ball-passing-inner-race frequency f_{bpfi}. In this case, the frequency f_{bpfi} is equal to 161 Hz. Thus, we select the frequency band from 2.5 kHz to 4 kHz for envelope analysis. Fig. 3.16 shows the spectrum of the bandpass filtered signal.

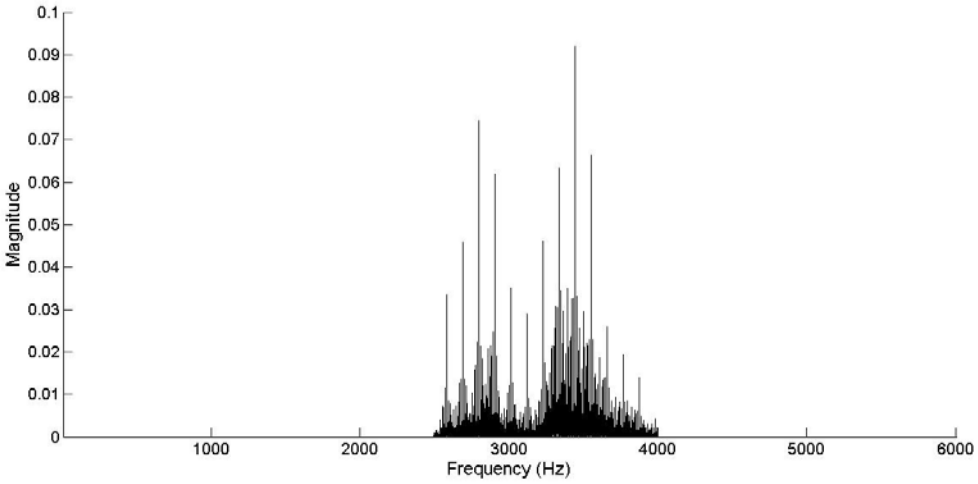


Figure 3.16 Spectrum of the bandpass filtered signal of Case one

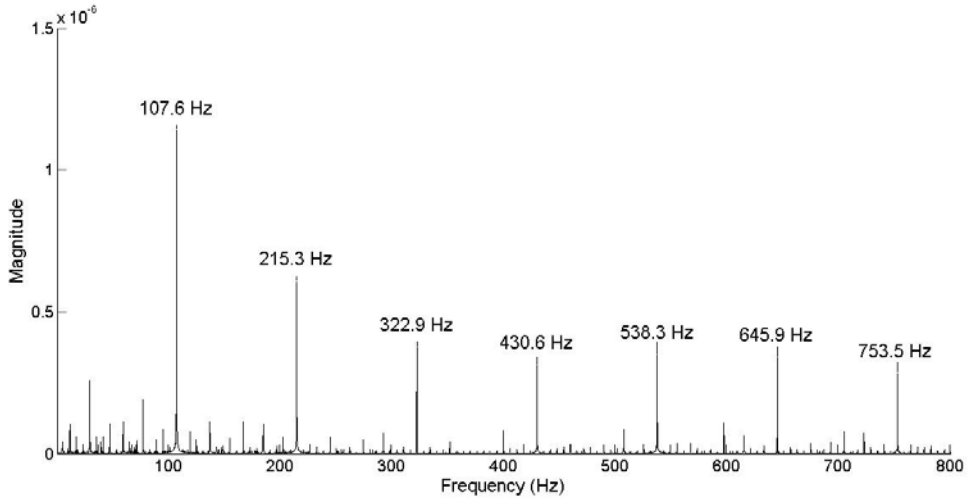


Figure 3.17 Spectrum of the envelope signal of Case one

The spectrum of the envelope signal in Fig. 3.17 includes a train of spectral lines without sidebands. These spectral lines are located at multiples of f_d=107.6 Hz which is approximately equal to f_{bpfo}. The envelope spectrum matches the spectral pattern of an outer race multiple point fault. Therefore, the diagnostic conclusion is that the bearing has outer race faults.

3.4.2 Case two: inner race fault

The vibration signal is generated by the same model of drive end bearing of the test stand introduced in Section 2.5. In this case, the bearing has a single pit on the inner race. The pit has the same size as in Case one. Shaft rotation frequency is 29.95 Hz. Motor load is 0. Sampling rate is 12 kHz. Fault frequency is ball-passing-inner-race frequency f_{bpfi} = 161 Hz. The spectrum of this signal is shown in Fig. 3.18.

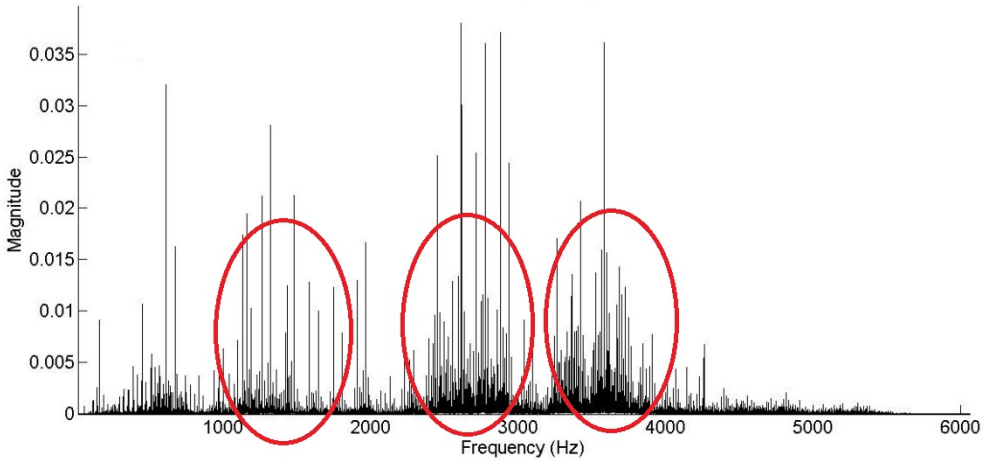


Figure 3.18 Spectrum of the vibration signal of Case two

There are three resonances circled in Fig. 3.18. Resonances with higher magnitude are better to use, thus we select the frequency band from 2.5 kHz to 4 kHz for envelope analysis. Fig. 3.19 shows the spectrum of the bandpass filtered signal.

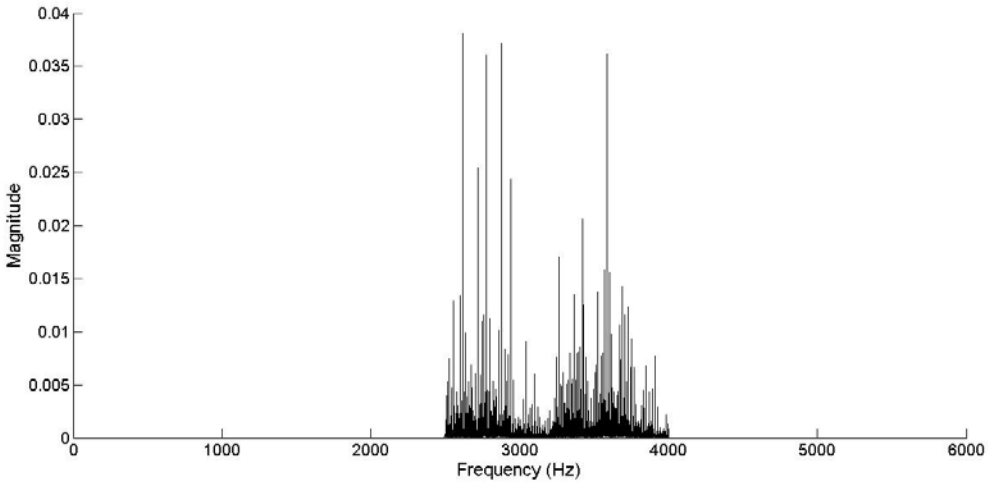


Figure 3.19 Spectrum of the bandpass filtered signal of Case two

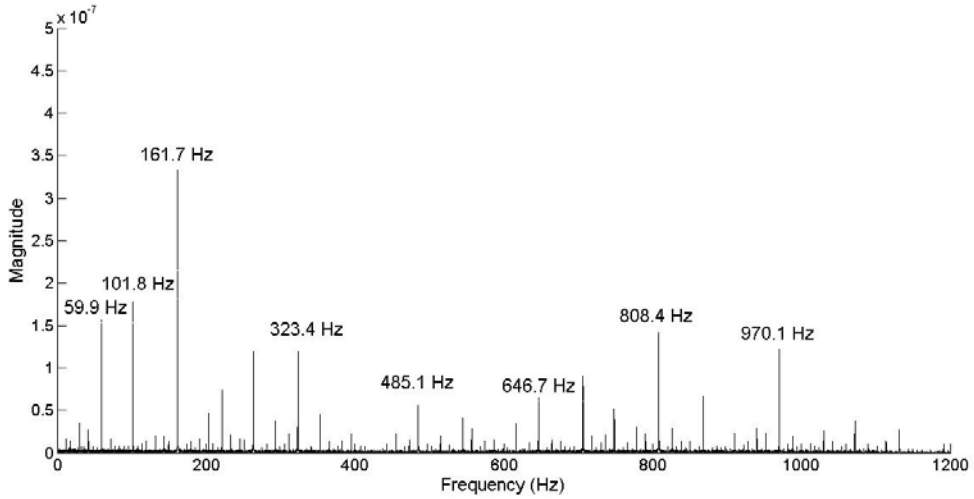


Figure 3.20 Spectrum of the envelope signal of Case two

The spectrum of the envelope signal in Fig. 3.20 shows that spectral lines appear at multiples of f_{bpfi} and there are sidebands around them. The spectrum matches spectral pattern of an inner race multiple point fault. The fault on the inner race is detected successfully.

As comparison with envelope spectra with faults, Fig. 3.21 shows the envelope spectrum of a healthy bearing. This case has same characteristic frequencies and same operation condition with Case 1 and Case 2. Fig. 2.12 shows the spectrum of the vibration signal generated by this bearing. The envelope spectrum is obtained using band [4 kHz, 4.5 kHz]. This spectrum includes two spectral lines at high order of shaft rotation frequency. The magnitude of this spectrum is much lower than that with faults.

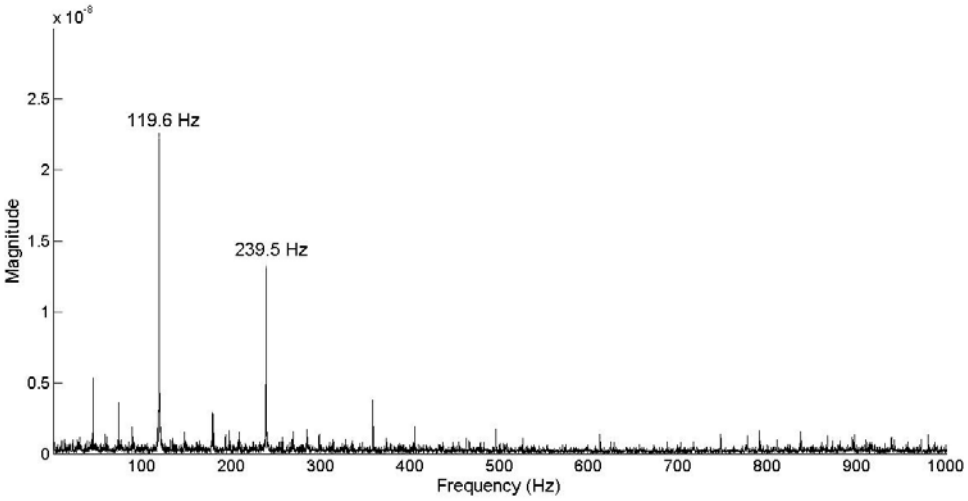


Figure 3.21 Envelope spectrum of a healthy bearing

Chapter Four: **AUTOMATIC SELECTION OF THE FREQUENCY BAND FOR ENVELOPE ANALYSIS**

Although there exist different bearing failure modes, all failures lead to localized faults. The localized faults can be regarded as the sum of a number of point defects of differing magnitude and position. From discussions in Chapter 3, we can see that envelope analysis can be very effective in diagnosing localized bearing faults. However, the frequency band for envelope analysis must be selected manually by an analyst. Although an algorithm was proposed to automatically select the frequency band by investigating the kurtosis parameter [42], this algorithm requires considerable computation.

We propose a new algorithm to determine the suitable frequency band for envelope analysis based on band energy distribution. As shown in Fig. 2.11, localized incipient bearing faults generate resonance signals with energy concentrated in narrow frequency bands. Empirically, it is recognized that bearing fault induced resonance is mostly located in frequency bands higher than 1 kHz. This algorithm is based on wavelet packet transform and Root Mean Square (RMS). Wavelet packet transform is used to decompose the residual signal into subbands, and the RMS values of these subbands are calculated to quantify energy distribution. The subband with the highest energy is considered suitable for envelope analysis.

4.1 Wavelet analysis overview

A wavelet is a waveform with limited duration in time and zero average. Fig. 4.1 shows Db20 wavelet in time domain as an example. Sinusoids extend from minus to plus

infinity. They have good frequency resolution but poor time resolution, being good at describing stationary signals. Instead, wavelets are irregular, of limited duration, and often non-symmetrical. They are better at describing anomalies, pulses, and other events that start and stop within a limited time duration [57].

Figure 4.1 Db20 wavelet (From [57])

Wavelet analysis consists of continuous wavelet transform (CWT), discrete wavelet transform (DWT) and wavelet packet transform (WPT). Wavelet analysis is mainly used for signal identification, decomposition [58, 59], compression and denoising. Considering CWT and DWT are the basis of WPT, it is beneficial to introduce CWT and DWT prior to introducing WPT.

4.1.1 Continuous wavelet transform

Continuous wavelet transform is defined as [60-62]

$$C_{\tau,s} = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) dt. \tag{4.1}$$

Function $\psi(t)$ is the mother wavelet. It must have zero mean

$$\int_{-\infty}^{\infty} \psi(t) dt = 0.$$

A wavelet dictionary is constructed from the mother wavelet which is dilated with a scale parameter s and translated by τ:

$$\left\{\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right)\right\} \tag{4.2}$$

where $\tau \in R$, $s \in R^+$. CWT is the orthogonal projection of signal $x(t)$ on basis $\{\psi_{\tau,s}(t)\}$.

