Direct Load Monitoring in Rolling Element Bearing by Using Ultrasonic Time of Fight

Wenqu Chen

The Department of Mechanical Engineering

The University of Sheffield

This dissertation is submitted for the degree of

Doctor of Philosophy

July 2015

Acknowledgements

I would really like to thank my supervisor, Professor Robert S Dwyer-Joyce for his encouragement, guidance and support. He always be optimism and patience in helping me during the project.

I would like to thank the University of Sheffield for the financial support. Without the fully funded scholarship, it would not have been possible to complete this study.

I would like to thank all PhDs, Postdocs and staffs in the Tribology group for their kindness and support. Thank Mr Dave Butcher for his excellent technical expertise, his provision of specimens used in this work. Thank Dr Robin Mills for his advice and help, especially his experimental skills. Thank Mr Tom Howard for his assistance in the experiments. Thank Mr Hiroyuki Suzuki for offering great tribological discussions.

Thanks to all my friends, especially Dr Huiming Kang, Dr Juanjuan Zhu and Dr Xinjin Hua for their encouragement at the initial stage of my PhD and their friendship throughout this project.

I would like to thank my parents for their love and support, without which it would not have been possible to this stage. Special thanks to my boyfriend Zhongyi, for your love, encouragement and patience!

Abstract

Rolling element bearings find widespread use in numerous machines and they are one of key components in involved systems. Bearing failures can cause catastrophic events if they are not detected in time and result in increasing downtime and maintenance cost. The need for longer endurance life with less cost drives research on bearing condition monitoring.

Load monitoring provides significant information for bearing design and residual service life prediction as load applied by each rolling element on a bearing raceway controls life friction and wear. It is possible to infer bearing load from load cells or strain gauges on the shaft or bearing housing. However this is not always simply and uniquely related to the real load transmitted by rolling elements directly to the raceway. Firstly, the load sharing between rolling elements in the raceway is statically indeterminate. And secondly, in a machine with non-steady loading the load path is complex and highly transient being subject to dynamic behavior of the transmission. This project develops a non-invasive, safe and portable technique to measure the load that transmitted directly by a rolling element to the raceway by using ultrasound.

The technique works by monitoring the time-of-flight (ToF) of ultrasound that travels in a raceway and reflects back from the contact face. A piezoelectric sensor was permanently bonded onto the external surface of the stationary raceway in a rolling element bearing. The ToF of an ultrasonic pulse from the sensor to the raceway-rolling element contact was measured which depends on the wave speed and the thickness of the raceway.

The speed of an ultrasonic wave in a component changes with the state of the stress; known as the acoustoelastic effect. The thickness of the element varies when deflection occurs as the contacting surfaces are subjected to load. Therefore, the ultrasonic ToF in a raceway is load dependent. In practical measurements, it was found that the phase of the wave reflected from rolling contacts varied with contact conditions. The phase was determined by the contact stiffness and in simple peak to peak measurement, this appeared as a change in the ToF. For typical rolling contacts, the ToF changes caused by deflection and acoustoelastic effect are of the order of nanoseconds, while the apparent time shift from the phase change effect is in the same order.

Despite the phase change having effect on reflected signals, it does not affect the enve-

lope of these signals. In this work the Hilbert transform was used to calculate the envelope of the reflected pulses and thus this contact dependent phase shift was eliminated. Time difference between the envelope of reflected pulses in unloaded and loaded state was a result of load effect alone.

Ultrasonic measurements have been carried out on a model line contact formed between a steel plate and a cylindrical bearing steel roller, and line contacts in a cylindrical roller bearing which was used for the planet gear of a wind turbine epicyclic gearbox, as well as on elliptical contacts in a radially loaded ball bearing (deep groove). The ToF changes under different contact loads were recorded and used to determine the deflection of the raceway. This was then related to load using a simple elastic contact model. Measured load from the ultrasonic reflection was compared with the applied load upon the contact and good agreement has been achieved. The ultrasonic ToF technique shows promise as an effective method for load monitoring in real bearing applications.

Contents

List of Figures

List of Tables

Nomenclature

Roman Symbols

A	Amplitude of ultrasonic wave, m
a	Semi-major axis of the elliptical contact area, m
a_l	Semi-width of the line contact, m
$A_R(f)$	Amplitude of reflected wave in loaded state
$A_{Rr}(f)$	Amplitude of reference
B	Bulk modulus, $\mathrm{N/m^2}$
b	Semi-minor axis of the elliptical contact area, m
c	speed of ultrasonic waves, m/s
c_0	Ultrasound speed in unloaded state, m/ s
c_d	Diametral clearance, m
c_P	Ultrasound speed in loaded state, m/ s
D	Diameter of the rolling element, m
d_0	Thickness in unloaded state, m
d_i	Diameter of the inner race, m
d_m	Bearing pitch diameter, m
d_o	Diameter of the outer race, m
d_P	Thickness in loaded state, m/ s

D_s Diameter of the circular sensor, m

E Modulus of Elasticity of the material, $\mathrm{N/m^2}$

e Eccentricity of the elliptical contact

$E^{'}$ Reduced modulus of elasticity, $\mathrm{N/m^2}$

F Complete elliptic integral of the first kind

h Film thickness, m

H_s Hersey number

h_{min} Minimum film thickness, m

I Incident wave

j Load-deflection exponent

K Stiffness, $\mathrm{N/m^2}$

k Elliptical parameter

K_j Load-deflection parameter

k_n Resonance number

k_w Wavenumber,m^{-1}

l The third order elastic constant

l_0 Eeffective length of the roller, m

L_s Length of the rectangle sensor, m

L_{zz} The acoustoelastic constant

m The third order elastic constant

N Number of rolling elements

n The third order elastic constant

N_{so} Length of the near field for the circular sensor, m

N_{sr} Length of the near field for the rectangle sensor, m

p	Pressure, $\mathrm{N/m^2}$
P_ψ	Load on the rolling element with the angle ψ, N
P_{max}	Maximum contact Load, N
p_{max}	Maximum contact pressure in the contact area, $\mathrm{N/m^2}$
R	Reflection coefficeint
$R^{'}$	Reduced radius of the contact bodies, m
R_1	Radius of medium 1, m
R_2	Radius of medium 2, m
R_q	Root mean square surface roughness, m
T	Transmission coefficient
t	Time, s
ToF	Time-of-flight, s
ToF_0	Time-of-flight in unloaded state, s
ToF_P	Time-of-flight in loaded state, s
U	Surface displacment, m
u	Displacement of ultrasonic waves, m
W	Load on bearing, N
W_h	Hydraulic load, N
x	Coordinate
y	Coordinate
Z	Acoustic impedance, $\mathrm{kg \cdot m^2 \cdot s^1}$
z	Coordinate
Z_w	Constant

Greek Symbols

λ	Wavelength of ultrasound, m
ω	Angular speed of ultrasonic wave, rad/s
$\Delta\varepsilon$	Strain change
Δc	Speed change, m/s
Δt_δ	ToF change caused by surface deflection, s
Δt_ϕ	ToF change caused by phase change, s
Δt_c	ToF change caused by acoustoelastic effect, s
δ	Surface deflection, m
δ_r	Radial displacement of the inner and the outer race, m
δ_{max}	Elastic deformation of the contact on the loading line, m
ε	Strain
η	Absolute viscosity, Pa· s
η_k	Kinematic viscosity, m^2/s
Λ	Dimensionless film parameter
λ	The third order elastic constant
μ	The second order elastic constant
μ_f	Friction coefficient
ν	Poisson's ratio
ω_r	Rotational speed, rpm
Φ_R	Phase of reflection coefficient
ϕ_r	Phase of reflection coefficient for reference, radians
ψ	Angle position of the ball to the load line, rad
ψ_l	Angular extent of the load zone, rad
ρ	Density, kg/m^3

ρ_0	Density in unloaded state
σ	Stress, $\mathrm{N/m^2}$
τ	Shear stress, $\mathrm{N/m^2}$
ε	Complete elliptic integral of the second kind
ξ	Pressure-viscosity coefficient, $\mathrm{GPa^{-1}}$

Subscripts

0	Subscript 0 refers to unloaded state
1	Subscript 1 refers to medium 1
2	Subscript 2 refers to medium 2
3	Subscript 3 refers to oil
P	Subscript P refers to loaded state
x	Subscript x refers to coordinate
y	Subscript y refers to coordinate
z	Subscript z refers to coordinate

Chapter 1

Introduction

1.1 Statement of the problem

Rolling element bearings make use of the rolling action of a complement of rolling elements (rollers or balls) which are placed between two rings to permit relative motion with minimum friction. They are key components in rotational machines, and bearing failure is a major cause of the machine breakdowns. Bearing failures can be disastrous and cause increasing downtime and maintenance cost if they are not detected in time and corresponding actions are taken, especially for those applications with limited accessibility.

Knowledge of load is essential information for the design and the prediction of bearing service life. In bearing design stage, the endurance life is commonly evaluated using the classical standard ISO 281 (International Organization for Standards (2010)) and this works effectively for most bearings in general applications. However, for those applications with highly fluctuating and dynamic load, like wind turbines, there are unexpected premature failures during operation (Al-Tubi and Long (2013)). The bearing failure in the gearbox has been reported to account for up to one-third of wind turbine failures, which significantly reduce the wind turbine availability (Christopher and Simon (2010)).

To properly predict bearing fatigue life, accurate load and load history data is needed. Bearing load can be directly measured with load cells or strain gauges on the bearing housing or the shaft (Chen and Chen (2005) and Nässelqvist et al. (2012)). A disadvantage of these techniques is these devices treat the bearing as a whole. As a result, the load dynamically transmitted by the rolling element to the race is not characterized. In rolling element bearings the load is shared by a number of rolling element and raceway contacts and the exact load in these contacts cannot always be uniquely derived from the load on the shaft or bearing. In some machines it may be difficult to determine the bearing load, particularly where the load paths are complex and/or highly transient. For instance, load on the bear-

ing in wind turbines is not easily derived from load on the blades or the shaft where it is measured due to complex load paths and the nature of highly varying loading.

Ultrasonic technique is a non-invasive and effective method for monitoring contact conditions. It has been commonly used since the 1920s for detecting the material composition and has been recently proved to be an effective means to characterize contact between surfaces. For instance, ultrasonic reflections from the contact interface can be used for determining the stiffness of the intermediate layer between the adhesive and the adherend (Pialucha and Cawley (1994)), and the stiffness of the contact interface between two rough surfaces (Drinkwater et al. (1996)). Moreover, they can be used for monitoring the presence of the lubricant layer, measuring the film thickness in both hydrodynamic and elastohydrodynamic lubrication (Dwyer-Joyce et al. (2003, 2004)). Besides the lubricant film thickness, load is another vital factor that controls bearing service life and a new approach needs to be explored by using ultrasound.

1.2 Project objectives

Based on these achievements in making use of ultrasound, this study aims at developing a novel approach for determining the load transmitted by rolling contacts in rolling element bearings via measuring the ToF change of ultrasonic reflections, and applying this approach to real bearing applications in industry. The key objectives were thus documented as:

- Develop the mathematical model for the ToF change and contact load. The ToF of ultrasonic pulses from the sensor to the contact face is governed by the speed of wave and the thickness of the raceway, both of them are load dependent and thus the ToF.
- Develop the signal processing software to measure the ToF change in ultrasonic reflections caused by the deflection and the speed change. As the phase change of the reflected wave introduces a significant apparent time shift, a technique to eliminate this effect was required.
- Perform experiments on a model line contact to assess the ability of the ultrasonic ToF technique for measuring load upon a line contact.
- Perform experiments on a cylindrical roller bearing to assess the ability of the ultrasonic ToF technique for measuring load in a roller bearing application. The cylindrical roller bearing was used in a 2-megawatt wind turbine gear boxes.
- Perform experiments on a ball bearing to assess the ability of the ultrasonic ToF technique for measuring load in ball bearings.

1.3 Layout of the thesis

Chapter 2 briefly describes the contact mechanism in rolling element bearings, such as the load distribution, the size of the contact patch, the surface deflection and stress distribution for two typical contacts. Also, the lubricant properties and lubrication regimes are introduced. A review of performance monitoring techniques for rolling bearing contacts is provided.

Chapter 3 describes the response of different boundaries to ultrasonic waves. The response of an ultrasonic wave at a boundary interface is governed by the acoustic impedance of the media, the stiffness of the boundary and the frequency of the wave. Three typical methods for determining the thin layer thickness between two surfaces are introduced. The phase change in the reflected wave is investigated as it affects the ultrasonic ToF determination. The acoustoelastic effect, the variation of ultrasound speed in a material with load, is then introduced as it is one of the main causes of the ToF change for a loaded raceway.

Following this, the mathematical model for the ultrasonic ToF method is developed in Chapter 4. The ToF in a bearing race when it is unloaded and loaded are given, and the ToF change caused by the deflection and the speed change are quantified. The apparent time shift caused by the phase change of the reflected wave in practical measurements, which is determined by the contact condition, is also investigated.

Chapter 5 gives a brief description of the ultrasonic load measurement system that has been used in the following Chapters. In the signal processing routine, the Hilbert transform is used for calculating the envelop of each signal where the phase change effect is excluded. The time difference between the envelope of the loaded signal and that of the unloaded signal is determined by using a cross-correlation technique. The measured time difference is the ToF change caused by the deflection and the speed change in the bearing race as it is loaded.

Having developed the measurement system, preliminary experiments have been done on a model line contact, they are given in Chapter 6. Ultrasonic reflections from a model line contact between a flat steel plate and a cylindrical bearing steel roller under a number of load were recorded and used to measure the ToF change.

Load measurements are then carried out on line contacts in a cylindrical roller bearing, they are detailed in Chapter 7. The ultrasonic ToF in the stationary inner ring was monitored during the bearing operation; it reduced for each roller pass as the raceway was loaded by the roller approach. By using the model that has been developed in Chapter 4, load upon the contact can be obtained from the measured ToF change.

Load measurements with the ultrasonic ToF approach were performed on an elliptical contact in a ball bearing as well and they are documented in Chapter 8. Again, a clear

reduction in the ultrasonic ToF was observed for each ball pass. Loads measured from the ultrasonic ToF changes are consistent with loads applied on the ball bearing contact.

Chapter 9 draws an overall conclusion arose from this work and gives general recommendations for future study.

1.4 Original contribution of the project

The novelty of the work within this thesis is developing and implementing a unique technique to monitor load transmitted by rolling contacts in rolling element bearings by using the ultrasonic ToF. The contribution of the elastic surface deflection, the change in speed of sound and the phase change effect on the ToF change have been modeled; all three components were significant and could not be neglected. The first two terms are load dependent whilst the last term is contact dependent. The ToF change contributed by load and that by contact are separated. A load monitoring system based on active piezoelectric sensors has been developed. The ultrasonic sensors have the advantage of cost-effectiveness and low profile causing less disruption of bearings. This technique has been applied to a model line contact and rolling contacts in a ball bearing, load from the ultrasonic ToF corresponded well to the applied load. Moreover, this technique has been evaluated on line contacts in a wind turbine gear box bearing, good agreement between measured and applied load has been achieved. This technique shows promise for application to large bearings when both the contact dimension and the raceway deflection are large and hence lead to readily measurable ToF change.

Chapter 2

Rolling Element Bearing Contacts Performance Monitoring

This chapter introduces the basics of rolling element bearings first: bearing types and bearing load distributions. It follows by the classical Hertz contact mechanism where the contact shape and size, the displacement and the stress distribution in contact surfaces are studied. It continues to provide the reader with a brief introduction of lubrication in bearings, followed by a review of most widely used techniques for monitoring bearing performance. Finally, a conclusion for this chapter is given.

2.1 Introduction to rolling element bearings

A bearing is a mechanical device which functions to provide relative smooth motion to one another machine element and transmit the force associated with machine elements efficiently (low friction). Bearings find extremely wide use in both domestic and industrial applications; today hundreds of billions pounds are spent on bearing production. There are several means to group bearings, according to the mode of operation, the direction of load they transmitted, or the geometry profile (Hamrock et al. (2004)).

According to the operation type, bearings can be classified as:

- Dry rubbing bearings. These bearings are the simplest and lowest cost, in which two bearing surfaces rub together in rolling or sliding motion or both.
- Hydrodynamic bearings. The two bearing surfaces in this case are effectively separated by a thin layer of fluid film which can be a gas or a liquid. The load carrying capacity relies on the fast rotational speed of the machine element (the journal) to build up a pressurized film between the two bearing surface.

- Hydrostatic bearings. As with hydrodynamic bearings, the two surfaces are separated by a fluid layer but their load carrying capacity relies on external pressurization.
- Rolling element bearings. By using the rolling motion of balls or rollers, these bearings minimize friction and provide strained relative motion of one component to another.

The selection of bearing can be based on the principles provided by (Engineering Sciences Data Unit (1965)). Compared with hydrodynamic bearings, rolling element bearings have extraordinary characteristic of low starting friction coefficient and the ability of supporting combined radial and thrust load and shorter axial length.

Rolling element bearings have widespread applications in engineering, therefore, the main interest of this work is this class of bearings. The majority of rolling bearings are adopted as they permit relative rotary motion between a shaft and some fixed structure, although some of them are adopted because they permit relative linear motion of a fixture in the stationary shaft direction (Harris (2001)). In this work, rolling element bearings refer to those providing rotary motion between two machine elements.

The standardized rolling element bearing commonly consists of four machine elements, an outer ring, a complement of rolling elements (balls or rollers), a cage and a inner ring. The rolling elements are placed between the two rings, constrained by the carrier cage which functions to separate these elements from each other at a specific distance. Reti (1971) shows that Leonardo da Vinci found it was important to space the rolling elements without touching otherwise it would cause opposite motions.

2.1.1 Ball bearings

Ball bearings have three subgroups according to the direction of load: radial, angular contact and thrust ball bearings.

2.1.1.1 Radial ball bearing

Radial ball bearings have been most popular bearings in industrial applications such as agriculture and automotive. They can carry comparably high radial load and the raceways have a deep-groove configuration on which balls roll. A deep groove ball bearing (single row) from SKF is given in figure 2.1a.

2.1.1.2 Angular contact ball bearing

Angular contact ball bearings are employed as they can accommodate both radial load and axial load. They are characterized by the contact angle which determines their load-carrying capacities. The angle between the line, which joins the contact points between the ball and the raceways (the loading transmission line) in bearing radial plane, and the line perpendicular to the bearing rotational axis, is defined as the contact angle. Figure 2.1b shows an angular contact bearing from SKF and figure 2.1 schematically illustrates the contact angle and the load path in the bearing. The contact angle can be increased to support large thrust load, however, it is commonly limited to 40 degrees.

The angular contact ball bearings can be classified into three subgroups: single row, double row and four-points contact ball bearings. The single row bearing can bear axial load in a specific direction and usually mounted in pairs. However, the double row angular contact bearing is able to support thrust load without direction limitation. For some particular applications, bearing rings are split to form four-point contact, as illustrated by figure 2.1d. This class of bearing can withstand radial, thrust and moment loads having less space and weight.

2.1.1.3 Thrust ball bearing

The thrust ball bearing is designed only for supporting axial loads. Figure 2.1e shows a thrust bearing assembly from SKF which includes a separable shaft washer, a housing washer and a cage assembly with balls.

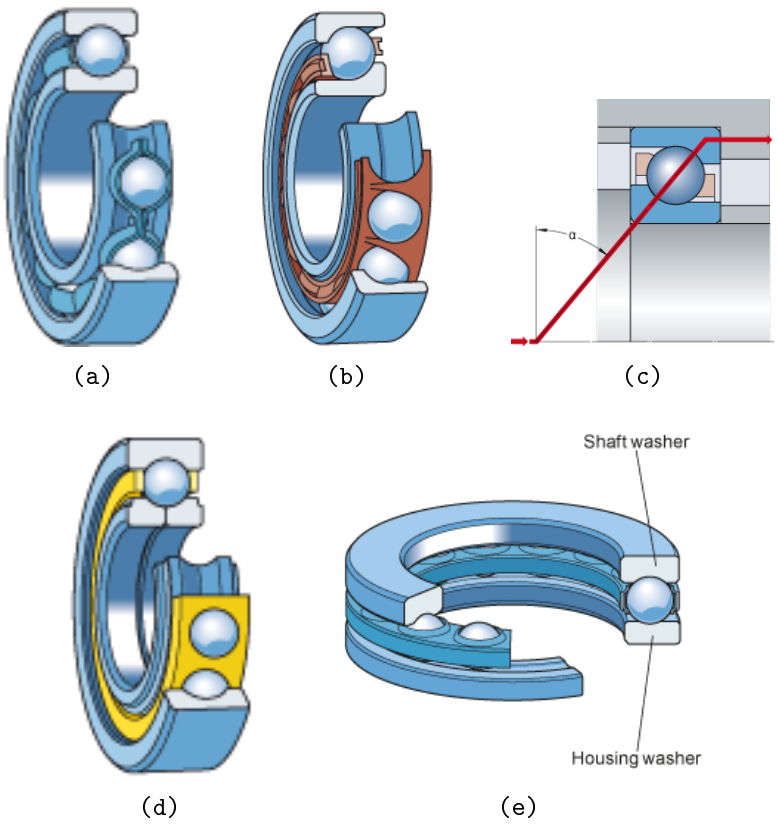


Figure 2.1. *Photograph of (a) a deep groove ball bearing; (b) an angular contact bearing; (c) Contact angle of an angular contact bearing; (d) a four-point contacts bearing; (e) a single row thrust bearing, reproduced from SKF (acessed on 28-02-2015).*

2.1.2 Roller bearings

Roller bearings utilizing rollers to transmit load and motion from the inner ring to the outer, or vice verse, as their name implies. Normally, they are used in those applications where the requirement of accommodating large load can not be satisfied by ball bearings. Compared with ball bearings, roller bearings with similar size can support higher load since their contact mechanisms are different. The contacts between balls and rings are considerable small ellipses while the contacts between the roller and rings are narrow strips in the roller length direction, which makes roller bearings outstanding for higher load-carrying capacity. Also, roller bearings have much stiffer structure than ball bearings which provide longer service life with more cost.

2.1.2.1 Radial roller bearings

Cylindrical roller bearings find their use in numerous applications, such as high speed operations, because of their exceptionally low friction torque characteristic, and applications requiring high radial load-carrying capacity. A photograph of a single row cylindrical roller bearing is given in figure 2.2a. By increasing the number of rows of rollers, higher radial load-carrying capacity can be achieved (Harris (2001)).

2.1.2.2 Tapered roller bearings

Figure 2.2b illustrates a tapered roller bearing which is employed in those applications where the bearing needs to carry both large radial and thrust loads. As the bearing name implies, the roller is tapered and it contacts with the inner and the outer race at different angles. The difference in contact angles generates high sliding friction force at the roller ends and the guide flange and hence is not suitable for applications with high operational speed except when the bearing is equipped with special cooling and lubricating system (Harris (2001)).

2.1.2.3 Spherical roller bearings

Spherical roller bearings can commonly internally self-aligning because their outer race are in the form of a part of an sphere, as illustrated by figure 2.2c. This type of bearings are mainly used for radial loads although a few of they can be used for radial load together with limited moment load.