CWT is mainly applied to identify non-stationary, transient and abnormal events within a signal [11, 12].

4.1.2 Discrete wavelet transform

CWT is extremely redundant and can produce tremendous amount of data. To reduce the computational cost and suppress the redundancy, DWT, which uses only those scales that are a power of 2, is used more commonly. For DWT, the wavelet dictionary is constructed with

$$\left\{\psi_{j,n}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t-2^j n}{2^j}\right)\right\}, \tag{4.3}$$

where $j \in Z$, $n \in Z$.

For any scale 2^j, $\{\psi_{j,n}(t)\}_{n\in Z}$ is an orthonormal basis. For all scales, $\{\psi_{j,n}(t)\}_{(j,n)\in Z^2}$ is an orthonormal basis of $L^2(R)$. This orthonormal basis removes all redundancy and is a subdictionary of the continuous wavelet transform dictionary.

The orthogonal projection of a signal on basis $\{\psi_{j,n}(t)\}_{(j,n)\in Z^2}$ is defined as the details of the signal. The details denotes a wavelet transform computed up to a scale 2^j, which is not a complete signal representation. To reconstruct the signal, it is necessary to add low frequency information of this signal corresponding to scales larger than 2^j. This

is obtained by introducing a scaling function $\phi(t)$, from which an orthonormal basis $\left\{\phi_{j,n}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t-2^j n}{2^j}\right)\right\}_{(j,n)\in Z^2}$ is constructed. The orthogonal projection of a signal on basis $\{\phi_{j,n}(t)\}$ is defined as the approximation of the signal.

From a different point of view, these bases are equivalent to conjugate mirror filters used in discrete multirate filter banks. Fig. 4.2 shows a two-level DWT filter bank. A signal S is convolved with the first level wavelet function, which is equivalent to

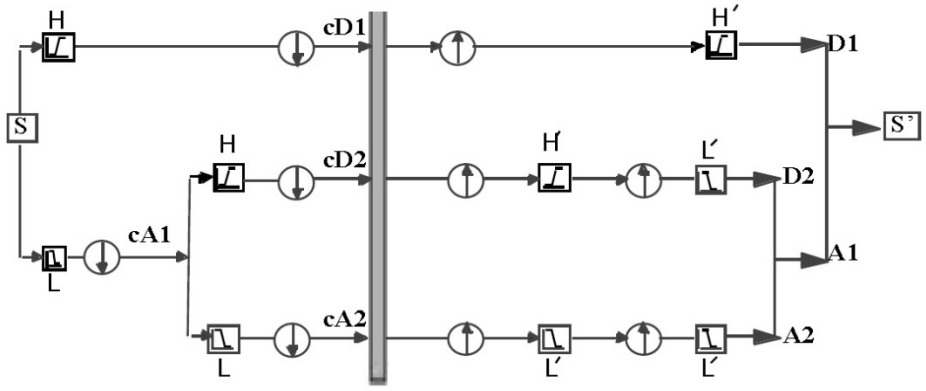


Figure 4.2 A two-level DWT filter bank (From [57])

being filtered by a highpass halfband filter H. After downsampling by factor 2, we get the first level detail of the signal cD1. The signal S is convolved with the first level scaling function, which is equivalent to be filtered by a lowpass halfband filter L. After downsampling, we obtain the first level approximation of the signal cA1. The signal S is decomposed into two subbands evenly by the first level DWT. The second level DWT divides the first level approximation further into two subbands while leaving the first

level detail untouched. The right hand side of Fig. 4.2 demonstrates a signal reconstruction process. A DWT filter bank is made of a set of filters called perfect reconstruction quadrature mirror filters. The signal reconstruction process will automatically cancel alias resulting from downsampling . Fig. 4.3 illustrates frequency allocation of a two-level DWT filter bank. At the first level of DWT decomposition, the frequency band of the original signal S is evenly split into two halves. The higher half band contains detail D1 and the lower half band contains approximate A1. The second level of DWT processes A1 just like the first level processes S. It divides A1 into A2 and D2. After two-level DWT, the original signal is decomposed into three bands – low frequency band A2 with one-quarter the bandwidth of S, middle frequency band D2 with same bandwidth as A2, and high frequency band D1 with half the bandwidth of S.

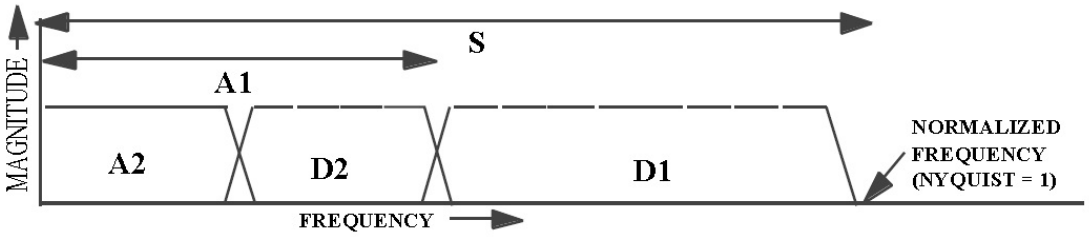


Figure 4.3 Frequency allocation of a two-level DWT filter bank (From [57])

4.1.3 Wavelet packet transform

DWT divides the frequency axis into intervals of 1 octave bandwidth. WPT generalizes this decomposition by further decomposing details of a signal into smaller frequency bands. The recursive decomposition is represented in a binary tree. A three-level WPT tree is illustrated in Fig. 4.4. An original signal S is denoted by node (0, 0) in Fig. 4.4.

The 1- level WPT is equivalent to a lowpass halfband filter and a highpass halfband filter and decomposes node (0, 0) into nodes (1, 0) and (1, 1) evenly. Likewise, nodes (1, 0) and (1, 1) are decomposed into (2, 0), (2, 1) and (2, 2), (2, 3) respectively. Thus, 2-level WPT decomposes (0, 0) into 4 subbands evenly. In summary, at each node (*j*, *k*), the frequency band is $[\frac{S}{2^j}(k-1), \frac{S}{2^j}k]$, where $k = 1, 2, \ldots, 2^j$ and j is the level index of WPT. For n level of WPT, the original signal is decomposed into 2^n bandwidth-equal frequency bands.

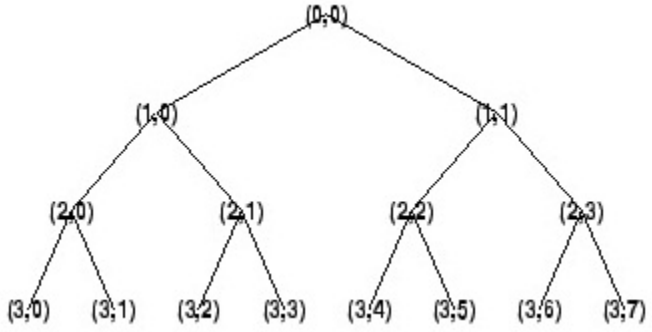


Figure 4.4 Three-level of binary wavelet packet tree

4.2 Frequency band selection for envelope analysis

The result of envelope analysis heavily depends on the selected frequency band. In the previous chapter, we showed the envelope spectrum of a signal with an outer race fault (Fig. 3.17). The envelope analysis was performed between 2.5 kHz and 4 kHz. This was manually chosen based on the observation of resonance from the spectrum in Fig. 2.11. Hence, envelope analysis based on this choice of band was successful in fault detection.

In contrast, if one uses 0 – 1 kHz band to conduct envelope analysis, results are shown in Fig. 4.5. Obviously, envelope analysis is not effective.

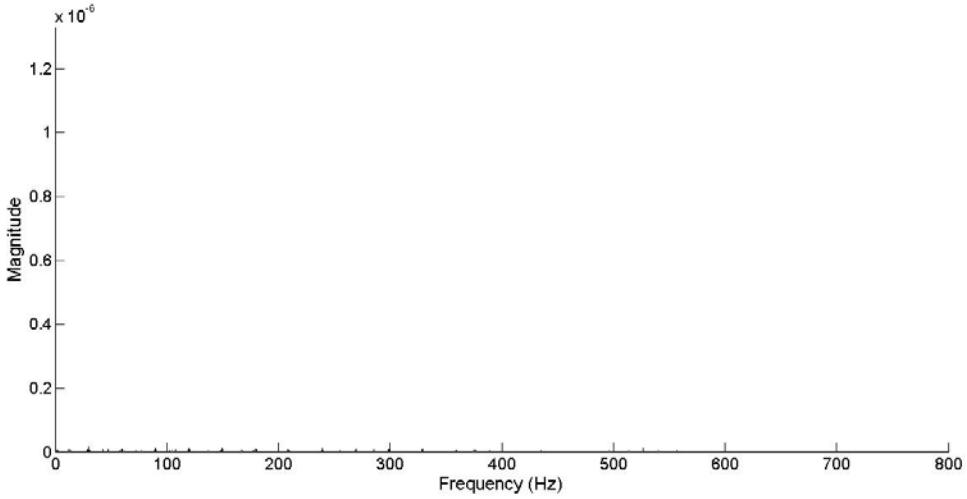


Figure 4.5 Envelope spectrum using band 0-1 kHz

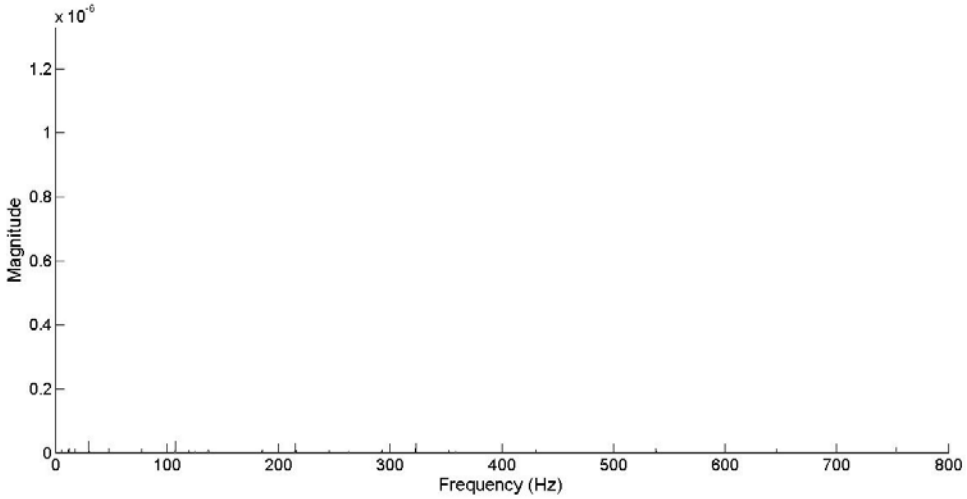


Figure 4.6 Envelope spectrum using band 1-2.5 kHz

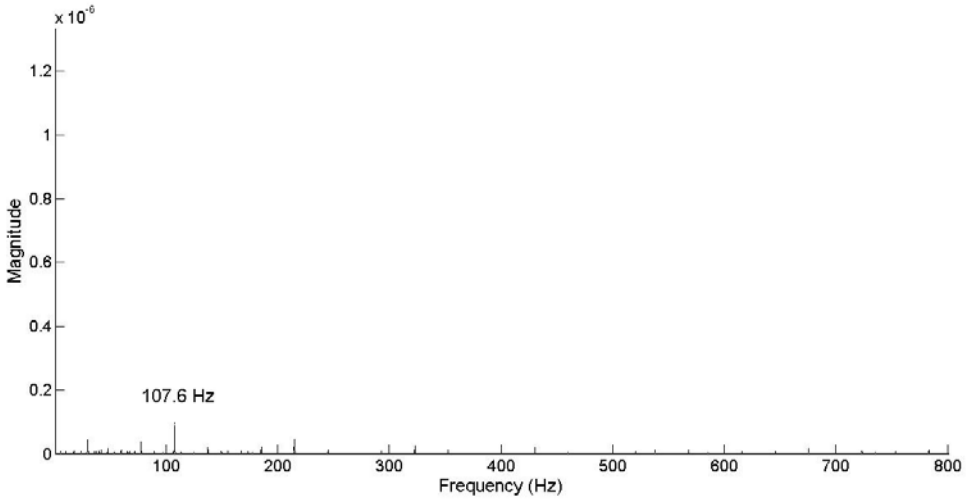


Figure 4.7 Envelope spectrum using band 4-6 kHz

We show two more envelope spectrums with the unsuitable band choices. Fig. 4.6 and Fig. 4.7 are the results of band 1 – 2.5 kHz and 4 – 6 kHz respectively. There is no spectral line in Fig. 4.6. Although there is one spectral line at f_{bpfo} in Fig. 4.7, its magnitude is very small and there are no harmonics.

To improve the effectiveness of envelope analysis, determination of a suitable band is crucial. As shown in Fig. 1.1, envelope analysis deals with residual signals obtained after order tracking and TSA, etc. Residual signals are dominated by bearing signals. The choice of the right band is guided by the appearance of resonance. Since high signal energy is a direct indication of resonance, we propose to search for frequency bands that contain high energy. Wavelet packet transform provides the advantages we needed for the purpose of automatic resonance search. It is efficient and accurate especially dealing with non-stationary signals. We design an algorithm of automatic band selection which

consists of WPT decomposition and subband recombination based on Root Mean Square. The procedure is shown in Fig. 4.8.