2.1.2.4 Thrust roller bearings

As with thrust ball bearing, the roller bearing is designed only for accommodating thrust load. The roller can be cylindrical, tapered or spherical roller, the geometry is similar to that of thrust bearing (replacing balls by rollers) and a cylindrical thrust bearing is shown in figure 2.2c. Thrust roller bearings find their use mostly in low speed applications as the presence of inherent sliding force.

Figure 2.2. *Photograph of (a) a cylindrical roller bearing bearing; (b) a tapered roller bearing (single row); (c) a spherical roller bearing (double rows)with brass cages; (d) a thrust roller bearing; reproduced from Timken (acessed on 28-02-2015).*

2.2 Static load distribution in radially loaded rolling bearings

By using a complement of balls or rollers, rolling bearings transmit load from one ring to another through the contacts between balls or rollers and raceways. In this work, the principal interest is given to radial load on rolling contacts without the loss of generality. Load on the bearing in such a case is shared by a serials of balls or rollers, and the load distribution in the bearing is determined by bearing geometry and load magnitude. The analysis of load distribution in bearings documented in this section will be the used later in Chapter 7 and 8.

In this work, rolling element bearings are in steady state, rotated by either the outer

(roller bearing in Chapter 7) or the inner ring (ball bearing in Chapter 8), where the rotational speed is not high enough to cause significant centrifugal forces. Therefore, centrifugal forces on rolling elements are ignored in the load distribution analysis.

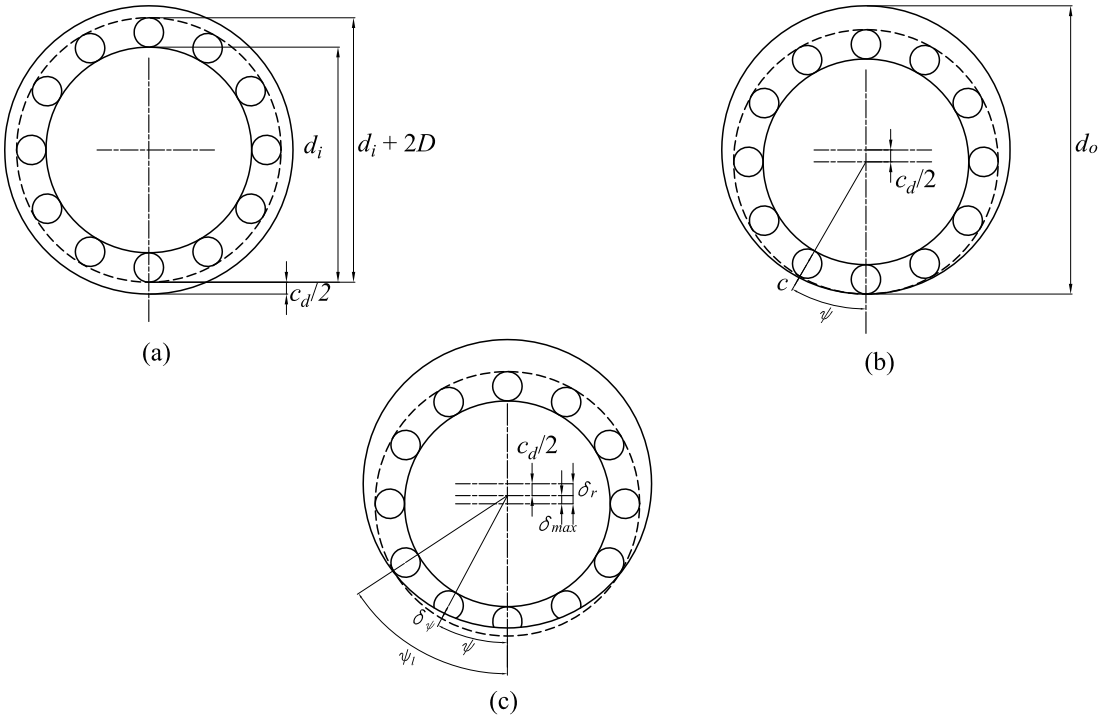


Figure 2.3. *Radially loaded rolling element bearing, reproduced from Hamrock and Anderson (1983)*

Contacts are made between rolling elements and races and deformation of surfaces take place when they are subjected to load. Figure 2.3 shows three conditions of a rolling element bearing under radial load, which having a diametral clearance denoted by c_d. The diametral clearance is determined by the diameter of the outer and inner ring, d_o and d_i respectively, and the diameter of the rolling element, D, as :

$$c_d = d_o - d_i - 2D. \tag{2.1}$$

Figure 2.3a shows the concentric state of the bearing with a uniform radial clearance before any load is applied. Figure 2.3b shows the initial contact state when a tiny load is applied on the shaft in radial direction, making the bearing inner ring moves toward the outer ring as the existence of radial clearance. The ball or roller on the load line initially contact with the bearing rings; however, the rolling element at an angle position ψ to the load line

still has a clearance away from the races, which is denoted by c and can be expressed by (Hamrock and Anderson (1983)):

$$c = (1 - \cos\psi)\frac{c_d}{2}. \tag{2.2}$$

Once the initial contact formed, further load application introduces elastic deformation in the rolling elements and the races. The load upon the rolling element on the load line (where $\psi = 0$) , P_{max}, is related to its deflection as (Hamrock and Anderson (1983)):

$$P_{max} = K_j \left(\delta_{max}\right)^j, \tag{2.3}$$

where K_j is the contact-deflection factor, δ_{max} represents the maximum elastic deformation. The magnitude of j is 1.5 for the cases of ball bearing and is unity for the cases of roller bearing.

In fact, the load is shared by some of rolling elements rather than all of them in radial ball or roller bearings. Provided the angular loading range in the bearing is designated by $2\psi_l$, a half of which is shown in figure 2.3c, the load then is only distributed on those elements located in the load arc $2\psi_l$. The magnitude of ψ_l depends on the applied load and the clearance as well as the material properties, expressed as (Hamrock and Anderson (1983)):

$$\psi_l = \cos^{-1}\frac{c_d}{2\left(\delta_{max} + \frac{c_d}{2}\right)}. \tag{2.4}$$

For a rolling element positioned at ψ to the load line, the deformation of contact denoted by δ_ψ, can be obtained from the elastic compression on the load line as :

$$\delta_\psi = (\delta_{max}\cos\psi - c) = \left(\delta_{max} + \frac{c_d}{2}\right)\cos\psi - \frac{c_d}{2}. \tag{2.5}$$

From figure 2.3, it is apparent that the total radial approaches between the inner and outer raceway, can be expressed as $\delta_r = \delta_{max} + \frac{c_d}{2}$. Therefore equation (2.5) can be written as :

$$\delta_\psi = \delta_r\cos\psi - \frac{c_d}{2}. \tag{2.6}$$

By using the load deflection relationship, as illustrated by equation (2.3), the load on the element can be obtained from :

$$P_\psi = K_j\left(\delta_r\cos\psi - \frac{c_d}{2}\right)^j. \tag{2.7}$$

According to the bearing static force equilibrium, the applied normal load equals to the

sum of the components of load on each element,

$$W = \sum P_{\psi} \cos \psi \,. \tag{2.8}$$

Therefore, the bearing load can be expressed as the summation of load on elements in the load zone $2\psi_l$. Substituting equation (2.7) into equation (2.8), the ball bearing load can be described by equation (2.9) and the roller bearing load can be described by (2.10) (Hamrock and Anderson (1983)), respectively.

$$W = \frac{N}{\pi} K_{1.5} \delta_r^{1.5} \int_0^{\psi_l} \left(\cos \psi - \frac{c_d}{2\delta_r} \right)^{1.5} \cos \psi \, \mathrm{d}\psi, \tag{2.9}$$

$$W = \left(\psi_l - \frac{c_d}{2\delta_r} \sin \psi_l \right) \frac{N K_1 \delta_r}{2\pi}. \tag{2.10}$$

Rearrange equation (2.3) and (2.9), the maximum load on the ball can be expressed as equation (2.11), while the maximum load on the roller can be expressed as equation (2.13) by rearranging euqation (2.3) and (2.10).

$$P_{max} = \frac{W Z_w}{N}. \tag{2.11}$$

$$Z_w = \frac{\pi \left(1 - \frac{c_d}{2\delta_r} \right)^{1.5}}{\int_0^{\psi_l} \left(\cos \psi - \frac{c_d}{2\delta_r} \right)^{1.5} \cos \psi \, \mathrm{d}\psi} \approx \frac{\pi \left(1 - \frac{c_d}{2\delta_r} \right)^{1.5}}{2.491 \left(\left[1 + \left(\frac{2\delta_r - c_d}{2.46\delta_r} \right)^2 \right]^{0.5} - 1 \right)}, \tag{2.12}$$

$$P_{max} = \frac{W \left(1 - \frac{c_d}{2\delta_r} \right)}{\frac{N}{2\pi} \left(\psi_l - \frac{c_d}{2\delta_r} \sin \psi_l \right)}. \tag{2.13}$$

For ball bearing cases, the calculation of maximum contact load requires iteration to achieve the constant Z_w. The procedure starts with assuming the initial value of 5 when ball bearings have a normal diametral clearance, and calculates the load on the most heavily loaded ball by equation (2.11) which can then be used to calculate the local elastic compression on the inner and outer race. The maximum deflection for contacts made on the inner and outer ring is then be used to calculate Z_w again. The iteration process continues until

the difference between two consequent values of Z_w satisfies the desired accuracy.

Comparing equations (2.3) and (2.7), it can be found that the load on the rolling balls or rollers at the angular position ψ can be calculate from the following equation:

$$P_\psi = P_{max}\left[1-(1-\cos\psi)\left(1-\frac{c_d}{2\delta_{max}}\right)^{-1}\right]. \tag{2.14}$$

Herein, for a specific load on a bearing, the static load distribution can be calculated by making use of equations (2.4), (2.11), (2.13) and (2.14).

2.3 Hertz contact

Figure 2.4 shows a initial contact between two non-conformal ellipsoids without any applied load, in which the two contacting solids have different radii of curvature in a pair of principal planes (x,y). In the following contact stress and deformation analysis, reduced curvature of the contact in principal planes are introduced and they are defined as :

$$\frac{1}{R_x} = \frac{1}{R_{1x}} + \frac{1}{R_{2x}}, \tag{2.15}$$

$$\frac{1}{R_y} = \frac{1}{R_{1y}} + \frac{1}{R_{2y}}, \tag{2.16}$$

where the curvature for a convex surface is positive and the curvature for a concave surface is negative. For instance, the curvature of the outer ring for a bearing is negative as its curvature centre is outside the ring while the curvature of the rolling element is positive as its curvature centre is within the element.

The initial contact is taken as the origin of a rectangular coordinate, the major axis of the contact is taken as the x axis and the minor axis of the contact is taken as the y axis. The z axis is normal to the common plane of the two contact bodies.

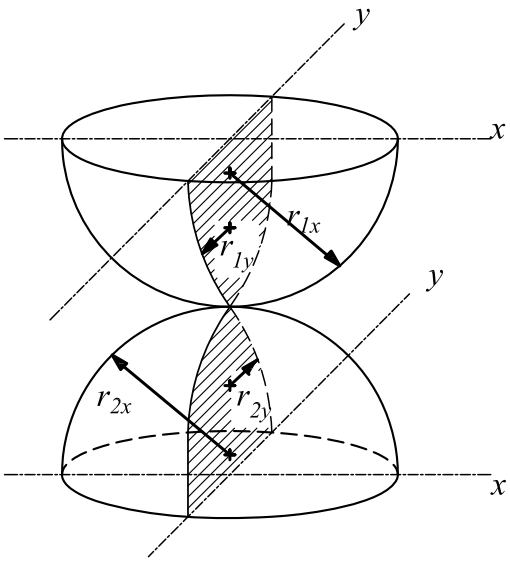


Figure 2.4. *Geometry of contacting bodies in general*

The initial contact enlarges as the contacting bodies are subjected to load, the contact shape and size are determined by load, the elastic properties and the curvatures of the two components. When two cylinders are inclined at 90 degree or two spheres are loaded against each other, they make contact over an area which is in the shape of a circle. The shape of the contact changes to an ellipse when a ring (outer or inner) is loaded by a ball and it becomes a rectangle when a ring (outer or inner) is loaded against a roller. The circle contact, the elliptical contact and the rectangle contact are three typical tribological contacts.