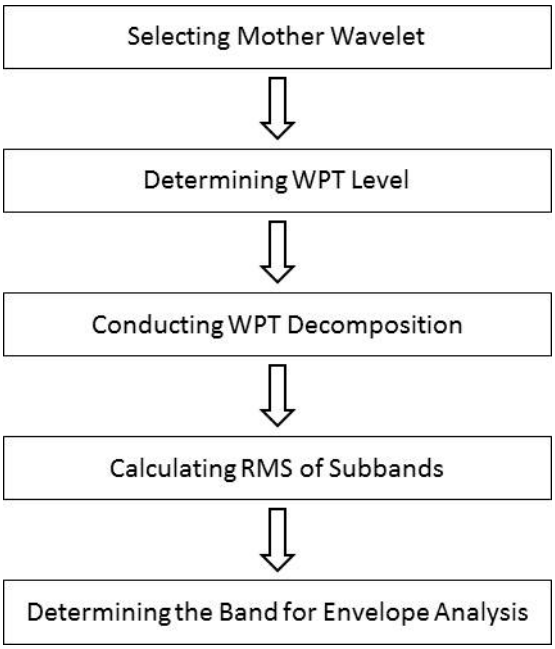


Figure 4.8 Procedure of automatic band selection

First, WPT decomposition is conducted. For WPT decomposition, two factors should be determined – mother wavelet and WPT level. Determination of WPT level is based on the Nyquist frequency i.e. the half of sampling rate, of the residual signal and bandwidth selection requirement for envelope analysis discussed in Section 3.2.2, i.e. approximate 10 times of f_{bpfi}. Assume the Nyquist frequency of the residual signal to be *f*. The *n*th level of WPT decomposes the frequency band of the residual signal into 2^n subbands with a bandwidth $\frac{f}{2^n}$. The bandwidth should be approximately 10 times f_{bpfi}. However, the *n*th level of WPT determined in this way cannot find the band with the

highest energy in most situations, considering WPT binary decomposition limitation. We propose $n + 2$ level of WPT to acquire finer decomposition.

In choosing a proper mother wavelet for signal decomposition, features such as the width and frequency selection capacity of a wavelet. A wavelet with narrow width, i.e. compact support, requires less computation than one with wider width. Good frequency selection capacity requires steep cut-off in frequency response. Figure 4.9 compares the frequency response of the discrete Meyer wavelet with Daubechies 8 wavelet. The discrete Meyer wavelet shows steeper cut-off rate.

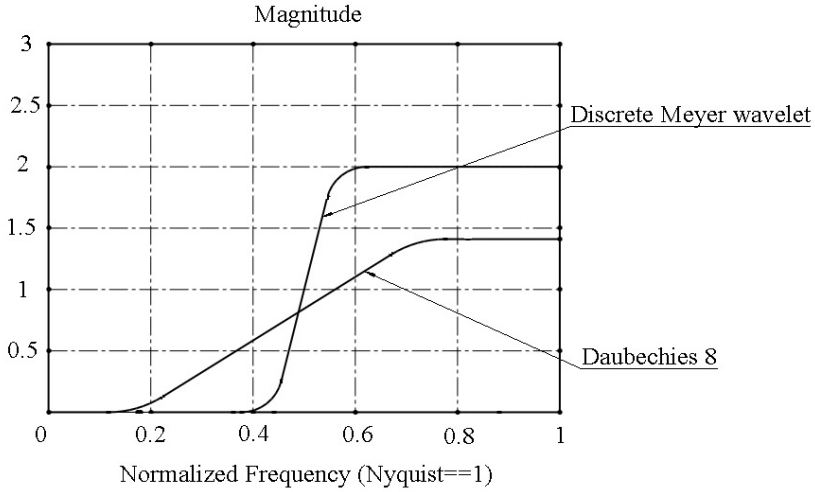


Figure 4.9 Frequency responses of two wavelet functions

In addition, discrete Meyer wavelet has equal bandwidth of 0.4 for pass band and stop band. This is good for a halfband filter. Further, magnitude of pass band of discrete Meyer wavelet function is higher than that of Daubechies 8. Therefore, the discrete Meyer wavelet possesses much better frequency selection capacity. Daubechies 8 is 8

points long while the discrete Meyer wavelet is 62 points long. WPT based on the discrete Meyer wavelet needs more computation. However, the level of WPT is not more than 10 in terms of bearing diagnosis normally. Thus, the computation is acceptable.

Fig. 4.10 shows a spectrum of a bandpass filtered signal with pass band between 1 kHz and 5 kHz, and 14 kHz Nyquist frequency. The signal is decomposed by 3-level WPT with both discrete Meyer wavelet and Daubechies 8 wavelet filter bank. The RMS values are calculated of each band. The results are compared and shown in Fig. 4.11.

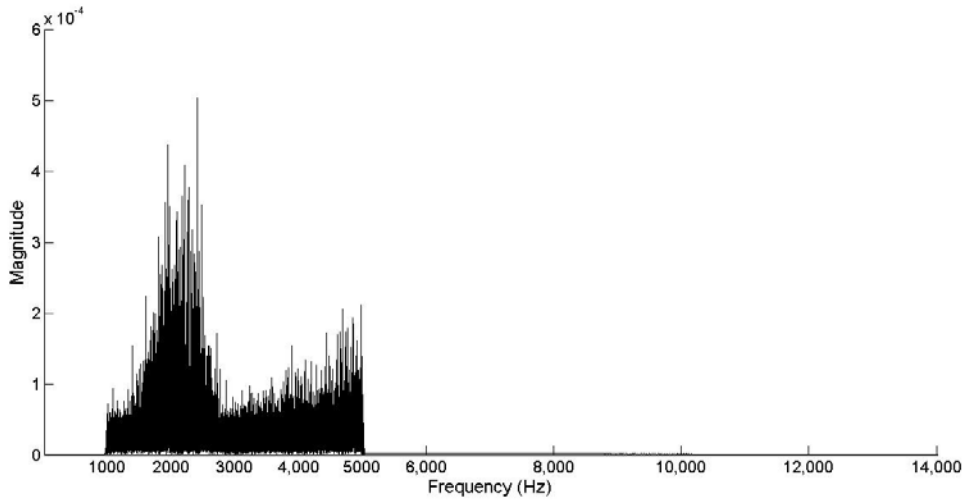


Figure 4.10 Spectrum of a bandpass filtered signal

As shown in Fig. 4.11, RMS values obtained by Daubechies 8 wavelet filter bank at the band [5.25 kHz, 7 kHz] and the band [7 kHz, 8.75 kHz] are much higher than those obtained by discrete Meyer wavelet filter bank. These two bands are in the range of stop band. This comparison indicates a better frequency selection capacity for discrete Meyer wavelet filter bank.

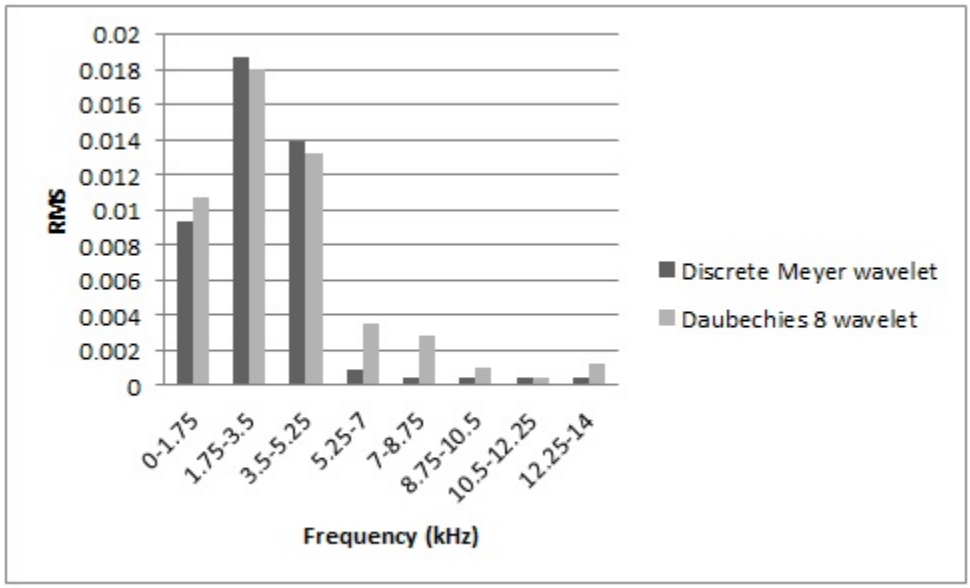


Figure 4.11 Comparison of energy distribution with different wavelet filter banks

After WPT decomposition, Root Mean Square (RMS) values of subbands are calculated to obtain energy distribution. RMS is a statistical measure of the magnitude of a fluctuating quantity. Since the sum of squares of a time series is related to the signal energy, an RMS value quantifies the signal energy in the same physical unit as the signal itself and removes the dependence on the signal length.

$$\text{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}, \tag{4.4}$$

where x_i is the sampled time series signal, N is the number of samples and $\bar{x}$ is the mean. Then the subband with highest RMS and its adjacent four bands are recombined according to bandwidth requirement of 10 times f_{bpfi}. Those with less than 30% of the

highest RMS among the adjacent bands are not combined in order to avoid including interference and noise.

4.3 Case studies

The same two cases studied in Section 3.4 are used here to demonstrate the procedure and effect of the automatic band selection method.

4.3.1 Case one: outer race fault

In this case, bearing characteristic frequency f_{bpfo}, f_{bpfi} and Nyquist rate are 107 Hz, 161 Hz and 6 kHz respectively. Bandwidth for envelope analysis requires $10f_{bpfi}$ which is 1.6 kHz. First, we use 2-level WPT with the discrete Meyer wavelet filter bank to result in a 1.5 kHz width for each band. The energy distribution among the four bands is shown in Fig. 4.12. We observe that the highest energy is located between 3 kHz and 4.5 kHz. Thus 3 to 4.5 kHz band is selected for envelope analysis. Fig. 4.13 shows the envelope spectrum using band [3 kHz, 4.5 kHz].

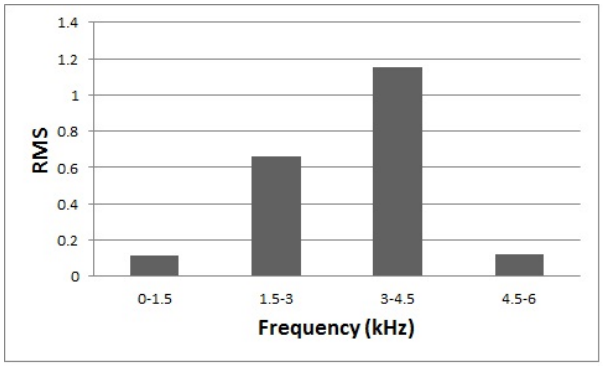


Figure 4.12 2-level WPT subband energy distribution of Case one

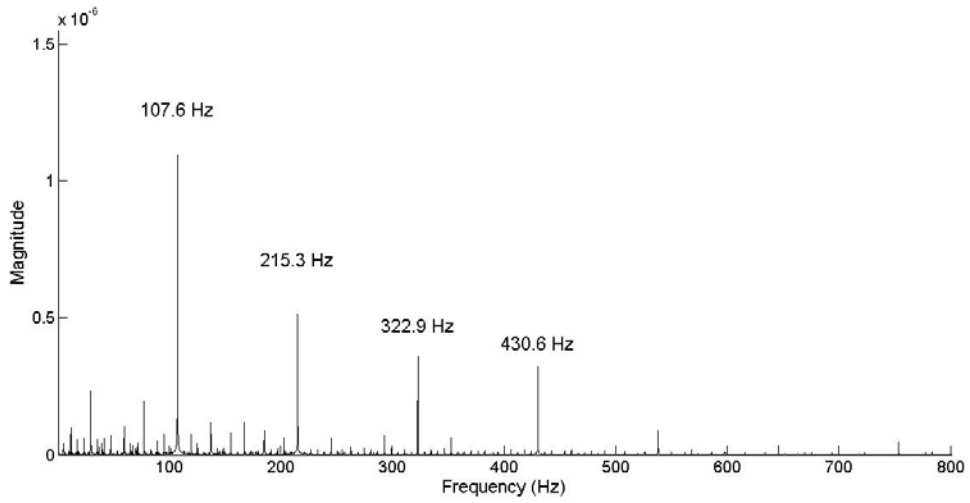


Figure 4.13 Envelope spectrum using band [3 kHz, 4.5 kHz] of Case one

Now, we use 4-level WPT to result in a 375 Hz width for each band. The energy distribution among the four bands is shown in Fig. 4.14. We observe that the highest energy is located between 3.375 kHz and 3.75 kHz. Among its four adjacent bands, bands [2.625 kHz, 3 kHz] and [3 kHz, 3.375 kHz] have more than 30% of the highest RMS. Thus, these two bands are combined with the highest energy band. The band from 2.625 to 3.75 kHz is selected for envelope analysis. Fig. 4.15 shows the envelope spectrum using band [2.625 kHz, 3.75 kHz].

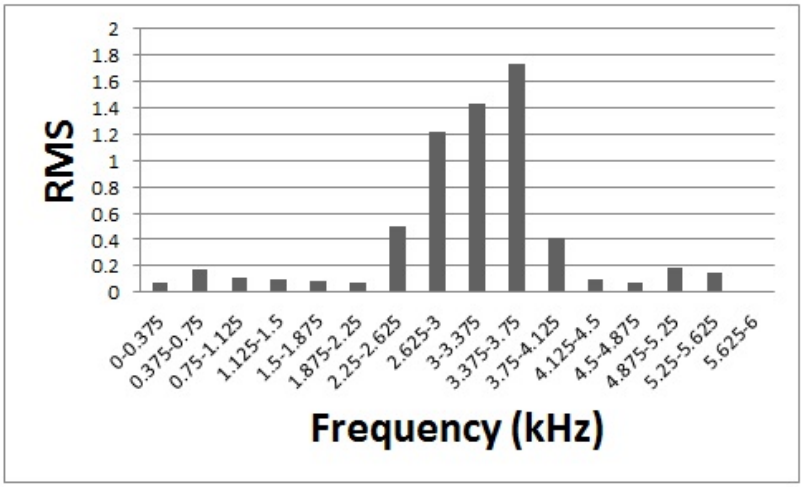


Figure 4.14 4-level WPT subband energy distribution of Case one

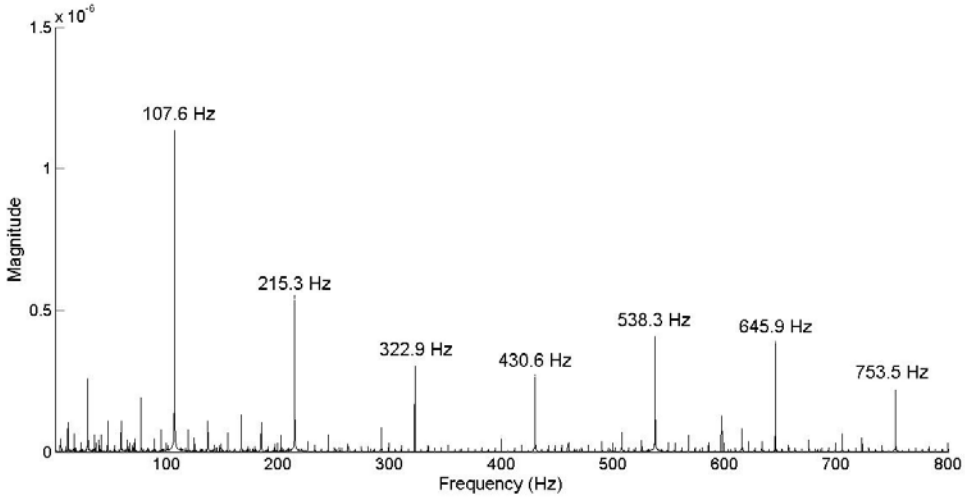


Figure 4.15 Envelope spectrum using band [2.625 kHz, 3.75 kHz] of Case one

The spectral pattern and fault frequency match those of outer race faults. The outer race fault is detected successfully. In Section 3.4.1, we observed and selected band [2.5 kHz, 4 kHz] manually for envelope analysis. The band selected automatically by 4-

level WPT is almost the same as that band. Compare Fig. 4.15 with Fig. 3.17. These two spectra are almost the same too. This proves that 4-level WPT performs very well for band selection. However, 2-level WPT selects band [3 kHz, 4.5 kHz]. This band includes only part of the resonance band, [3 kHz, 4 kHz]; band [4 kHz, 4.5 kHz] is not resonance band and may introduce interference and noise in the envelope spectrum. This is proved by the envelope spectrum in Fig. 4.13 which has fewer harmonics and slightly lower magnitude than that in Fig. 4.15. Enough magnitude and harmonics are important for diagnosis when the residual signal has much interference and noise. Further, they are particularly significant for calculating cepstrum of the envelope signal which is used for fault signature identification in Chapter 5.

4.3.2 Case two: inner race fault

In this case, the shaft rotation frequency is 29.95 Hz. The frequency f_{bpfi} and Nyquist rate are 161 Hz and 6 kHz respectively. 4-level WPT using discrete Meyer wavelet filter bank is applied. Bandwidth of each subband is 375 Hz. The energy distribution of subbands is shown in Fig. 4.16. Subband [3.375 kHz, 3.75 kHz] has the highest energy. All of its adjacent bands but [4.125 kHz, 4.5 kHz] have more than 30% of its RMS. Therefore, band [2.625 kHz, 4.125 kHz] is selected for envelope analysis.

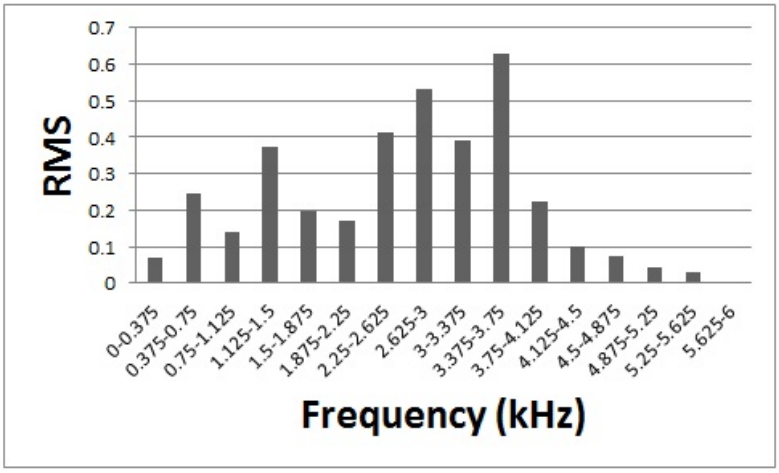


Figure 4.16 Subband energy distribution of Case two

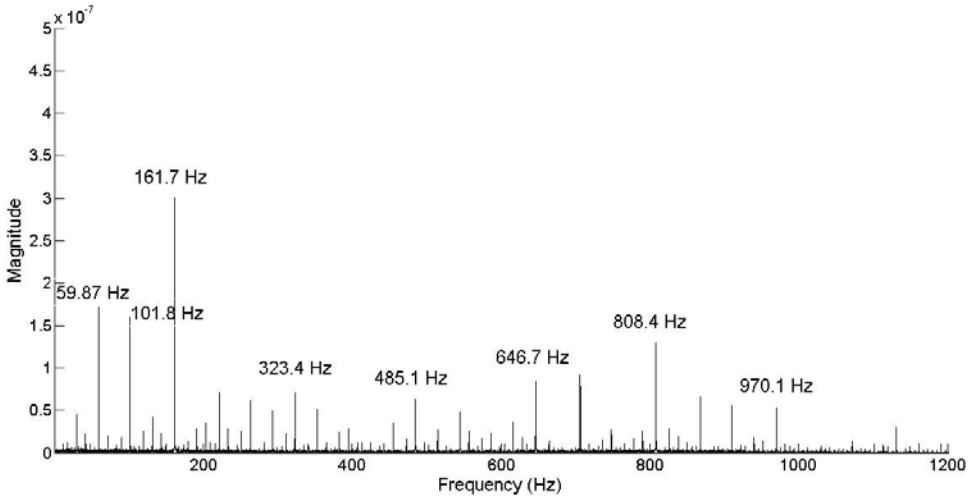


Figure 4.17 Envelope spectrum of Case two

Fig. 4.17 shows the envelope spectrum using band [2.625 kHz, 4.125 kHz]. The spectral pattern and fault frequency match those of inner race faults. The inner race fault is detected successfully. The spectrum has same number of harmonics of slightly lower

magnitude than the envelope spectrum using band [2.5 kHz, 4 kHz] shown in Fig. 3.20. It should be noted that an excessively wide band is not good for envelope analysis. When a machine has gears and operates under variable speed, order tracking and TSA is likely to be performed first. These manipulations inevitably introduce interference and noise in residual signals. An excessively wide band gives more chance for interference and noise to be shown in the residual signals, weakening the subject signal of envelope analysis.

Chapter Five: **AUTOMATIC IDENTIFICATION OF BEARING FAULTS**

After envelope analysis, the spectrum of the envelope signal is normally processed by a knowledgeable vibration analyst to determine whether a bearing has faults and the type and severity of each fault. This is accomplished by correlating the spectral pattern with bearing characteristic frequencies. In order to automate the entire diagnosis process, it is desirable that interpretation of the enveloped signal be automated as well. To do so, we propose to apply cepstrum analysis to the enveloped signal. Cepstrum has been reported to be successful in detecting harmonics normally buried in noisy signals. However, it is less commonly implemented for bearings. A possible reason is that bearing fault induced harmonics are not easily separated. In this thesis, we propose to apply ceptrum to the bearing signal after envelope analysis. This will allow ceptrum analysis to be more effect and enable an automatic detection of harmonics in the enveloped signals.

5.1 Cepstrum overview

Cepstrum is defined as:

$$C(\tau) = \mathcal{F}^{-1}\left[\log\left(X(f)\right)\right] \tag{5.1}$$

where

$$X(f) = \mathcal{F}[x(t)] = A(f)exp\left(j\emptyset(f)\right) \tag{5.2}$$

is the spectrum of $x(t)$. $\mathcal{F}(.)$ represents Fourier transform. $A(f)$ is magnitude and $\emptyset(f)$ is phase. Although eq. (5.1) implies a transform back to time domain, variable τ is called "quefrency" because the transformation back to the time domain is not the original signal. When $X(f)$ is a complex cepstrum as defined in eq. (5.1) is known as the ‘complex

cepstrum'. However, $C(\tau)$ is always real valued since $\log\big(A(f)\big)$ is even and $\emptyset(f)$ is odd. Hence equivalently

$$C(\tau) = \mathcal{F}^{-1}\big[\log\big(A(f)\big)\big] \tag{5.3}$$

is often preferred. It is also termed as the 'real cepstrum'.

The value in the cepstrum lies in the logarithmic conversion of the spectrum $X(f)$. It turns multiplication of two variables into addition. A bearing vibration signal in the absence of noise as shown in Fig. 5.1 can be regarded as the impulse response of a mechanical system $f_1(t)$ convolved with a train of weighted delta function $f_2(t)$ representing fault induced impact functions.

$$f_2(t) = \sum_{k=-\infty}^{\infty} a_k \delta(t - kT_d) \tag{5.4}$$

Hence the response signal is

$$x(t) = f_1(t) * f_2(t)$$

In frequency domain, this means:

$$X(f) = F_1(f)F_2(f)$$

where $F_1(f)$ and $F_2(f)$ are the Fourier transforms of $f_1(t)$ and $f_2(t)$, respectively. Therefore:

$$\log[X(f)] = \log[F_1(f)] + \log[F_2(f)]$$

$$F_2(f) = f_d \sum_{k=-\infty}^{\infty} a_k \delta(f - kf_d) \tag{5.5}$$

Here $f_d = \frac{1}{T_d}$ is the bearing fault frequency. Because of the additive components in $\log[X(f)]$, the cepstrum of the response signal contains features from $f_1(t)$ plus a train of spikes with spacing T_d representing features from $f_2(t)$. It allows the separation of

forcing function of bearing faults from the transfer function of mechanical structure. By detecting f_d in the cepstrum waveform and correlating it with bearing characteristics frequencies, bearing fault can be identified.

Figure 5.1 A bearing signal without noise

5.2 Automatic identification method of bearing fault signatures

To automate bearing fault diagnosis, we have first achieved automatic amplitude demodulation as described in Chapter 4. We now set up a window in the envelope spectrum. It spans the range of 0 to $(1.05)f_{bpfi}$ and it covers all bearing characteristics frequencies including ball-passing-outer-race frequency f_{bpfo}, ball-passing-inner-race frequency f_{bpfi} and ball spinning frequency f_{bsf}. By identifying the highest peak in this window and matching it with bearing characteristics frequency, the fault type may be determined. This is now demonstrated using the example shown in Fig. 4.15. In this case, f_{bpfi} =161 Hz. A window ranging from 0 – 169 Hz is applied to the envelope spectrum,

as shown in Fig.5.2. Outer race fault is identified by the highest peak at 107.6 Hz which matches $f_{bpfo} = 107$ Hz.

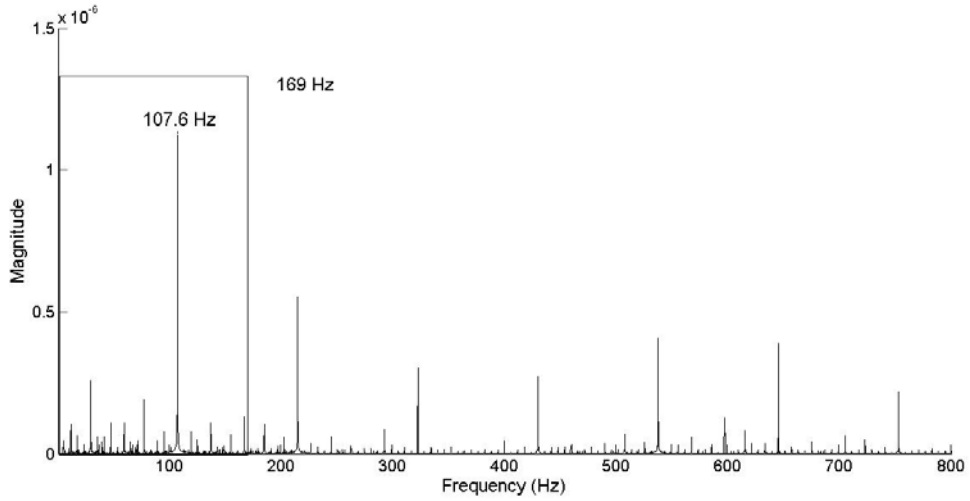


Figure 5.2 Windowed envelope spectrum from an outer race fault

This simple method works well for outer race fault cases for which the enveloped spectrum is already a train of spikes with the highest magnitude located at the fundamental frequency. However, it is less effective in detecting inner race or rolling element faults. This is due to the existence of sidebands around the spectral lines at multiples of f_d. Even worse, sidebands may possess higher magnitude than the spectral line at f_d , as shown [55] in McFadden's multiple point fault vibration model.

We now propose the use of cepstrum on enveloped signals in order to promote the significance of the forcing function associated with bearing faults. The reason is that cepstrum can just detect a harmonic family passing through zero frequency. Hence, spectral lines at multiples of f_d can be detected in envelope spectra while sidebands

centered at these spectral lines are removed out. Sidebands may pass through zero frequency also in some situations and cepstrum does detect them, however they are located at multiples of T. The frequency $f_s = \frac{1}{T}$ is shaft rotation frequency. Altogether, cepstrum of envelope signals avoids interference of sidebands as in the envelope spectra. Therefore, cepstrum of envelope signals is effective in suppressing sideband influence.

For automatic identification, we search for the fault related peak in a time window that spans to include all bearing characteristics periods. It ranges from $(0.95)\frac{1}{f_{bpfi}}$ to $(1.05)\frac{1}{f_{bpfo}}$. This is because $f_{bpfo} < f_{bsf} < f_{bpfi} < f_{bpfi.}$ The first spike is used to extract bearing fault signatures. It is located at $\tau = T_d$, so the bearing fault frequency is $f_d = \frac{1}{T_d} = \frac{1}{\tau}$. The time window is applied to the cepstrum of the envelope signal and it includes only the first spike of the cepstrum. The peak in this time window is detected and the bearing fault diagnosis can be concluded automatically.