When two contact media are brought to contact by a load, stresses and elastic deformations occur in the contacting surfaces and they can be described by the classical Hertz theory. Before making use of the well known Hertz theory, the assumptions are summarized here (Hertz (1881)):

1. The two contacting media are perfectly smooth and continuous;
2. The materials are homogeneous and the stresses in the contacting media not exceed the material yield stress;
3. The contact solids do not introduce tangential forces;
4. The contact patch is relative small compared with the radii of the ellipsoids;
5. The contact solids are in their static equilibrium states.

2.3.1 Elliptical contacts in ball bearings

2.3.1.1 Contact size

Contacts between balls and the outer or inner ring are elliptical in shape and the stress and deformation only occur in this area. When a ball is under a load of P ($P = P_\psi$ or $P = P_{max}$), the length of the contact in the major and the minor axes, $2a$ and $2b$, respectively, can be expressed as (Hertz (1881)):

$$a = \left(\frac{6k^2\varepsilon PR^{'}}{\pi E^{'}}\right)^{\frac{1}{3}}, \tag{2.17}$$

$$b = \left(\frac{6\varepsilon PR^{'}}{\pi k E^{'}}\right)^{\frac{1}{3}}, \tag{2.18}$$

where k is the elliptical parameter, ε is the complete elliptic integral of the second kind, $R^{'}$ is the reduced radii and $E^{'}$ is the reduced elasticity of the contacting media. They are defined by :

$$k = \frac{a}{b}, \tag{2.19}$$

$$\varepsilon = \int_0^{\frac{\pi}{2}} \left[1 - \left(1 - \frac{1}{k^2}\right)\sin^2\phi\right]^{\frac{1}{2}} \mathrm{d}\phi, \tag{2.20}$$

$$R^{'} = \left(\frac{1}{R_x} + \frac{1}{R_y}\right)^{-1}, \tag{2.21}$$

$$E^{'} = 2\left(\frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2}\right)^{-1}. \tag{2.22}$$

2.3.1.2 Surface stress

The pressure distribution in the elliptical contact area follows an ellipsoid profile as pointed out by Hertz (1881), which can be expressed as

$$p = p_{max}\left[1 - \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2\right]^{\frac{1}{2}}, \tag{2.23}$$

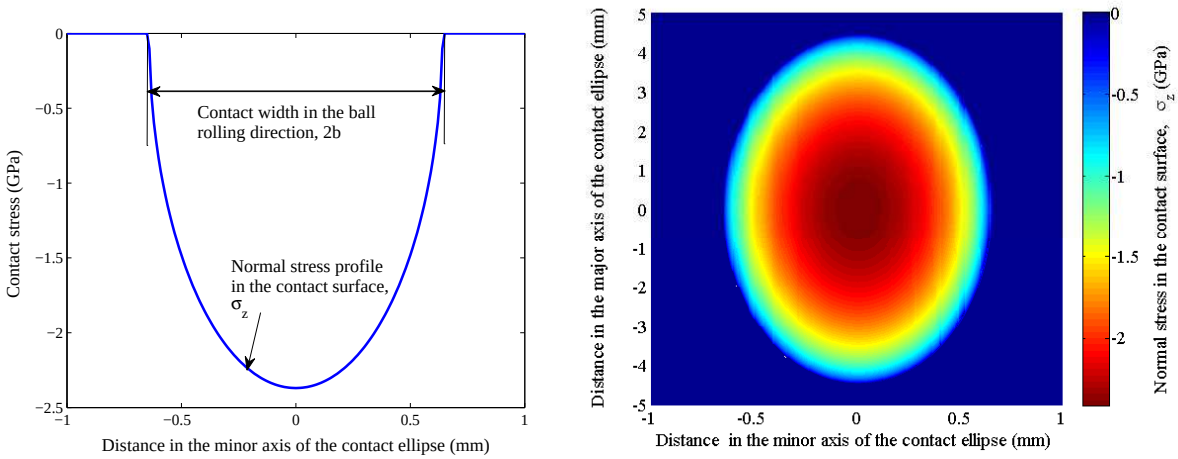
where p_{max} represents the maximum normal pressure in the contact. It takes place at the geometrical contact centre and can be calculated from load upon the contact as :

$$p_{max} = \frac{3P}{2\pi ab}. \tag{2.24}$$

The magnitude of normal surface stress in the contact is equal to that of the pressure but its sign is negative since the contact surface is compressed and it reduces to zero for surface outside the contact, expressed as

$$\sigma_z = \begin{cases} -p_{max}\left[1-\left(\frac{x}{a}\right)^2-\left(\frac{y}{b}\right)^2\right]^{\frac{1}{2}} & \left(\frac{x}{a}\right)^2+\left(\frac{y}{b}\right)^2 \leq 1 \\ 0 & \left(\frac{x}{a}\right)^2+\left(\frac{y}{b}\right)^2 > 1 \end{cases}. \tag{2.25}$$

Figure 2.5 gives an example of the surface distribution in an elliptical contact in a deep groove ball bearing. Although the bearing is under a medium load of 22.7 kN, the contact stress is rather high; the maximum stress in the contact center goes up to 2.34 GPa. This is because the load is only distributed in a small elliptical area. As it can be seen from the figure, the semi-major axis of the contact ellipse is about 4.4 mm while the semi-minor axis is only 0.65 mm.



`Figure 2.5.` *Normal surface stress in an elliptic contact in a deep groove ball bearing 6410 under a radial load of 22.7 kN.*

Within the contact, the surface stress distribution along the x axis can be expressed as

(Johnson (1987)) :

$$\frac{\sigma_x}{p_{max}} = -2v\gamma - (1-2v)\frac{b}{e^2 a}\left[\left(1-\frac{b\gamma}{a}\right) - \frac{x}{ae}\tanh^{-1}\left(\frac{ex}{a+b\gamma}\right)\right], \tag{2.26}$$

$$\frac{\sigma_y}{p_{max}} = -2v\gamma - (1-2v)\frac{b}{e^2 a}\left[\left(\frac{a\gamma}{b}-1\right) + \frac{x}{ae}\tanh^{-1}\left(\frac{ex}{a+b\gamma}\right)\right]. \tag{2.27}$$

and the stress distribution along the y axis can be expressed as (Johnson (1987)):

$$\frac{\sigma_x}{p_{max}} = -2v\gamma - (1-2v)\frac{b}{e^2 a}\left[\left(1-\frac{b\gamma}{a}\right) - \frac{y}{ae}\arctan\left(\frac{aey}{b(b+a\gamma)}\right)\right], \tag{2.28}$$

$$\frac{\sigma_y}{p_{max}} = -2v\gamma - (1-2v)\frac{b}{e^2 a}\left[\left(\frac{a\gamma}{b}-1\right) + \frac{y}{ae}\arctan\left(\frac{aey}{b(b+a\gamma)}\right)\right], \tag{2.29}$$

where $e = \sqrt{1-1/k^2}$ and $\gamma = \sqrt{1-(x/a)^2-(y/b)^2}$.

At the contact centre, the original of the coordinates, the stress is :

$$\frac{\sigma_x}{p_{max}} = -2v - (1-2v)\frac{b}{a+b}, \tag{2.30}$$

$$\frac{\sigma_y}{p_{max}} = -2v - (1-2v)\frac{a}{a+b}. \tag{2.31}$$

For a point outside the elliptical contact zone, the surfaces stresses are (Johnson (1987)):

$$\frac{\sigma_x}{p_{max}} = -\frac{\sigma_y}{p_{max}} = -(1-2v)\frac{b}{e^2 a}\left[1 - \frac{x}{ae}\tanh^{-1}\left(\frac{ex}{a}\right) - \frac{y}{ae}\arctan\left(\frac{aey}{b^2}\right)\right], \tag{2.32}$$

$$\tau_{xy} = -(1-2v)\frac{b}{e^2 a}\left[\frac{y}{ae}\tanh^{-1}\left(\frac{ex}{a}\right) - \frac{x}{ae}\arctan\left(\frac{aey}{b^2}\right)\right]. \tag{2.33}$$

2.3.1.3 Subsurface stress

The stress distribution in the contact surfaces has been studied by many researchers and it is quite complicated. The stress component for a general point in the solid is not straightforward, however, it can be obtained from finite element analysis. Johnson (1987) has summarized the expression for points along the z axis from the work of Thomas and Hoersch (1930), Belajev (1917) and Lundberg and Sjövall (1958), and they can be expressed as :

$$\frac{\sigma_x}{p_{max}} = \frac{2b}{e^2 a}\left(\Omega_x + \nu\Omega_x^{'}\right), \tag{2.34}$$

$$\frac{\sigma_y}{p_{max}} = \frac{2b}{e^2 a}\left(\Omega_y + \nu\Omega_y^{'}\right), \tag{2.35}$$

$$\frac{\sigma_z}{p_{max}} = -\frac{b}{e^2 a}\left(\frac{1-T^2}{T}\right), \tag{2.36}$$

where

$$\Omega_x = -\frac{1}{2}(1-T) + \frac{z}{a}(F-\varepsilon), \tag{2.37}$$

$$\Omega_x^{'} = \left(1 - \frac{a^2 T}{b^2}\right) + \frac{z}{a}\left(\frac{a^2}{b^2}\varepsilon - F\right), \tag{2.38}$$

$$\Omega_y = \frac{1}{2} + \frac{1}{2T} - \frac{a^2 T}{b^2} + \frac{z}{a}\left(\frac{a^2}{b^2}\varepsilon - F\right), \tag{2.39}$$

$$\Omega_y^{'} = -1 + T + \frac{z}{a}(F-\varepsilon), \tag{2.40}$$

$$T = \left(\frac{b^2+z^2}{a^2+z^2}\right)^{\frac{1}{2}}, \tag{2.41}$$

where F is the complete elliptic integral of the first kind, a is the semi-major axis, b is the semi-minor axis and z is the coordinate in the normal load direction.

$$F = \int_0^{\frac{\pi}{2}} \left[1 - \left(1 - \frac{1}{k^2}\right)\sin^2\phi\right]^{-\frac{1}{2}} \mathrm{d}\phi\,. \tag{2.42}$$

The simplified approximation equations for k, F , ε are obtained by Brewe and Ham-

rock (1977) with remarkably high accuracy, the errors are less than 3% when the elliptical parameter is between 1 to 10.

$$k \approx 1.0339 \left(\frac{R_x}{R_y} \right)^{0.636}, \tag{2.43}$$

$$F \approx 1.5277 + 0.6023 \ln \left(\frac{R_x}{R_y} \right), \tag{2.44}$$

$$\varepsilon \approx 1.0003 + 0.5968 \left(\frac{R_x}{R_y} \right)^{-1}. \tag{2.45}$$

As an example, figure 2.6 gives the stress distribution in the outer race of the ball bearing 6410, through the race thickness direction. The normal stress reduces from 2.42 GPa at the contact surface, to 1 GPa at the depth of 1.36 mm, and it reduces to 0.1 GPa at the outer surface of the outer race.