5.3 Case study

The same two cases in Section 4.3 are studied here to demonstrate the automatic detection method of bearing fault signatures. Envelope signals obtained by the automatic band selection method are used to calculate cepstra for these two cases.

5.3.1 Case one: outer race fault

This case is an outer race fault with f_{bpfo}= 107 Hz and f_{bpfi} = 161 Hz. The spectrum of the envelope signal is shown in Fig. 4.15. A time window ranging from 5.9 – 9.8 ms is applied to the cepstrum of the envelope signal.

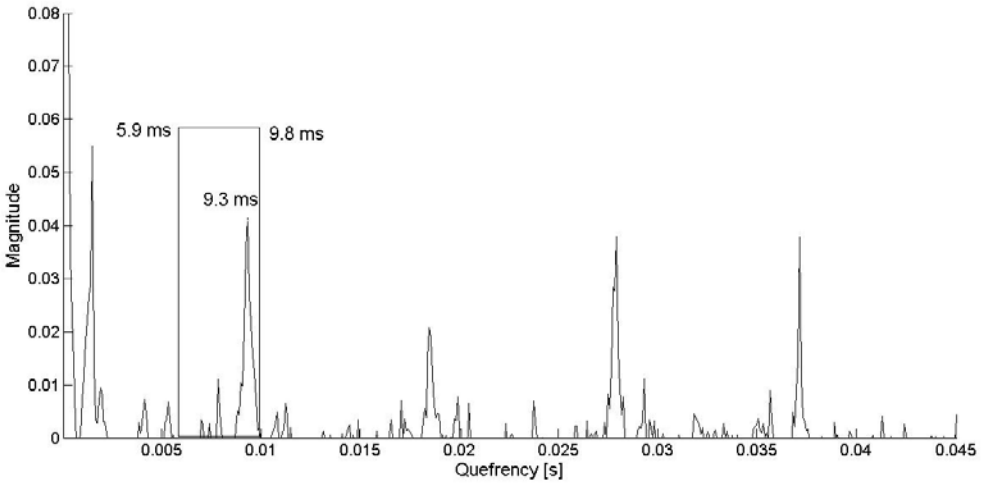


Figure 5.3 Cepstrum of the envelope signal of Case one

As shown in Figure 5.3, the highest peak is located at $t = 9.3$ ms. The bearing fault frequency $f_d = \frac{1}{t} = \frac{1}{9.3\ ms} = 107.5\ Hz$. The outer race fault is detected successfully by this method.

5.3.2 Case two: inner race fault

This case is an inner race fault with f_{bpfo}= 107 Hz and f_{bpfi} = 161 Hz. The spectrum of the envelope signal is shown in Fig. 4.17. A time window ranging from 5.9 – 9.8 ms is applied to the cepstrum of the envelope signal.

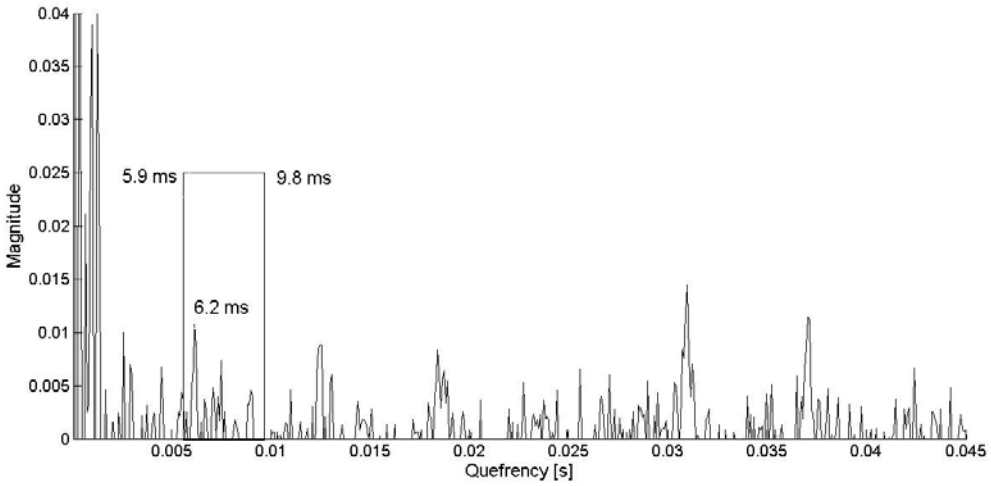


Figure 5.4 Cepstrum of the envelope signal of Case two

As shown in Fig. 5.4, the highest peak is located at $t = 6.2$ ms. The bearing fault frequency $f_d = \frac{1}{t} = \frac{1}{6.2\ ms} = 161.3$ Hz. The inner race fault is detected successfully by this method.

From the two cases, it is noted that the first spike does not always have the highest magnitude among all spikes, but still has considerable magnitude. Other spikes may have high magnitude in one case, but may not in another. For example, the fourth spike has high magnitude in Fig. 5.3 but disappears in Fig. 5.4. That is why the first spike is selected for this method. The other reason is that the time window must include only one spike. If other spikes are selected, the time window is too large to ensure this. Take Case two as an example. If we select the second spike for this algorithm, the time window is from 11.8 – 19.6 ms. The second spike at 12.4 ms and the third spike at 18.6 ms appear in the time window at the same time. Therefore, the first spike in cepstrum of envelope signals is used for this algorithm.

Fig. 3.21 shows the envelope spectrum of a healthy bearing. Cepstrum of its envelope signal is shown in Fig. 5.5 as comparison with above cases. The frequency $f = \frac{1}{t} = \frac{1}{8.2\ ms} = 121.9$ Hz. This frequency is about 4 times of shaft frequency.

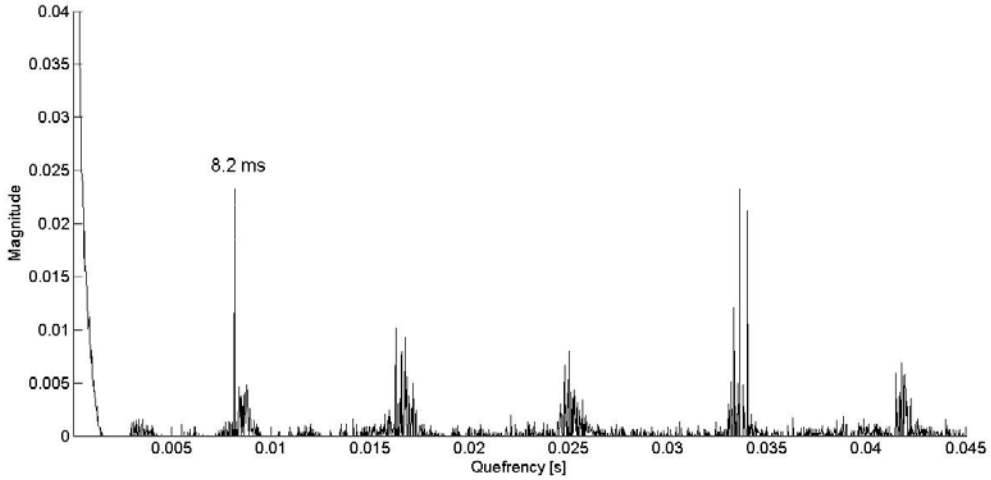


Figure 5.5 Cepstrum of the envelope signal of a healthy bearing

Chapter Six: **EXPERIMENTAL VERIFICATION**

Based on the development from Chapters 3 and 4, we construct an automatic fault diagnostic algorithm for rolling element bearing fault diagnosis – automatic envelope analysis. It combines traditional envelope analysis with the automatic frequency band selection and the automatic identification of bearing fault signatures. In this chapter, three groups of benchmark data are used to validate this technique and to compare its performance with results generated by traditional envelope analysis.

6.1 Data description

Two cases have been discussed in previous chapters. Information of them is summarized in Table 6.1.

Table 6.1 Summary of two discussed cases

	Sampling rate (kHz)	Fault size (mm)	Motor load (hp)	Shaft speed (Hz)	Fault location
Case 1	12	0.18×0.28	0	29.95	Outer race
Case 2	12	0.18×0.28	0	29.95	Inner race

Three more cases are analyzed in this chapter to investigate effectiveness of the automatic fault diagnostic algorithm under different loads, different sampling rates and different fault sizes. The three cases are based on five groups of benchmark data generated by the drive end bearing of the same test stand introduced in Section 2.5. The

bearing is a deep groove ball bearing and the model is 6205-2RS JEM SKF. Sampling duration time is around 10 second and 5 second for sampling rate 12 kHz and 48 kHz respectively. Table 6.2 lists information of the three cases.

Table 6.2 Summary of three cases

	Sampling rate (kHz)	Fault size (mm)	Motor load (hp)	Shaft speed (Hz)	Fault location
Case 3	48	0.18×0.28	0	29.95	Outer race
Case 4	12	0.18×0.28	3	28.83	Inner race
Case 5	12	0.53×0.28	0	29.95	Inner race

Case 3 has different sampling rate from Case 1. Case 4 has different motor load compared to Case 2. Case 7 has different fault size compared to Case 2. The following demonstrates validation of automatic envelope analysis through the three typical cases.

6.2 Case 3

In this case, the same bearing has the same fault and runs under the same conditions as Case 1. The difference with Case 1 is that the sampling rate is 48 kHz instead of 12 kHz. This case is used to investigate effectiveness of automatic envelope analysis under different sampling rates. Characteristic frequencies of this case are listed in Table 6.3.

Table 6.3 Characteristic frequencies of Case 3

	Shaft frequency	f_{bpfi}	f_{bpfo}	f_{ftf}	f_{bsf}
Frequency (Hz)	29.95	161	107	12	141

6.2.1 Traditional envelope analysis

Fig. 6.1 shows the spectrum of the vibration signal of this case. Fig. 6.2 shows a part of the spectrum from 0 to 6 kHz.

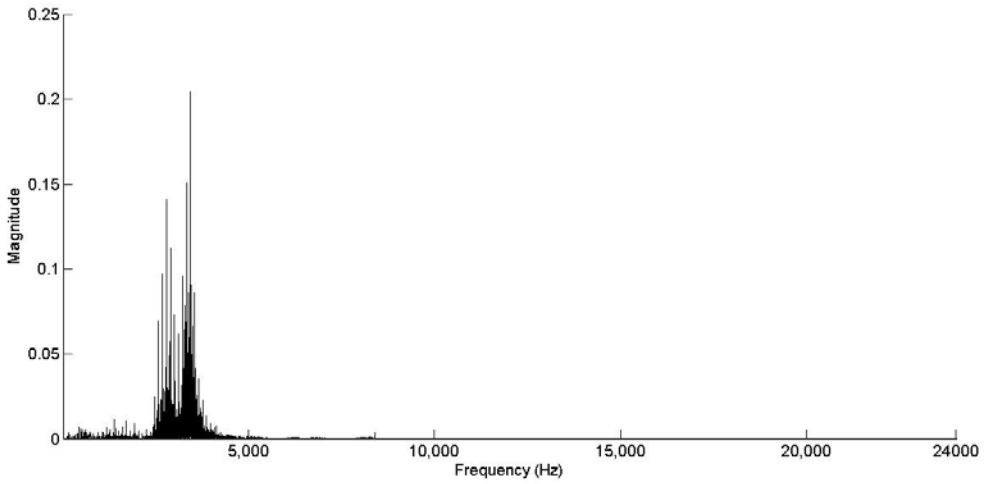


Figure 6.1 Spectrum of the vibration signal of Case 3

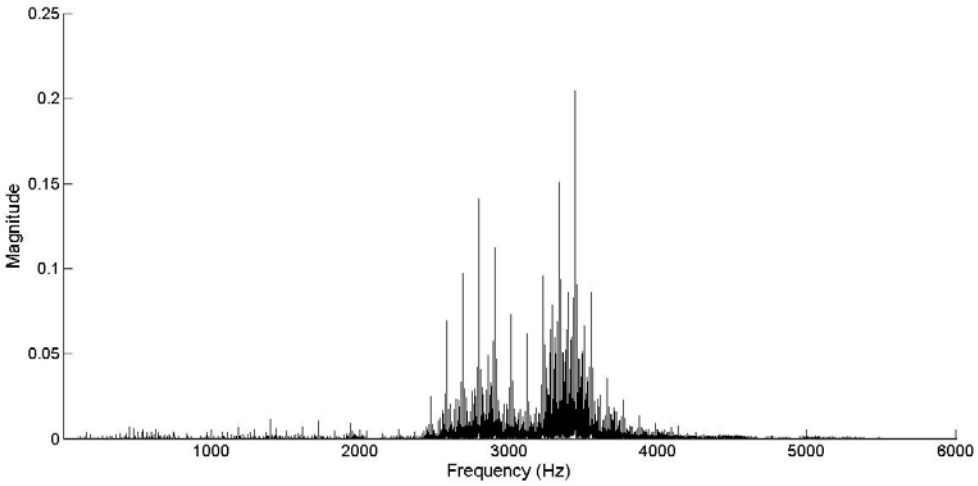


Figure 6.2 Part of the spectrum of Case 3

The frequency band of 2.5-4 kHz has the highest energy, so it is selected manually for envelope analysis just as Case 1. The envelope spectrum is shown in Fig. 6.3.

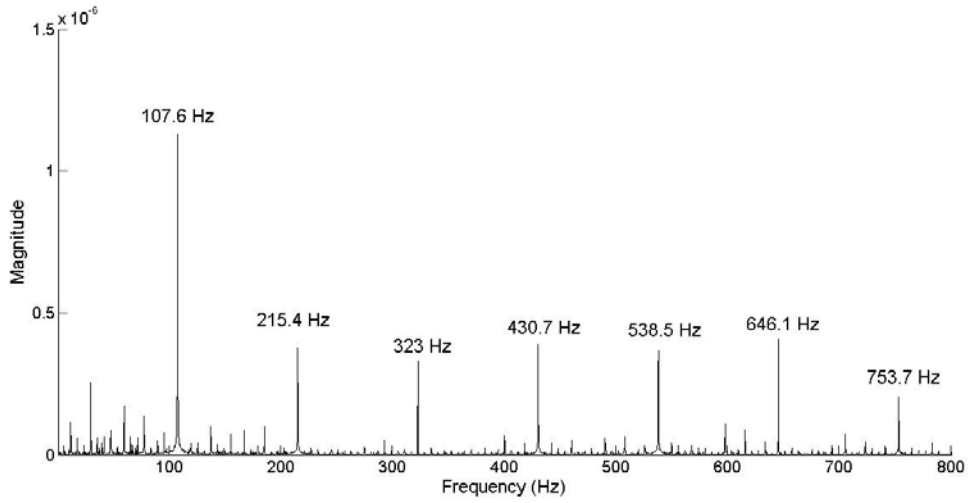


Figure 6.3 Traditional envelope spectrum of Case 3

The outer race fault is detected. The spectrum in Fig. 6.3 is same with that of Case 1 in Fig. 3.17. This proves that change of sampling rates does not affect the result of traditional envelope analysis.

6.2.2 Automatic envelope analysis

In this case, Nyquist rate is 24 kHz and 10 times of f_{bpfi} is 1.6 kHz, so 6-level WPT is used for frequency band selection. The energy distribution of subbands is shown in Fig. 6.4. The frequency band of 2.625-3.75 kHz is selected for envelope analysis.

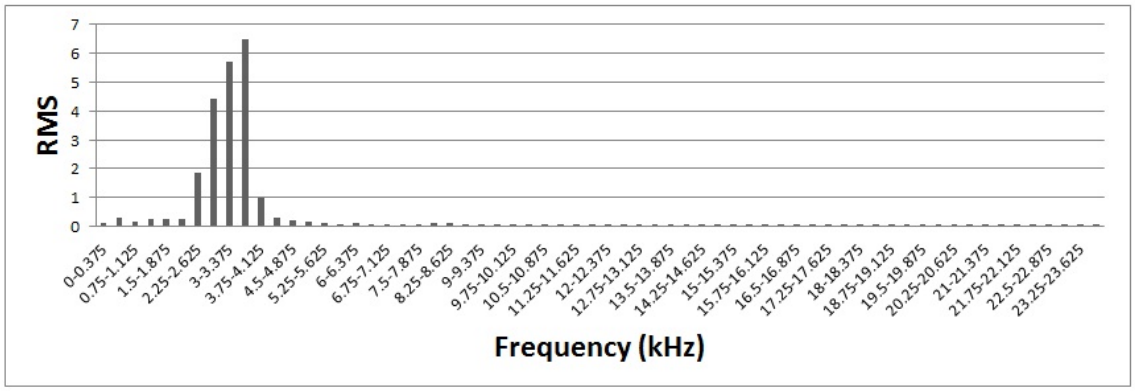


Figure 6.4 Subband energy distribution of Case 3

Fig. 6.5 shows the envelope spectrum using band [2.625 kHz, 3.75 kHz]. This spectrum is same as that in Fig. 6.3 but a little lower magnitude. This indicates that change of sampling rates does not affect automatic envelope analysis. It should be noted that the automatic envelope spectrum is shown here for comparison with the result of

traditional envelope analysis. Automatic envelope analysis does not derive this spectrum. Instead, cepstrum of envelope signals is used for fault signature identification.

Cepstrum of the envelope signal is indicated in Fig. 6.6. A time window ranging from 5.9 – 9.8 ms is applied to the cepstrum. The highest peak in this window is detected at $\tau = 9.3$ ms. The bearing fault frequency $f_d = \frac{1}{\tau} = \frac{1}{9.3\ ms} = 107.5$ Hz. The outer race fault is detected successfully.

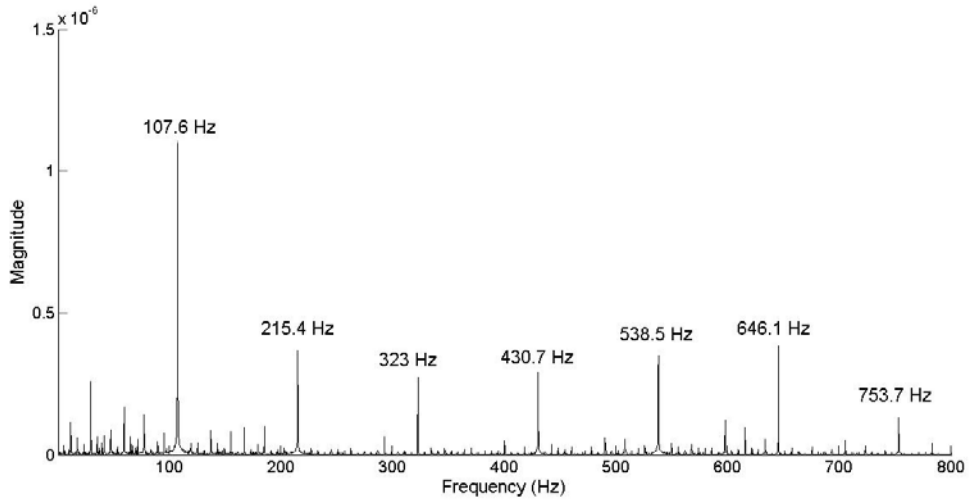


Figure 6.5 Automatic envelope spectrum of Case 3

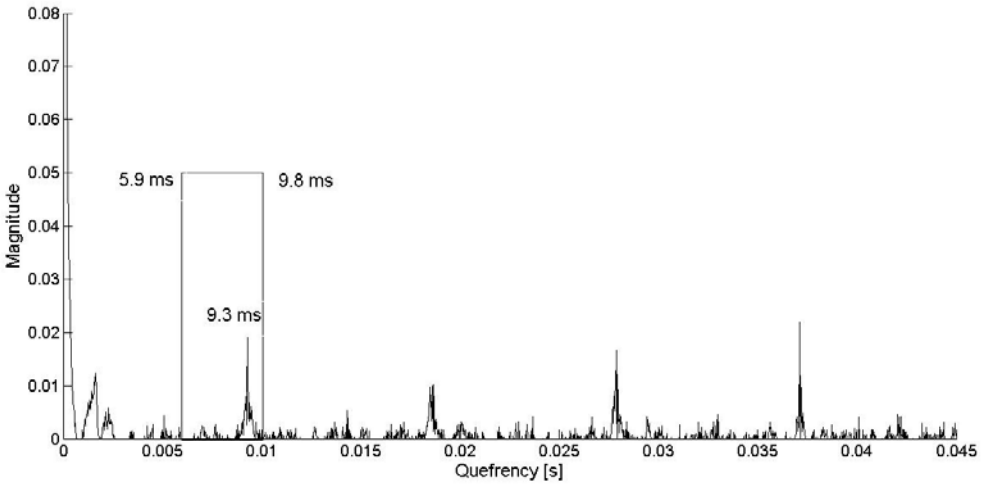


Figure 6.6 Cepstrum of Case 3

Comparing with cepstrum of Case 1 in Fig. 5.3, cepstrum of Case 3 has lower magnitude. This shows that the magnitude of cepstra of envelope signals decreases as sampling rates increase. Therefore, sampling rates may not be excessively large for automatic envelope analysis, meanwhile at least one resonance must be included.

6.3 Case 4

In this case, the same bearing has the same fault with the same sampling rate as Case 2. The main difference with Case 2 is that the motor load is 3 hp instead of 0 hp. The shaft speed is slightly different also, but it does not make much difference except changing characteristic frequencies. This case is used to investigate effect of work load on bearing fault diagnosis. Characteristic frequencies of this case are listed in Table 6.4.

Table 6.4 Characteristic frequencies of Case 4

	Shaft frequency	f_{bpfi}	f_{bpfo}	f_{ftf}	f_{bsf}
Frequency (Hz)	28.83	156	103	12	136

6.3.1 Traditional envelope analysis

Fig. 6.7 shows the spectrum of the vibration signal of this case. Comparing with Fig. 3.18, the only difference is a little higher magnitude of the spectrum. The frequency band of 2.5-4 kHz is selected manually for envelope analysis as in Case 2.

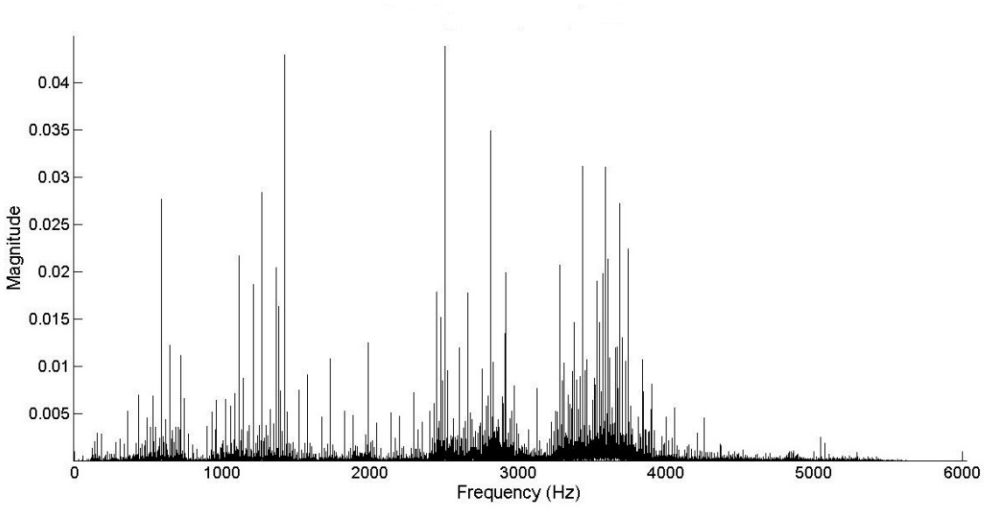


Figure 6.7 Spectrum of the vibration signal of Case 4

Fig. 6.8 shows the envelope spectrum of the signal. This spectrum matches the spectral pattern of inner race faults and the fault frequency is approximately f_{bpfi}. The inner race fault is detected successfully. This envelope spectrum has same magnitude and pattern with that of Case 2 in Fig. 3.20. The only difference is spectral line location due

to different shaft speeds. The reason is that, in this case, work load is a torque and bearing load is a vertical force. They are orthogonal, thus work load change does not affect bearing fault diagnosis. In fact, this is common for rotating machines. Therefore, envelope analysis can work well under changing work load for rotating machines.

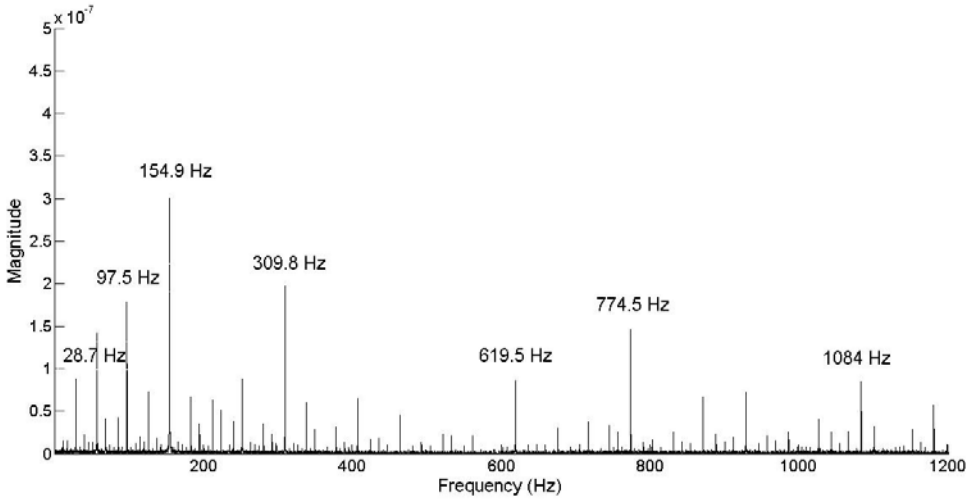


Figure 6.8 Traditional envelope spectrum of Case 4

6.3.2 Automatic envelope analysis

In this case, Nyquist frequency is 6 kHz and 10 f_{bpfi} = 1.56 kHz, so 4-level WPT is used for frequency band selection. The energy distribution of subbands is shown in Fig. 6.9. The band of 2.625-4.125 kHz is selected for envelope analysis.

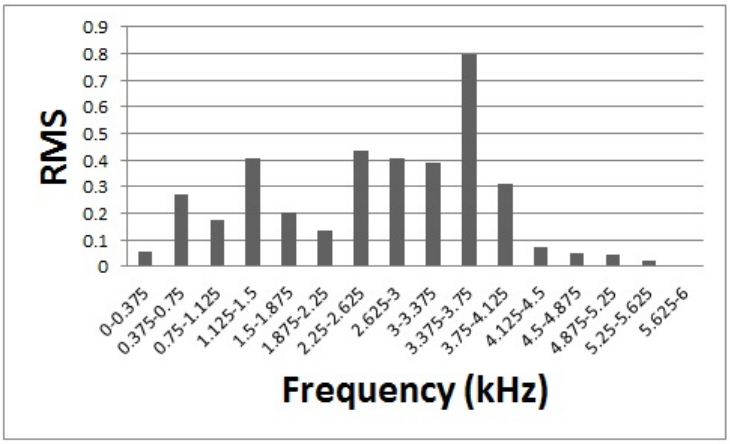


Figure 6.9 Subband energy distribution of Case 4

Fig. 6.10 shows the envelope spectrum of the signal using band [2.625 kHz, 4.125 kHz]. It is almost same with the traditional spectrum in Fig. 6.8. Cepstrum of the envelope signal is shown in Fig. 6.11.