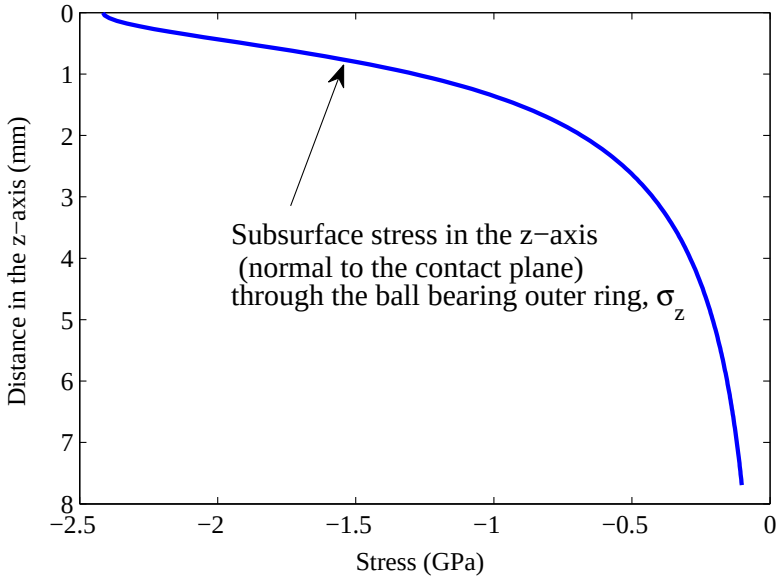


Figure 2.6. *Normal contact stress in the subsurface of the ball bearing outer ring. The bearing is under a radial load of 22.7 kN.*

2.3.1.4 Surface deflection

According to the load deflection relationship which is illustrated by equation (2.3), the maximum deformation of the contact can be derived as :

$$\delta_{max} = F\left[\frac{9}{2\varepsilon R'}\left(\frac{P_{max}}{\pi k E'}\right)^2\right]^{\frac{1}{3}}. \tag{2.46}$$

At the geometrical contact centre ($x = y = z = 0$), the deflection of the contact is the total displacement of both surfaces $U_1(0,0,0) + U_2(0,0,0) = \delta_{max}$, thus, the surface deflection is half of the contact deformation. For a general surface point $(x,y,0)$ within the contact area, the surface displacement is related to the contact deformation by (Johnson (1987)):

$$U_1(x,y,0) + U_2(x,y,0) = \delta_{max} - \hat{A}x^2 - \hat{B}y^2, \tag{2.47}$$

where the parameter $\hat{A}$, $\hat{B}$ are determined by the first and second kind of complete elliptic integrals, they are expressed as :

$$\hat{A} = \frac{2bp_{max}}{E'e^2a^2}(F - \varepsilon), \tag{2.48}$$

$$\hat{B} = \frac{2bp_{max}}{E'e^2a^2}\left(\frac{a^2}{b^2}\varepsilon - F\right). \tag{2.49}$$

By using equation (2.47), (2.48) and (2.49), the surface deformation in the contact area can be predicted.

2.3.2 Line contacts in roller bearings

2.3.2.1 Contact size

When a roller is loaded against the raceway by a normal load of P, the length of the contact zone in the major axis direction is relative large with compared with the length in the direction of the minor axis, resulting in the elliptical parameter approaches to infinity and a rectangle contact is obtained. The semi-width of the contact patch can be predicted from (Hertz (1881)) :

$$a_l = R_x\left(\frac{8P}{\pi l_0 R_x E'}\right)^{\frac{1}{2}}, \tag{2.50}$$

where l_0 is the effective length of the roller.

2.3.2.2 Surface stress

The compressive surface stress in the contact area follows a semi-cylindrical form distribution and can be given by (Hertz (1881)) :

$$p = p_{max}\left[1-\left(\frac{x}{a_l}\right)^2\right]^{\frac{1}{2}}, \tag{2.51}$$

where p_{max} is the maximum stress in the contact, expressed as

$$p_{max} = \frac{2P}{\pi l_0 a_l}. \tag{2.52}$$

Within the contact area, the stresses are $\sigma_x = \sigma_y = -p(x)$; they become zero at the surface outside the contact.

The principal stress along the z axis are expressed as (Johnson (1987)):

$$\frac{\sigma_x}{p_{max}} = -\frac{1}{a_l}\left[\left(a_l^2+2z^2\right)\left(a_l^2+z^2\right)^{-\frac{1}{2}}-2z\right], \tag{2.53}$$

$$\frac{\sigma_y}{p_{max}} = -\frac{\nu}{a_l}\left[\left(a_l^2+2z^2\right)\left(a_l^2+z^2\right)^{-\frac{1}{2}}-2z\right]-a_l\nu\left(a_l^2+z^2\right)^{-\frac{1}{2}}, \tag{2.54}$$

$$\frac{\sigma_z}{p_{max}} = -a_l\left(a_l^2+z^2\right)^{-\frac{1}{2}}, \tag{2.55}$$

and the principal shear stress is

$$\frac{\tau}{p_{max}} = -\frac{1}{a_l}\left[z-z^2\left(a_l^2+z^2\right)^{-\frac{1}{2}}\right]. \tag{2.56}$$

2.3.2.3 Subsurface stress

Away from the symmetric axis the stresses can be predicted in many ways such as in the work of Hamilton and Goodman (1966), Hamilton (1983), Sackfield and Hills (1983). McEwen (1949) proposed the expression for the stress components for a general point (x,z)(the load is evenly distributed along the y axis except at the very edge of the roller) as:

$$\frac{\sigma_x}{p_{max}} = -\frac{1}{a_l}\left[\hat{m}\left(1+\frac{z^2+\hat{n}^2}{\hat{m}^2+\hat{n}^2}\right)-2z\right], \tag{2.57}$$

$$\frac{\sigma_y}{p_{max}} = -\frac{\nu}{a_l}\left[\hat{m}\left(1+\frac{z^2+\hat{n}^2}{\hat{m}^2+\hat{n}^2}\right)-2z\right]-\frac{\nu\hat{m}}{a_l}\left(1-\frac{z^2+\hat{n}^2}{\hat{m}^2+\hat{n}^2}\right), \tag{2.58}$$

$$\frac{\sigma_z}{p_{max}} = -\frac{\hat{m}}{a_l}\left(1-\frac{z^2+\hat{n}^2}{\hat{m}^2+\hat{n}^2}\right), \tag{2.59}$$

$$\frac{\tau_{xz}}{p_{max}} = -\frac{\hat{n}}{a_l}\left(\frac{\hat{m}^2-z^2}{\hat{m}^2+\hat{n}^2}\right), \tag{2.60}$$

where $\hat{m}$ and $\hat{n}$ are dimension parameters the sign of which coincides with that of z and x, respectively. They are expressed as (McEwen (1949), Johnson (1987)) :

$$\hat{m}^2 = \frac{1}{2}\left\{\left[\left(a_l^2-x^2+z^2\right)^2+4x^2z^2\right]^{\frac{1}{2}}+\left(a_l^2-x^2+z^2\right)\right\}, \tag{2.61}$$

$$\hat{n}^2 = \frac{1}{2}\left\{\left[\left(a_l^2-x^2+z^2\right)^2+4x^2z^2\right]^{\frac{1}{2}}-\left(a_l^2-x^2+z^2\right)\right\}. \tag{2.62}$$

This model can be used for those contacts involving a tangential friction component over the contact area, which is often the case in practice, just need to add a friction coefficient component of $\mu(x-\hat{n})$, $\mu\hat{n}\frac{z^2-\hat{m}^2}{\hat{m}^2+\hat{n}^2}$, $\mu(\hat{m}-2z)+\mu\hat{m}\frac{z^2+\hat{n}^2}{\hat{m}^2+\hat{n}^2}$ for equations (2.57), (2.59) and (2.60)(McEwen (1949)), respectively.

Equations (2.57), (2.58), (2.59) and (2.60) can be used to check whether the material has the plastic deformation or not.

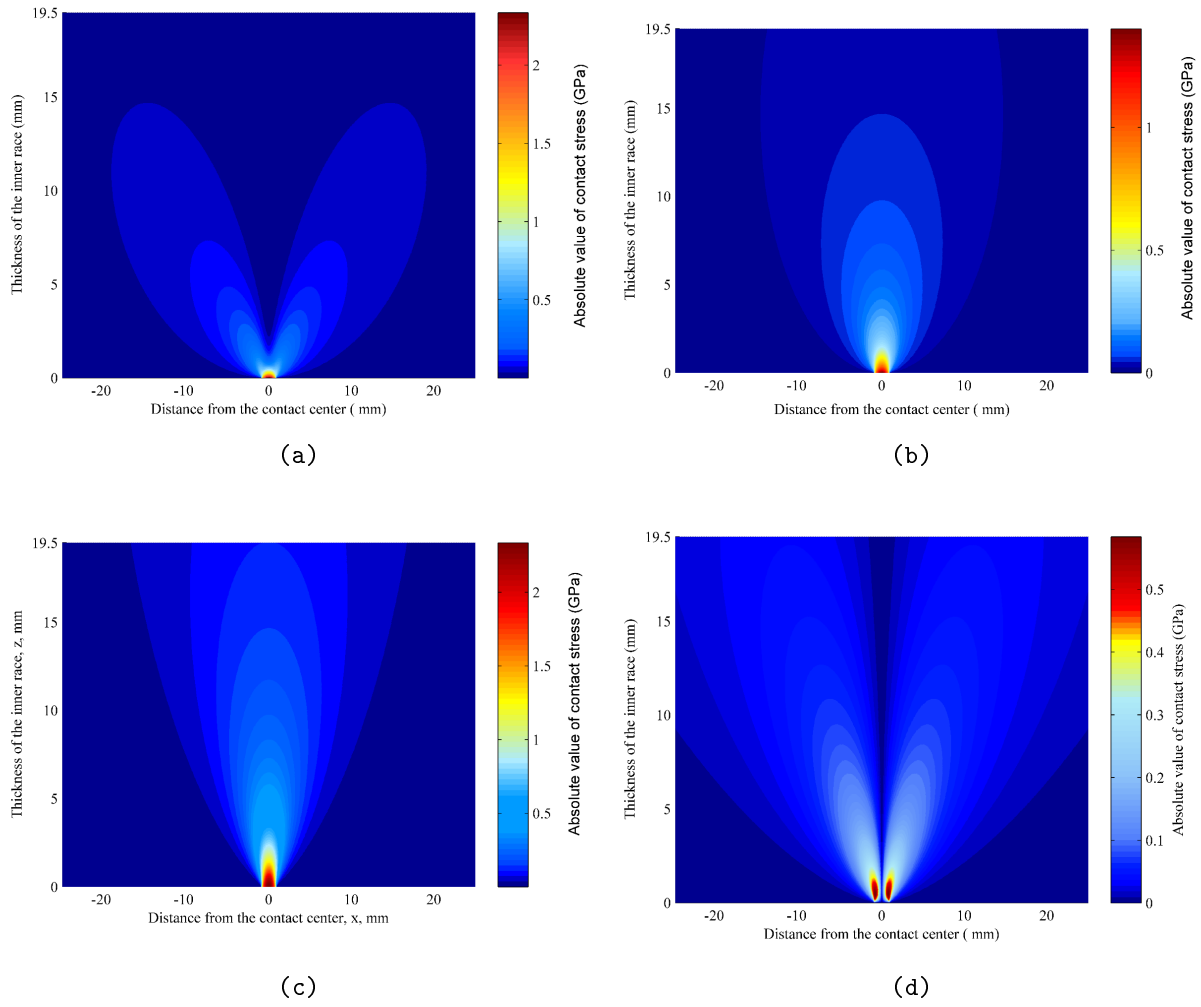


Figure 2.7. *The distribution of stresses in the inner race for the most loaded roller-race contact of the roller bearing, with the bearing load of 980 kN; (a) σ_x, (b) σ_y, (c) σ_z and (d) τ_{xz}. The x axis is in the roller rolling direction, the y axis is in the roller length direction and the z axis is directed into the inner race.*

As an example for illustration, the distribution of the stresses in the inner race through the thickness is calculated when it is loaded by a roller (on the load line). The inner race has a thickness of 19.5 mm with a width of 108 mm, and the roller has a diameter of 54 mm with an effective length of 79 mm; other parameters of the roller bearing will be detailed in Chapter 4. The maximum contact pressure at the contact centre of the most loaded contact is is 2.35 GPa when the bearing is under a radial load of 980 kN. The corresponding subsurface stresses can be calculated from equation (2.57) - (2.60). This has been done in Matlab and the results are shown in figure 2.7. It can be seen that the stresses are only distributed in a quite small contact area, about 2 mm in the rolling direction. Except the shear

stress, all the stresses achieved the maximum at the contact center. They reduce rapidly from the maximum value at the contact surface through the thickness. The shear stress achieve the minimum at the contact center and the maximum at the edges of the contact. It reduce rapidly from the maximum value at the contact surface (contact edges) through the thickness.