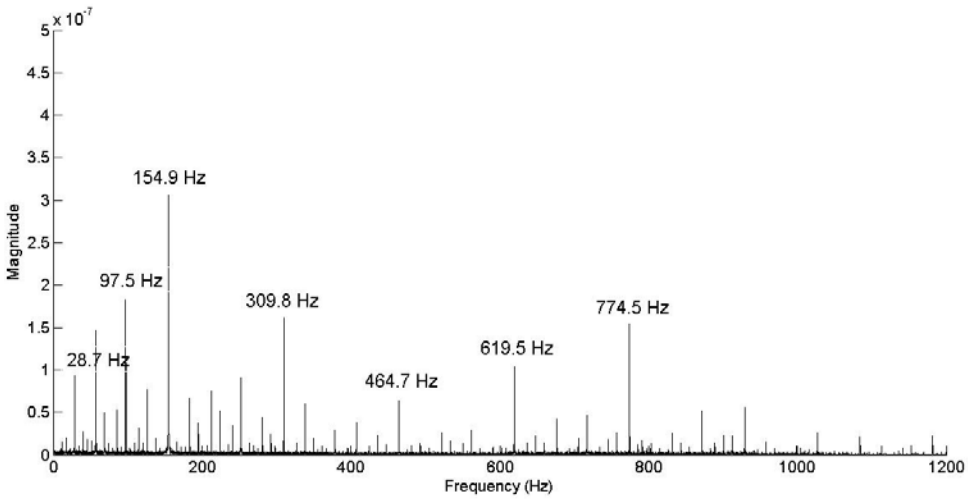


Figure 6.10 Automatic envelope spectrum of Case 4

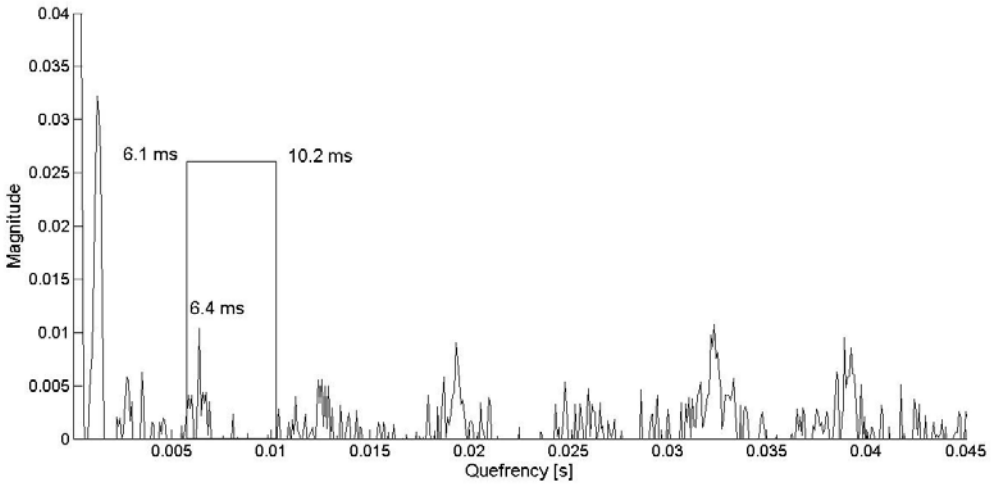


Figure 6.11 Cepstrum of Case 4

A time window ranging from 6.1 – 10.2 ms is applied to the cepstrum. The highest peak in this window is detected at $\tau = 6.4$ ms. The bearing fault frequency $f_d = \frac{1}{\tau} = \frac{1}{6.4\ ms} = 156.3\ Hz$. The inner race fault is detected successfully. Comparing with cepstrum of Case 2 in Fig. 5.4, we know that work load change does not affect automatic envelope analysis.

6.4 Case 5

In this case, the same bearing has same fault type under same operation condition and same sampling rate as Case 2. The only difference is that the fault size is bigger than Case 2. This case is used to investigate effect of fault size on automatic envelope analysis. Characteristic frequencies of this case are listed in Table 6.5.

Table 6.5 Characteristic frequencies of Case 5

	Shaft frequency	f_{bpfi}	f_{bpfo}	f_{ftf}	f_{bsf}
Frequency (Hz)	29.95	161	107	12	141

6.4.1 Traditional envelope analysis

Fig. 6.12 shows the spectrum of the vibration signal of this case. Comparing with the spectrum of Case 2 in Fig. 3.18, the spectrum of Case 5 has higher magnitude because the bigger size of fault generates heavier impacts when it contacts balls. Further, the resonance located between 1 kHz and 2 kHz in Fig. 3.17 disappears. This shows that different size of faults may excite different resonances. The frequency band of 2.5-4 kHz is selected manually for envelope analysis as in Case 2.

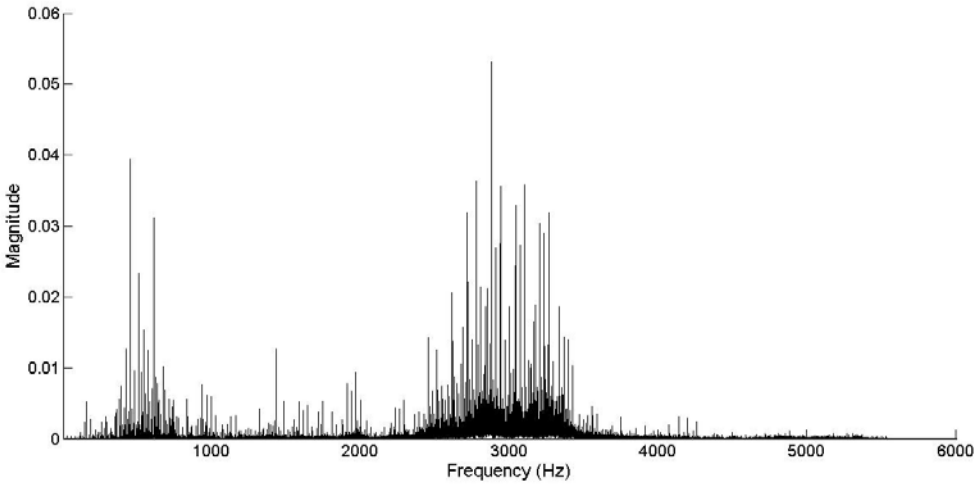


Figure 6.12 Spectrum of the vibration signal of Case 5

Fig. 6.13 shows the envelope spectrum of the signal. This spectrum matches the spectral pattern of inner race faults and the fault frequency is approximately equal to f_{bpfi}. The inner race fault is detected successfully. This envelope spectrum has higher magnitude and fewer harmonics than that of Case 2 in Fig. 3.20.

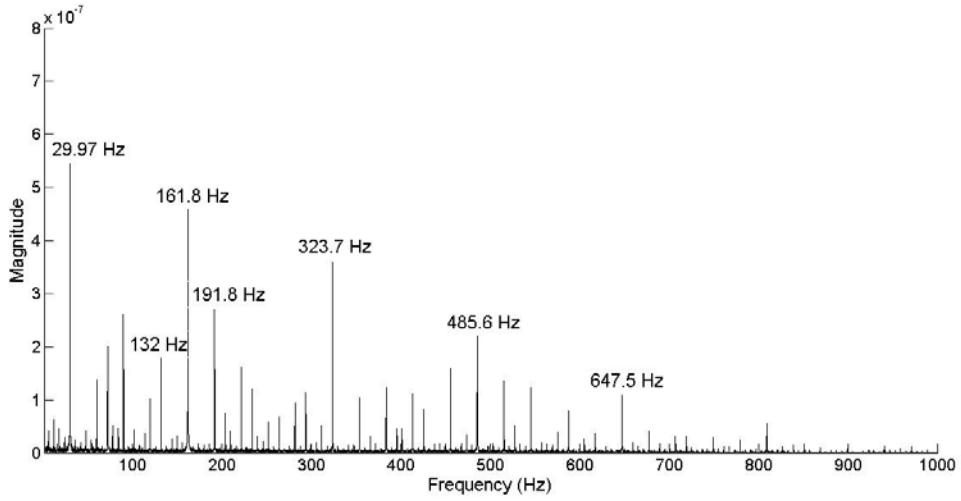


Figure 6.13 Traditional envelope spectrum of Case 5

6.4.2 Automatic envelope analysis

Nyquist rate is 6 kHz and 10 f_{bpfi} = 1.6 kHz, so 4-level WPT is used for frequency band selection. The energy distribution of subbands is shown in Fig. 6.14. The band [2.25 kHz, 3.75 kHz] is selected for envelope analysis.

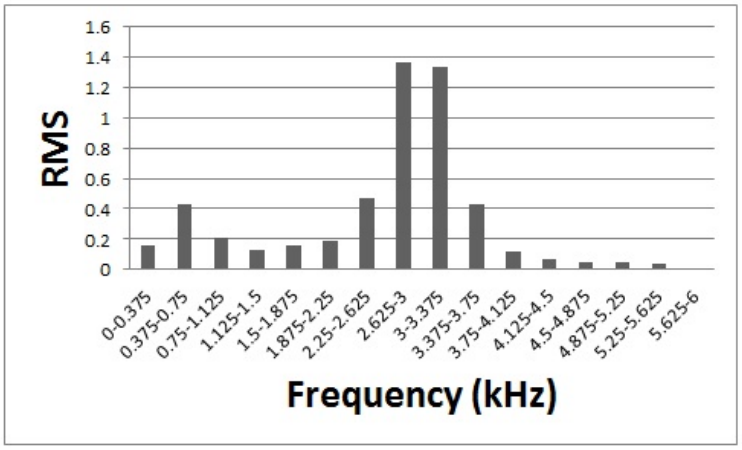


Figure 6.14 Subband energy distribution of Case 5

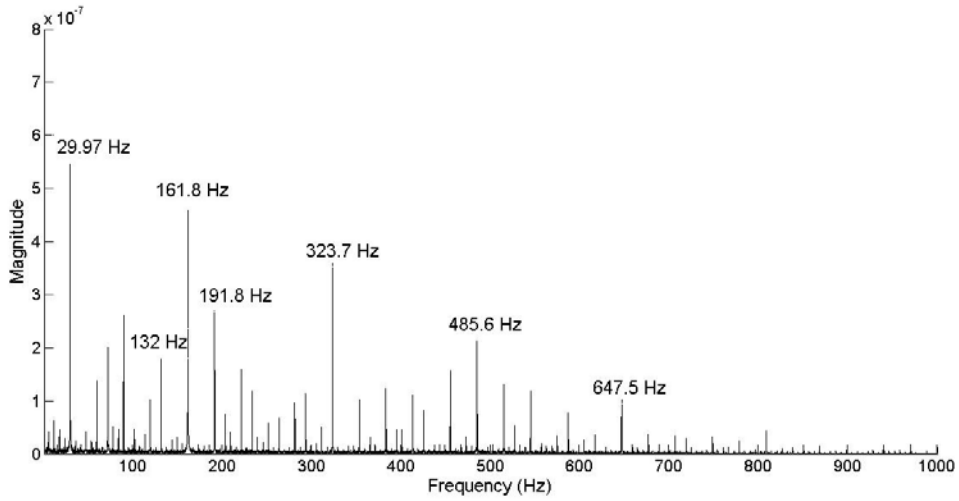


Figure 6.15 Automatic envelope spectrum of Case 5

Fig. 6.15 shows the envelope spectrum of the signal using band [22.5 kHz, 3.75 kHz]. This spectrum is same with that in Fig. 6.13.

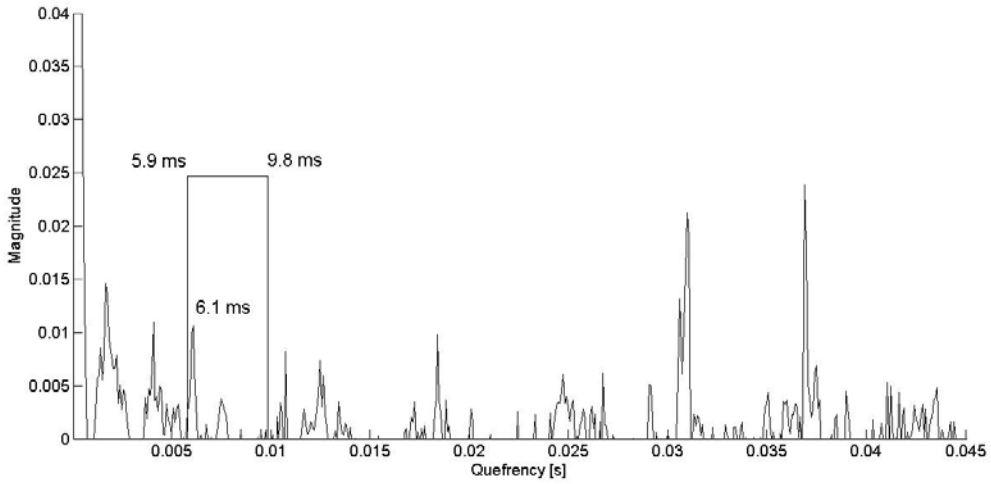


Figure 6.16 Cepstrum of Case 5

Cepstrum of the envelope signal is shown in Fig. 6.16. A time window ranging from 5.9 – 9.8 ms is applied to the cepstrum. The highest peak in this window is detected at $\tau = 6.1$ ms in this window. The bearing fault frequency $f_d = \frac{1}{\tau} = \frac{1}{6.1\ ms} = 163.9$ Hz. The inner race fault is detected successfully.

Chapter Seven: **CONCLUSIONS, DISCUSSION AND FUTURE WORK**

7.1 Conclusions

In order to improve the efficiency of fault diagnosis for rolling element bearings, two algorithms have been developed in this thesis. One algorithm is devoted to selecting a suitable resonance frequency band for envelope analysis. The resonance frequency band is determined by the highest energy in a residual signal dominated by bearing signals. Wavelet Packet Transform with the discrete Meyer wavelet as kernel function is applied to decompose the residual signal into subbands and RMS is calculated to quantify the signal energy in each band. The other algorithm aims to identify the location of the bearing fault. Cepstrum is used to detect harmonics in the enveloped signals and to extract fault frequencies. Classification of a faulty bearing is concluded when the fault frequency matches a bearing characteristic frequency. A completely automatic fault diagnostics for rolling element bearings, named automatic envelope analysis, is enabled by combining the two algorithms with envelope analysis. To show the effectiveness of automatic envelope analysis, we applied it to five sets of benchmark data with different faults and different operation conditions.

7.2 Discussions

In order for the envelope analysis to be effective, signals generated by shafts, blades and gears must be removed from the vibration signal as pre-processing. This is because signal energy from these components is much higher than that from bearings. TSA and LP filtering etc. can be applied to remove gear signals. However, they are sensitive to

operation speed variation or shaft speed change since harmonics of the gear meshing frequency can extend to the high frequency band. Operation speed variation will be amplified to a large extent in the high frequency band. Therefore order tracking must be conducted to convert the signal to a constant speed reference. On the other hand, for any system without gears, order tracking is not necessary if speed fluctuation are less than 5% of the mean frequency from our experience and this makes bearing fault diagnosis easier. The reason is that side bands produced by a bearing collecting around the resonance frequency do not get amplified as much as gear induced harmonics. Therefore, after a high pass filtering to remove shaft and blade signals, the proposed method can be applied to produce reliable results.

As mentioned in Section 4.2, a subband will not be combined into the band for envelope analysis if its RMS is less than 30% of the highest. The criterion is proposed by trial and error based on benchmark data. Benchmark data are generated by a test rig without gears and under constant speed operation. Vibration signals produced by a machine with gears and operating at variable speed include much more interference and noise. The criterion may need to be adjusted.

In Section 5.2, a parameter 5% is used in setting up a time window for fault signature identification. This parameter is selected from our experience. Theoretical characteristic frequencies are not accurate as mentioned in Section 2.3. Fault frequency cannot match them exactly; however the gap is not over 5% commonly.

7.3 Future work

The automatic bearing diagnostics proposed in this thesis offers a convenient and efficient way to diagnose faults for rolling element bearings. This technique has been verified in a test stand running at constant speed, which includes shafts and bearings with single point faults. To extend this work, tests in a more complicated mechanical system with shafts, bearings and gears at variable speed should be conducted to validate this technique and to test the interference extent among techniques described in Figure 1.1. Moreover, the automatic bearing diagnostics should be used to diagnose real bearing faults to test its performance further. When this algorithm is proved to be effective and reliable, it can be integrated into diagnostic instruments.

References

1. White, G.D., *Introduction to Machine Vibration*. 2008: Reliabilityweb.com.

2. Institute, E.P.R., *Manual of Bearing Failures and Repair in Power Plant Rotating Equipment*. 1991: EPRI.

3. Neale, M., Associates, and G.B.C.f. Terotechnology, *A Guide to the Condition Monitoring of Machinery*. 1979: H.M. Stationery Office.

4. Cibulka, J., et al., *A Review on Approaches for Condition Based Maintenance in Applications with Induction Machines Located Offshore.* Modeling Identification and Control, 2012. **33**(2): p. 69-86.

5. Daneshi-Far, Z., G.A. Capolino, and H. Henao. *Review of Failures and Condition Monitoring in Wind Turbine Generators*. in *2010 XIX International Conference on Electrical Machines (ICEM), 6-8 Sept. 2010*. 2010. Piscataway, NJ, USA: IEEE.

6. Hameed, Z., et al., *Condition Monitoring and Fault Detection of Wind Turbines and Related Algorithms: A Review.* Renewable and Sustainable Energy Reviews, 2009. **13**(1): p. 1-39.

7. Gong, X., W. Qiao, and Ieee, *Current-Based Online Bearing Fault Diagnosis for Direct-Drive Wind Turbines via Spectrum Analysis and Impulse Detection*. 2012 Ieee Power Electronics and Machines in Wind Applications. 2012, Los Alamitos: Ieee Computer Soc.

8. Braun, S. and B. Datner, *Analysis of Roller-Ball Bearing Vibrations.* Journal of Mechanical Design-Transactions of the Asme, 1979. **101**(1): p. 118-125.

9. Antoni, J., *Cyclic Spectral Analysis of Rolling-Element Bearing Signals: Facts and Fictions.* Journal of Sound and Vibration, 2007. **304**(3–5): p. 497-529.

10. Li, H., L.H. Fu, and H.Q. Zheng, *Bearings Fault Detection and Diagnosis Using Envelope Spectrum of Laplace Wavelet Transform*. Proceedings of the 2009 2nd International Congress on Image and Signal Processing, Vols 1-9, ed. P.H. Qiu, et al. 2009. 4143-4147.

11. Sun, Q. and Y. Tang, *Singularity Analysis Using Continuous Wavelet Transform for Bearing Fault Diagnosis.* Mechanical Systems and Signal Processing, 2002. **16**(6): p. 1025-1041.

12. Tang, Y. and Q. Sun, *Application of the Continuous Wavelet Transform to Bearing Defect Diagnosis.* Journal of Tribology-Transactions of the Asme, 2003. **125**(4): p. 871-873.

13. Xi, F.F., Q. Sun, and G. Krishnappa, *Bearing Diagnostics Based on Pattern Recognition of Statistical Parameters.* Journal of Vibration and Control, 2000. **6**(3): p. 375-392.

14. Oppenheim, A.V., R.W. Schafer, and J.R. Buck, *Discrete-Time Signal Processing*. Vol. 5. 1999: Prentice Hall Upper Saddle River.

15. McFadden, P.D., *Examination of a Technique for the Early Detection of Failure in Gears by Signal-Processing of the Time Domain Average of the Meshing Vibration.* Mechanical Systems and Signal Processing, 1987. **1**(2): p. 173-183.

16. J.J., B., R.F. Burchill, and H.L. Clark, *Design and Fabrication of Prototype System for Early Warning of Impending Bearing Failure.* NASA CR 123717, 1972.

17. McFadden, P.D. and J.D. Smith, *Vibration Monitoring of Rolling Element Bearings by the High-Frequency Resonance Technique — A Review.* Tribology International, 1984. **17**(1): p. 3-10.

18. Braun, S.G., *The Signature Analysis of Sonic Bearing Vibrations.* Ieee Transactions on Sonics and Ultrasonics, 1980. **27**(6): p. 317-328.

19. Ho, D. and R.B. Randall, *Optimisation of Bearing Diagnostic Techniques Using Simulated and Actual Bearing Fault Signals.* Mechanical Systems and Signal Processing, 2000. **14**(5): p. 763-788.

20. Randall, R.B., *Detection and Diagnosis of Incipient Bearing Failure in Helicopter Gearboxes.* Engineering Failure Analysis, 2004. **11**(2): p. 177-190.

21. Randall, R.B., *State of the Art in Monitoring Rotating Machinery-Part 2.* Sound and Vibration, 2004. **38**(5): p. 10-17.

22. Amirat, Y., et al., *Hilbert Transform-Based Bearing Failure Detection in DFIG-Based Wind Turbines.* International Review of Electrical Engineering-Iree, 2011. **6**(3): p. 1249-1256.

23. Bossley, K.M., et al., *Hybrid Computed Order Tracking.* Mechanical Systems and Signal Processing, 1999. **13**(4): p. 627-641.

24. Villa, L.F., et al., *Angular Resampling for Vibration Analysis in Wind Turbines under Non-Linear Speed Fluctuation.* Mechanical Systems and Signal Processing, 2011. **25**(6): p. 2157-2168.

25. Bonnardot, F., *Enhanced Unsupervised Noise Cancellation Using Angular Resampling for Planetary Bearing Fault Diagnosis.* International Journal of Acoustics and Vibration, 2004.

26. Bonnardot, F., et al., *Use of the Acceleration Signal of a Gearbox in order to Perform Angular Resampling (with Limited Speed Fluctuation).* Mechanical Systems and Signal Processing, 2005. **19**(4): p. 766-785.

27. Urbanek, J., et al., *Comparison of Amplitude-Based and Phase-Based Methods for Speed Tracking in Application to Wind Turbines.* Metrology and Measurement Systems, 2011. **XVIII**(2).

28. Zimroz, R., *Measurement of Instantaneous Shaft Speed by Advanced Vibration Signal Processing - Application to Wind Turbine Gearbox.* 2011.

29. Combet, F. and R. Zimroz, *A New Method for the Estimation of the Instantaneous Speed Relative Fluctuation in a Vibration Signal based on the Short Time Scale Transform.* Mechanical Systems and Signal Processing, 2009. **23**(4): p. 1382-1397.

30. McFadden, P.D., *Interpolation Techniques for Time Domain Averaging of Gear Vibration.* Mechanical Systems and Signal Processing, 1989. **3**(1): p. 87-97.

31. Fyfe, K.R. and E.D.S. Munck, *Analysis of Computed Order Tracking.* Mechanical Systems and Signal Processing, 1997. **11**(2): p. 187-205.

32. McFadden, P.D. and J.D. Smith, *A Signal-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, 1985. **199**(4): p. 287-292.

33. McFadden, P.D., *A Revised Model for the Extraction of Periodic Waveforms by Time Domain Averaging.* Mechanical Systems and Signal Processing, 1987. **1**(1): p. 83-95.

34. McFadden, P.D., *Window Functions for the Calculation of the Time-Domain Averages of the Vibration of the Individual Planet Gears and Sun Gear in an Epicyclic Gearbox.* Journal of Vibration and Acoustics-Transactions of the Asme, 1994. **116**(2): p. 179-187.

35. Wang, W.Y. and A.K. Wong, *Autoregressive Model-Based Gear Fault Diagnosis.* Journal of Vibration and Acoustics-Transactions of the Asme, 2002. **124**(2): p. 172-179.

36. Antoni, J. and R.B. Randall, *Unsupervised Noise Cancellation for Vibration Signals: Part I—Evaluation of Adaptive Algorithms.* Mechanical Systems and Signal Processing, 2004. **18**(1): p. 89-101.

37. Antoni, J. and R.B. Randall, *Unsupervised Noise Cancellation for Vibration Signals: Part II—A Novel Frequency-Domain Algorithm.* Mechanical Systems and Signal Processing, 2004. **18**(1): p. 103-117.

38. Urbanek, J., T. Barszcz, and T. Uhl, *Comparison of Advanced Signal-Processing Methods for Roller Bearing Faults Detection.* Metrology and Measurement Systems, 2012. **19**(4): p. 715-726.

39. Antoni, J., *The Spectral Kurtosis: A Useful Tool for Characterising Non-Stationary Signals.* Mechanical Systems and Signal Processing, 2006. **20**(2): p. 282-307.

40. Howard, I., *A Review of Rolling Element Bearing Vibration Detection, Diagnosis, and Prognosis.* 1994.

41. Antoni, J., *Fast Computation of the Kurtogram for the Detection of Transient Faults.* Mechanical Systems and Signal Processing, 2007. **21**(1): p. 108-124.

42. Sawalhi, N., R.B. Randall, and H. Endo, *The Enhancement of Fault Detection and Diagnosis in Rolling Element Bearings Using Minimum Entropy Deconvolution Combined with Spectral Kurtosis.* Mechanical Systems and Signal Processing, 2007. **21**(6): p. 2616-2633.

43. Shigley, J.E., Mischke, C.R. and Budynas, R.G., *Mechanical Engineering Design, 7th Edition*. 2004.

44. Kurfess, T., S. Billington, and S. Liang, *Advanced Diagnostic and Prognostic Techniques for Rolling Element Bearings*, in *Condition Monitoring and Control for Intelligent Manufacturing*, L. Wang and R. Gao, Editors. 2006, Springer London. p. 137-165.

45. Beer, F.P., E.R. Johnston, and J.T. DeWolf, *Mechanics of Materials, 4th*. 2006: McGrow Hill.

46. Tsushima, N., *Rolling-Contact Fatigue and Fracture-Toughness of Rolling Element Bearing Materials.* Jsme International Journal Series C-Dynamics Control Robotics Design and Manufacturing, 1993. **36**(1): p. 1-8.

47. Rosado, L., et al., *Rolling Contact Fatigue Life and Spall Propagation of AISI M50, M50NiL, and AISI 52100, Part I: Experimental Results.* Tribology Transactions, 2010. **53**(1): p. 29-41.

48. Keba, J.E., *Component Test-Results from the Bearing Life Improvement Program for the Space-Shuttle Main Engine Oxidizer Turbopumps*. Rotating Machinery Dynamics, ed. J.H. Kim and W.J. Yang. 1992. 303-318.

49. Eschmann, P., L. Hasbargen, and K. Weigand, *Ball and Roller Bearings - Theory, Design and Application*. 2nd Edition ed. 1985: John Willey and Sons.

50. Oppenheim, A.V., A.S. Willsky, and S.H. Nawab, *Signals and Systems*. Vol. 2. Prentice-Hall Englewood Cliffs, NJ.

51. Darlow, M.S., R.H. Badgley, and G. Hogg, *Application of High-Frequency Resonance Techniques for Bearing Diagnostics in Helicopter Gearboxes*, 1974, DTIC Document.

52. Kschischang, F.R., *The Hilbert Transform.* University of Toronto, 2006.

53. Randall, R.B., *Vibration-Based Condition Monitoring : Industrial, Automotive and Aerospace Applications*. 2010.

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

55. McFadden, P.D. and J.D. Smith, *The Vibration Produced by Multiple Point-Defects in a Rolling Element Bearing.* Journal of Sound and Vibration, 1985. **98**(2): p. 263-273.

56. Randall, R.B. and J. Antoni, *Rolling Element Bearing Diagnostics-A Tutorial.* Mechanical Systems and Signal Processing, 2011. **25**(2): p. 485-520.

57. Fugal, D.L., *Conceptual Wavelets in Digital Signal Processing*. 2009: Space & Signals Technologies. 250.

58. Hu, W., Q. Sun, and C.K. Mechefske, *Condition Monitoring for the Endurance Test of Automotive Light Assemblies.* International Journal of Advanced Manufacturing Technology, 2013. **66**(5-8): p. 1087-1095.

59. Sun, Q., *Feature Extraction with Discrete Wavelet Transform for Drill Wear Monitoring.* Journal of Vibration and Control, 2005. **11**(11): p. 1375-1396.

60. Mallat, S., *A Wavelet Tour of Signal Processing (Third Edition)* 2009: p. 267-279, 377-380.

61. Akansu, A.N. and P.R. Haddad, *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets*. 2000: Access Online via Elsevier.

62. Torrence, C. and G.P. Compo, *A Practical Guide to Wavelet Analysis.* Bulletin of the American Meteorological society, 1998. **79**(1): p. 61-78.