The maximum shear stress is half of the difference between principal stresses which can be calculated using equations (2.57), (2.59) and (2.60). The contours of maximum shear stress is given in figure 2.8 when the bearing is under a radial load of 980 kN. The maximum shear stress is 0.7 GPa, which take place at the depth of 0.72 mm (that is equal to $0.78a_l$). The inner ring of this roller bearing has a quite high yield stress, the normal value is ~1.7 GPa for 100Cr6 in EuropeLelikov (2009). According to the Tresca's maximum shear stress criterion Johnson (1987), the onset of plastic yield occurs when the maximum shear stress reaches the half of the yield stress which is 0.85 GPa for the 100Cr6 in this case. Therefore, the inner race should only have elastic deformation when the roller bearing is under a radial load of 980 kN.

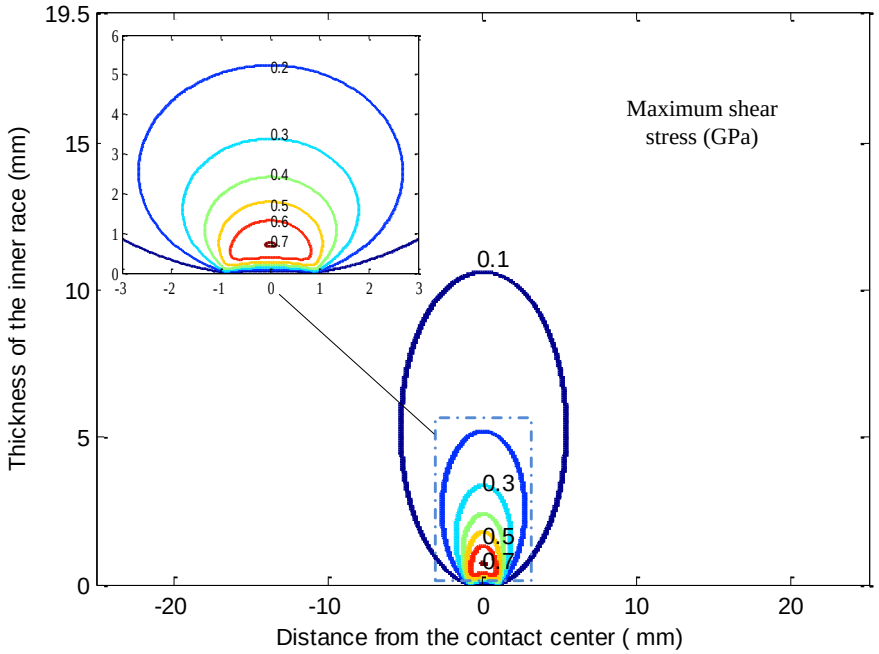


Figure 2.8. *Contours of maximum shear stress τ_{max} in the inner race for the most loaded roller-race contact of the roller bearing, with the bearing load of 980 kN.*

2.3.2.4 Surface deflection

For an ideal line contact formed between an infinite roller and an infinite plane with the same material, the maximum contact deformation has been determined by (Lundberg and Sjövall (1958)) :

$$\delta_{max} = \frac{2P_max\left(1-v^2\right)}{\pi El} \ln\left[\frac{\pi El^2}{P_max\left(1-v^2\right)\left(1 \mp \frac{D}{d_m}\right)}\right]. \tag{2.63}$$

To avoid the stress concentration at roller ends, they are normally crowned in practice. For these crowned rollers, Palmgren (1959) developed a formula for the contact deformation on the basis of laboratory testing a range of crowned rollers loaded against raceways, which can be expressed as :

$$\delta_{max} = 3.84 \times 10^{-8} \frac{P_{max}^{0.9}}{\left(l_0 \times 1000\right)^{0.8}}. \tag{2.64}$$

Once the approach of the two mating bodies is achieved, the surface displacement for points lying within the contact area can be obtained from

$$U_1(x) + U_2(x) = \delta_{max} - \frac{1}{2R_x}x^2. \tag{2.65}$$

2.4 Lubrication of rolling element bearings

2.4.1 An introduction to fluid film lubrication

In rolling element bearings, the lubrication has a remarkable influence on the bearing durability and load capacity. A lubricant is always required to separate balls or rollers from the raceways to avoid metal-to-metal contact. If the lubricant film fails, high friction and noise take place and wear and seizure can occur. Besides, the lubricant also functions to cool the bearing to maintain a proper temperature, prevent the bearing from contaminants and prevent surface corrosion of balls or rollers and bearing races.

The film thickness of the lubricant is determined by the lubricant viscosity, the geometry feature of the mating surfaces, and load and speed conditions. Ideally, the lubricant film should be thick enough to separate the mating surfaces but not too thick to increase unnecessary work for overcoming the shear effect in the lubricant (Dwyer-Joyce et al. (2003)). The following sections aim to provide the reader with an brief understanding of the lubrication of rolling element bearings.

2.4.2 Viscosity of lubricants

Most lubrication is making use of a fluid film which can be from mineral oils, synthetic esters, silicone fluids, water or gas. The thickness of the fluid film in bearings can be ranging from less than a micron thick for non-conformal contact in ball or roller bearings to hundreds of microns thick for conformal contact in journal bearings and thrust bearings.

The viscosity of a lubricant is a significant parameter as indicates the lubricant's internal friction resist to flow. A higher viscosity means harder to flow, for example, water flows more easily compared with honey as the latter has a higher viscosity.

2.4.2.1 Absolute viscosity of lubricants

The absolute viscosity of a liquid was introduced by Newton in 1687, which is stated by Hamrock and Anderson (1983) and can be expressed as :

$$\eta = \frac{\frac{F}{A}}{\frac{u}{h}} = \frac{\tau}{\dot{\gamma}} \tag{2.66}$$

where τ and $\dot{\gamma}$ are the shear stress and shear strain rate, which are the ratio between the friction force F and the lubricated area A , and the ratio between the relative surface velocity u and the film thickness h, respectively. The SI unit for the viscosity is $\mathrm{N}\cdot\mathrm{s}\cdot\mathrm{m}^{-2}$ or $\mathrm{Pa}\cdot\mathrm{s}$ and a common unit employed in engineering is centipoise (cP) which is a thousandth of $\mathrm{N}\cdot\mathrm{s}\cdot\mathrm{m}^{-2}$.

2.4.2.2 Kinematic viscosity of lubricants

The kinematic viscosity, which is more convenient in many applications, is defined as the ratio of absolute viscosity to its density :

$$\eta_k = \frac{\eta}{\rho} \tag{2.67}$$

where ρ is the lubricant density. The unit of kinematic viscosity is then $\mathrm{m}^2\cdot\mathrm{s}^{-1}$; however, a more commonly used unit is centistoke (cSt) which is a millionth of $\mathrm{m}^2\cdot\mathrm{s}^{-1}$.

The viscosity of a fluid is not a constant; it increases with pressure but decreases with temperature. The effect of pressure and temperature on the lubricant viscosity will be briefly discussed in the following subsections.

2.4.2.3 Pressure effect on the absolute viscosity

The dependence of the viscosity of lubricants on pressure is proposed by Barus (1893) as :

$$\ln \frac{\eta}{\eta_0} = \xi p \tag{2.68}$$

where η_0 is the absolute viscosity at atmospheric pressure with constant temperature, $\mathrm{N \cdot s \cdot m^{-2}}$, ξ is the pressure-viscosity coefficient of the lubricant depends on the temperature, $\mathrm{m^2 \cdot s^{-1}}$, and p the pressure, $\mathrm{N \cdot m^{-2}}$, . The equation (2.68) is only suitable to moderate pressure applications.

Roelands (1966) developed a formula from a experimental study for the isothermal condition, as :

$$\log \eta + 1.2 = (\log \eta_0 + 1.2) \left(1 + \frac{p}{2000}\right)^{Z_1} \tag{2.69}$$

where attention is required for units; centipoise cP used for absolute viscosity and $\mathrm{kg \cdot cm^{-2}}$, for pressure. Z_1 is a dimensionless viscosity-pressure index.

The pressure-viscosity coefficients in equation (2.68) and the dimensionless viscosity pressure index in equation (2.69) for some common fluids has been determined by Jones et al. (1975).

2.4.2.4 Temperature effect on the absolute viscosity

Mineral and synthetic oils are most common lubricants in engineering applications and the viscosity of them decreases with increasing temperature. Roelands (1966) described the temperature effect on the absolute viscosity by the following equation :

$$\log (\log \eta + 1.2) = -S_0 \log \left(1 + \frac{T}{135}\right) \log G_0 \tag{2.70}$$

where η for the absolute viscosity in centipoise (cP) and T for temperature in in degree Celsius (°C); G_0 and S_0 are dimensionless constants.

2.4.2.5 Combined pressure-temperature effect on the absolute viscosity

For applications involving with both pressure and temperature variations which are more practical, the absolute viscosity of lubricants associated with the combined pressure-temperature

effect was proposed by Roelands (1966) as:

$$\log \eta + 1.2 = G_0 \frac{\left(1+\frac{p}{2000}\right)^{-C_2 \log\left(1+\frac{T}{135}\right)+D_2}}{\left(1+\frac{T}{135}\right)^{S_0}} \tag{2.71}$$

where η for the absolute viscosity in centipoise (cP), T for temperature in degree Celsius (°C) and pressure p in kilograms per square centimeter ($\mathrm{kg \cdot cm^{-2}}$); G_0, S_0, C_2 and D_2 are parameters determined from experiment.

2.4.3 Density of lubricants

Compared to the viscosity, the change in density with pressure and temperature are relative small for a mediate pressure range, which is considered as a incompressible lubricant. However, pressure in rolling element bearings are commonly extremely high, of the order of $\sim$ 1GPa, where the lubricant is compressible. The pressure effect on the lubricant density has been investigated by many researchers.

The compressibility of the lubricant is described by a dimensionless density, and the expression from Dowson and Higginson (1966) is suitable for low pressure, less than 0.4 GPa, expressed as :

$$\bar{\rho} = \frac{\rho}{\rho_0} = 1 + \frac{0.6p}{1+1.7p} \tag{2.72}$$

Bridgman (1949) derived the expression from Tait equation and experimentally validated the compressibility expression :

$$\bar{\rho} = \frac{\rho}{\rho_0} = \frac{1}{(1-1/(1+B_0^{'})ln[1+p/B_0(1+B_0^{'})])} \tag{2.73}$$

Murnaghan (1944) derived the compressibility equation for isothermal condition as :

$$\bar{\rho} = \frac{\rho}{\rho_0} = \left(1+\frac{B_0^{'}}{B_0}p\right)^{\frac{1}{B_0^{'}}} \tag{2.74}$$

Later, Jacobson and Vinet (1987) developed a formula for the pressure-density from experiments with pressure up to 2.2 GPa, the expression is:

$$p = \frac{3B_0}{\bar{\rho}^{-\frac{2}{3}}}\left(1-\rho^{-\frac{1}{3}}\right)e^{\eta\left(1-\bar{\rho}^{-\frac{1}{3}}\right)} \tag{2.75}$$

Hamrock et al. (1987) investigated six base oil behaviors when the pressure was over their

solidification pressure (denoted by p_s) and proposed a pressure density formula as :

$$\bar{\rho} = \begin{cases} \frac{\rho}{\rho_0} = \frac{1}{1-C_1 p^2 - C_2 p} & \text{p} < \text{p}_\text{s} \\ \frac{\rho}{\rho_0} = \frac{1}{1-C_3 p + C_4} & \text{p} > \text{p}_\text{s} \end{cases} \tag{2.76}$$

2.4.4 Lubrication Regimes

Depending on the intervening film, four lubrication regimes can be defined by the well known Stribeck curve (Stribeck (1902)) which describes relationship between the friction coefficient and the dimensionless Hersey number. The Hersey number is given as

$$H_s = \frac{\omega_r \eta}{p} \tag{2.77}$$

where ω_r is the rotational speed in revolution per second (rps), η is the absolute viscosity in Pascal second (Pa·s) and p is pressure in Pascal (Pa).

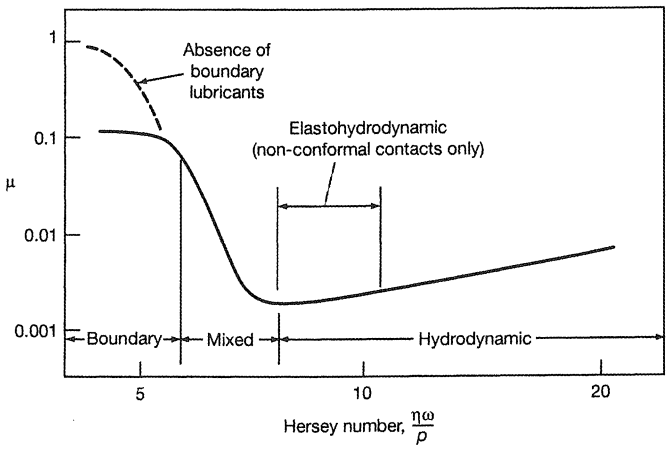


Figure 2.9. *A Stribeck curve which illustrates the effects of Hersey number on friction coefficient (adapted from Hamrock et al. (2004)).*

2.4.4.1 Boundary lubrication

In figure 2.9, the first lubrication regime is the boundary lubrication where the Hersey number is so low that rarely real lubricant film is developed. Since the asperity contact dominant in this regime, the friction coefficient is quite high (~ 0.1). This can be observed in applications involving low rotational speed, low lubricant viscosity and high applied load, for instance, most machines will see boundary lubrication during their start-up or shutdown stages.

2.4.4.2 Mixed lubrication

With increasing Hersey number, a lubricant film is built up to partially separate the two contacting surfaces and this case is referred as mixed lubrication. The load in the contact is shared by contacts between asperities and the fluid film, then the friction coefficient is determined by their contributions. In this regime, the friction coefficient can achieve its minimum when the film thickness is thick enough to effectively separate two surface. Mixed lubrication is encountered in contacts between surfaces having roughness similar to the thickness of lubricant.

2.4.4.3 Hydrodynamic lubrication

Hydrodynamic lubrication is not identified until the Hersey number becomes considerable large (over 10 in figure 2.9). Since the opposing surfaces are completely separated by the pressurized fluid film, the friction coefficient remains low (less than 0.01) in this regime although it increases slowly. Again, this can be attributed to the excessive work need to drag a comparable large amount of oil to the contact area. The hydrodynamic lubrication is commonly seen in journal bearings.

2.4.4.4 Elastohydrodynamic lubrication

The elastohydrodynamic lubrication regime can only be observed in non-conformal contacts which is a common case encountered in rolling element bearings. As the thickness of the fluid film increases, the contact surfaces are fully separated from each other and the load is only accommodated by the thin sandwiched film. The friction coefficient is considerable low although there is a slight increase with the Hersey number. This is caused by slightly increasing fluid drag on the moving surfaces; however, this effect is not significant.

2.4.4.5 Specific film thickness for different regimes

As the lubrication regime depends on the thickness of the fluid film and the roughness of the contacting surfaces, a dimensionless film parameter, Λ, the specific film thickness is introduced by Hamrock. It is defined as the ratio of the lubricant film thickness to the composite surface roughnessHamrock and Dowson (1981), given as :

$$\Lambda = \frac{h_{min}}{\left(R_{q1}^2 + R_{q2}^2\right)^{1/2}} \tag{2.78}$$

where h_{min} is the minimum film thickness, R_{q1} and R_{q2} are the rms surface roughness of contact surfaces. Then the dimensionless film parameter for the corresponding four regimes are roughly estimated as :

- Boundary lubrication, $\Lambda < 1$
- Mixed lubrication, $1 < \Lambda < 5$
- Elastohydrodynamic lubrication, $3 < \Lambda < 10$
- Hydrodynamic lubrication, $5 < \Lambda < 100$

2.5 Review of rolling bearings performance monitoring techniques

Although a single bearing cost is low, there are tremendous expense associated with bearing failures when they cause long shutdown time for the involved system. Studies (IAS motor reliability working group (1985)) in motor reliability has showed that bearing failures accounted for over 40 percent of these machine failures. There are many techniques can be used for monitoring bearing failure and they can be classified into vibration measurement, wear debris analysis, acoustic emission measurement, lubricant film thickness measurement and load measurement.

2.5.1 Vibration monitoring

The bearing vibration mechanism and corresponding noise have been explained by researchers (Choudhury and Tandon (1998), Meyer et al. (1980), Sunnersjö (1978), Tallian and Gustafsson (1965)), it is attributed to the varying compliance and/or the defects. Even for a new bearing without imperfectness, vibration takes place when it is under radial load cause the bearing load is shared by a finite number of balls or rollers. Different rolling elements are brought into the load zone during bearing rotation, which results in a periodical variation of bearing stiffness and vibrations (Sunnersjö (1978), Tallian and Gustafsson (1965)) ; meanwhile, the direction of the contact force changes with the rolling elements position and ring-mode vibrations are generated (Choudhury and Tandon (1998); Meyer et al. (1980)).

Although the new perfect bearing has vibration, a defect enlarges the vibration to a much higher extent, which makes vibration analysis to be the most widely used method for detecting bearing faults. Bearings under normal operation conditions commonly start fatigue

failure from subsurface where the maximum shear stress occurs. It gradually propagates to the surface until fragments of metal breaks away leaving a small pit or spall which results in detectable vibrations and increasing noise levels (Eschmann et al. (1958)). There are many vibration analysis techniques and they can be roughly classified into two groups: time domain and frequency domain.

2.5.1.1 Analysis in the time domain

The crest factor is defined as the ratio of the peak value to the root mean square, and it is a simple measure of how extreme the peaks in the signal in time domain are. Tandon and Choudhury (1999) stated that this method had limited success for defects diagnostic.

Then Kurtosis value was used by Dyer and Stewart (1978) to detect bearing defects. It is a mathematical description of amplitude probability deviation from the Gaussian distribution, i.e., the spread of amplitude distribution, given by

$$Kurtosis = \frac{\int_{-\infty}^{+\infty} (x-\bar{x})^4 A(x)\,\mathrm{d}x}{\sigma^4} \tag{2.79}$$

where $A(x)$ is the probability density function (PDF) of x , $\bar{x}$ is the mean and σ is the standard deviation. They stated that the Kurtosis value for a new rolling element bearing with Gaussian distribution was approximately 3, and the magnitude rose when the bearing got defects. This method has been also successfully applied to detect bearing defects by Prabhu (1996) while other researchers (Mathew and Alfredson (1984)) argued its reliability to detect the incipient damage. The advantage of this method is it requires minimum historical signal information, however, it is limited by the fact that the Kurtosis value for a bearing with well advanced defects reduces to the level for a undamaged bearing.

A short duration pulse is generated whenever the rolling elements come into contact with a surface having defects or with debris in the bearing and the pulse excites the natural frequencies of surrounding bearing components. As a consequence, an increase in vibration at these frequencies is observed. Balderston (1969) found that the frequency of structure resonance, excited by impulsive loading caused from bearing defects, was rather high, in the order of 10 kHz. Therefore, bandpass filtering techniques have been used to extract the high frequency components that associated with bearing faults.

The shock pulse method as an example of making use of this principle has been developed. It uses an accelerometer having a resonant frequency of 32 kHz and the transducer is excited by the shock pulse generated by the impacts in the bearing. Through a bandpass filter, the low frequency components caused by other sources are filtered out and the remain-

der is compared with a reference value which is determined by the bearing bore diameter and speed. The difference between the measured value (caused by defects) and the reference value (caused by surface roughness of undamaged bearing) indicates health condition of the defective bearing.

A number of researchers (Alfredson and Mathew (1985); Kim (1984); Mathew and Alfredson (1984); Prabhu (1996)) have reported the success of this method in sensing the faults in rolling bearings. The shock pulse method is so simple that it only requires only two input parameters, the bearing diameter and speed, and outputs a single value which indicates the bearing health condition, which means it has found widespread use in industry, such as the shock pulse meter from SKF and the spike energy from IRD.

2.5.1.2 Analysis in the frequency domain

Spectral analysis approaches have been popular in vibration analysis. By using fast Fourier transform, the spectrum of vibration signal can be obtained in which both low and high frequency component can be useful to detect bearing defects.

It is the presence of the defect corresponding characteristic rotational frequency in the low frequency range that is used in the bearing defects detection. These defect characteristic frequencies can be calculated from references (Broderick et al. (1972); Gustafsson and Tallian (1962); Igarashi and Hamada (1982); Prasad (1987); Tandon and Nakra (1992)), which are commonly in the order of hundreds of Hertz (less than 500 Hz under normal operation conditions). Many reports can be found on successfully detecting bearing defects by finding these characteristic frequencies (Igarashi and Hamada (1982); Tandon and Nakra (1993); Taylor (1980)). However, some researches (Broderick et al. (1972); Osuagwu and Thomas (1982)) have found it is a challenge to identify these defect-associated frequencies for small defects cause they may be masked by vibrations from other sources.

Advanced signal processing techniques have been developed with the aim of improving the effectiveness of spectral analysis. High frequency resonance technique or envelope analysis is a powerful signal processing approach in vibration spectral analysis, which has find wide use in detecting bearing faults since 1970s. A short duration pulse is generated when the defective surface contact with its mating component, resulting in resonance at the defect characteristic frequencies, i.e., the resonances are amplitude modulated(AM) at these characteristic frequencies. The high frequency resonance technique involves a sequence of signal operations on vibration signal to extract the resonances of an influence associated with the bearing defect.

Many researchers (Burchill (1973); Darlow and Badgley (1975); Darlow et al. (1974); Harting (1978); Martin and Thorpe (1992); Su and Lin (1992)) adopted this technique for

early bearing fault detection and analysis and a review has been reported by McFadden and Smith (1984b). McFadden and Smith (1984a) developed the classical model for describing the vibration excited by a single point defect on the inner ring of a rolling bearing. The procedure starts with acquiring a vibration signal from the bearing or bearing housing, followed by applying bandpass filter to the signal, and the filtered signal is then demodulated by an envelope detector in which the signal is rectified and smoothed by a low pass filter in order to exclude the carrier resonant frequency. By using fast Fourier transform, the spectrum of the signal is obtained and a line at the defect corresponding characteristic frequencies is observed. In this model, the impacts produced while the rolling elements strike the defect is circumferential evenly distributed in the bearing and the vibration is periodic.

The limitation of high frequency resonance technique is that associated frequencies for advanced defects may be not large enough to isolate them from the increasing background level of the spectrum (Burgess (1988); McFadden and Smith (1984b)). Therefore, the wavelet transform has been adopted by many researchers to extract weak features in vibration signals (Luo et al. (2003); Mori et al. (1996); Prabhakar et al. (2002); Qiu et al. (2006); Rubini and Meneghetti (2001); Yang et al. (2005)). For instance, Mori et al. (1996) have made use of discrete wavelet transform to study the impulsive responses in vibration whenever balls strike the defective outer raceway to predict the spalling and Prabhakar et al. (2002) have extended the success in predicting both the inner race defects and the outer race defects of a ball bearing.

2.5.2 Oil wear debris monitoring

Oil wear debris analysis is one useful diagnostic tool for monitoring rolling element bearings health because bearing component wear generate debris in the lubrication system. The most interesting features in wear particle analysis are quantity, size, composition and morphology because they imply the severity, rate, type and source of wear. Figure 2.10 shows the relationship between wear particle features and wear characteristics (Roylance and Hunt (1999)). Several techniques have been utilized for wear particle analysis, including spectrometric oil analysis, particle counting, magnetic plug and ferrography.

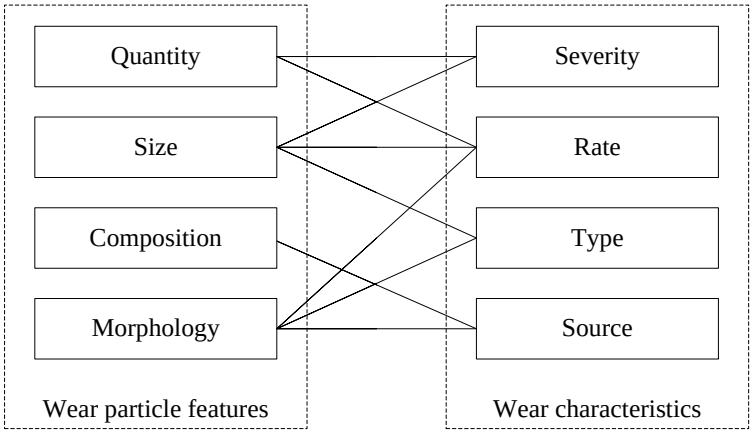


Figure 2.10. *The relationship between wear particle features and wear characteristics, reproduced from Roylance and Hunt (1999).*

Spectrometric oil analysis has been routinely used for wear particle analysis in aviation industry to monitor rotating components. This technique can identify and quantify particles in oil samples less than 10 microns in size. Wang and Zhang (2005) developed a model for predicting the remain service life for aircraft engines on the basis of spectrometric oil analysis, which assumed the predicted residual life was proportional to the wear severity.

Magnetic inductive debris sensors have been utilized by researchers (Chambers (1988); Dempsey et al. (2006); Flanagan et al. (1990); Miller and Kitaljevich (2000)) and particles with size between 100 to 1000 microns can be detected. Besides particle quantity, they can also give the size distribution and morphology information as well as composition. The success in detecting the size and number of debris particles by using inductance type oil debris sensors has been reported by Hunt (1993). Miller and Kitaljevich (2000) developed an in-line, full flow oil debris monitoring system for rolling bearings in aircraft jet, turbofan and turboprop engines by using an inductive sensor which was capable to detect both ferromagnetic and non-ferromagnetic metallic particles. Dempsey et al. (2006) used an inductance type oil debris sensor to monitor the failure progression on a tapered roller bearing used for helicopter transmission. The measurement system in Dempsey's work consists of target lubricant samples (off line), magnetic chip detectors which capture the ferrous particles in the lubricant, and inductive debris sensors which examines the disturbance effect of debris particles passage on the magnetic field online. They found that the accumulated oil debris mass was able to indicate bearing defects.

Roylance (2005) compared these techniques in terms of their ability of observing particle features and ferrography outlines itself from other wear particle analysis methods, since

ferrography can provide all particle features and also separate debris having size range from 1-100 microns (size range from 20- 80 microns depicts the transition from benign to active wear). The first documented paper for this technique was reported in 1972 and Seifert and Westcott (1972) developed a device for investigating wear mechanisms of a ball bearing failure in a jet engine by examining lubricants. Since then, ferrography technique has been widely used for wear particle analysis.

Monitoring wear debris in lubricants is an informative method from which the mode, source and severity of wear in rolling element bearings can be known. However, the limitation of this method is very time consuming and requires expert examination of wear debris compared with vibration analysis which has a number of software packages allowing for automated analysis.

2.5.3 Acoustic emission monitoring

Acoustic emission refers to the generation of transient elastic waves caused by a rapid release of strain energy resulted from a structure change (deformation or damage) in a solid material (Mathews (1983); Pao et al. (1979)). When a structure is subjected to a sudden change in pressure, load or temperature, stress waves are produced and they propagate to the material surface and are detected by the transducer. Plastic deformation, crack initiation and growth are common sources of acoustic emission in rolling element bearings. Therefore, acoustic emission has been widely used for cracks detection.

Acoustic emission monitoring has the advantage of detecting small cracks beneath the surface; by contrast, vibration analysis cannot detect them until they propagate to the surface. The maximum shear stress in the mating components occurs beneath the surface, as a consequence, subsurface cracks are more likely to be initialed. In such a case, acoustic emission method can detect these cracks before they can be detected by the vibration accelerator.

Many researchers have reported the success in making use of acoustic emission technique to detect bearing defects. One of the earliest documented acoustic emission measurement is carried out by Balderston (1969) to identify artificially seeded defects in rolling bearings. This technique later was adopted by Rogers (1979) on a anti-friction bearing used for slewing cranes in offshore gas production with low speed. He successfully identified detectable acoustic emission sources in bearings, such as rubbing of the crack faces, impacts between rolling elements and defective surfaces in the loaded zone. Yoshioka and Fujiwara (1982) found that acoustic emission technique was able to give earlier disclosure of bearing defects compared with vibration analysis from experiment on a thrust loaded ball bearing, and that positions of acoustic emission sources corresponded to positions of failures caused

by rolling contact fatigue. Choudhury and Tandon (2000); Tan (1990); Tandon and Nakra (1990) studied acoustic emission counts and peak values on bearings with different sizes and found that the number of acoustic counts was lower for bearings without defects than those with defects. However, they found that the number of acoustic emission event relies on the predetermined threshold value.

Shiroishi et al. (1997) used both an accelerometer and an acoustic emission sensor to detect bearing defects with a signal analysis routine combining the high frequency resonance technique with an adaptive line enhancer technique. They stated the peak ratio of the signal (the ratio of the sum of the peak values at the defect frequency harmonics to the average value of the spectrum) well indicated the presence of the defect, and established the relationship between the acceleration peak value and the defect width. More recent correlation between defect size and acoustic emission parameters can be found in references (Al-Dossary et al. (2009); Al-Ghamd and Mba (2006); Elforjani and Mba (2006, 2010)).

2.5.4 Lubricant film thickness monitoring

The thickness of the lubricant layer in rolling bearing contacts is a significant parameter as it controls the friction, fatigue and thus wear of rolling element bearings. Ideally, the lubricant should be just thick enough to separate the mating surfaces but not to cause unnecessary work by overcoming churning losses. In rolling element bearing, the film thickness is commonly less than a micron. A large wealth of investigations have been reported on lubricant film thickness measurement. Basically, they can be grouped into three classes according to the method they utilized: electrical measurement, optical measurement and active ultrasonic measurement.

Film thickness can be measured by obtaining the lubricant electrical properties such as resistance, capacitance and inductance. El-Sisi and Shawki (1960) measured the film resistance at a known thickness when it was subjected to specified potential and developed the relationship between resistance and thickness. This was used in turn to indicate the thickness once the resistance was measured.

Astridge and Longfield (1967) measured capacitance of oil films to predict the thickness in a nominal point contact formed between a large radius disc and ring machine. Safa et al. (1982-1983) measured the pressure with thin-film manganin pressure transducers in lubricated contacts, and the temperature with a resistance transducer and also the film thickness with a capacitive transducer. The measured film thickness generally agree with prediction from the well known Dowson and Higginson's equation. Lubecheck from SKF was a developed as a result of the capacitance technique.

Using either resistance or capacitance of the film requires electrical isolation of the op-

posing surfaces and temperature calibration. Also, the sensor needs to be directly in contact with oil film which may involve mounting sensor on one of opposing surfaces, which may disrupt the profile of contact surface or oil film (Dwyer-Joyce et al. (2004)).

An eddy current transducer is an example of using inductance. Glavatskih et al. (2001) used an eddy current sensor to measure the oil film thickness in a pad thrust bearing under different operation conditions. In this investigation, they considered temperature compensation for the sensor by using the simultaneous temperature data. The principle of the eddy current sensor is that a crystal oscillator with high oscillating frequencies of 1 MHz is used to excite the senor coil to produce a vibrating electromagnetic field where eddy current can be induced in any conductive component surface; the interaction between the two electromagnetic fields (from the sensor coil and from the component surface) generates a voltage in the circuit which depends upon the distance between the sensor coil and the component.

Optical methods have been proved to be very powerful to measure thin film thickness in elastohydrodynamic lubrication. This technique has been adopted by researchers since the early 1960s in model contacts, with which they can investigate the film distribution in the contact and film thickness-dependent parameters (such as the applied load, the operational speed and the material properties). This method was first used by Kirk (1962), followed by Cameron and Gohar (1966) and Gohar (1971). Cooper and Moore (1994) used the optical means to measure elastohydrodynamic film thickness formed between a transparent flat plate and a super-finished steel ball; the oil thickness was down to 10 nanometers. They found that the measured data agreed well with prediction from (Hamrock and Dowson (1977)) and the measured data was then used to derive the pressure-viscosity coefficient. Johnston et al. (1991) utilized optical interferometry with a solid spacer (by coating a thin layer of transparent solid on the glass disc) to investigate very thin film in a point contact between a glass disc and a steel ball.

Optical technique gained great success in the elastohydrodynamic lubrication development, however, it has inherent limitations as each method does. It requires one of the opposing subject in contact to be transparent, which is unrealistic in industry service equipment and limits its application mainly to laboratory studies.

2.5.5 Active ultrasonic monitoring

As an effective non-destructive method, ultrasound has been widely used in cracks detection and recently has been used for measurements in rolling contacts. Drinkwater and Dwyer-Joyce made use of ultrasonic reflections from dry and lubricated contacts to study the stress and the film thickness. The principle is simple: an ultrasonic wave will be partly reflected back and partly transmitted when it strikes on boundary between two media, and the ratio

of the reflected pulse to the incident pulse, refereed to as the reflection coefficient, is determined by the acoustic properties of media on either side, the contact stiffness and the frequency of waves.

Ultrasound reflections was used by Drinkwater et al. (1996) to determine the stiffness of a dry partially contact interface between two aluminum surfaces. During the experiment, they observed the plastic flow effect on the reflection coefficient during the first loading cycle even the pressure in the material did not exceed its yield limit. They stated that measured stiffness for all load cases agreed reasonable well with predictions, within an order of theoretical values. Dwyer-Joyce et al. (2009) measured the concentrated contact between wheel and rail by using an ultrasonic array to measure reflections from the contact interface. Two-dimensional contact maps under several load conditions were achieved and compared with the results from a pressure sensitive film; it was found that measurements from ultrasonic method corresponded well to that from pressure film method.

In the case when the wavelength of ultrasound is less than the oil thickness, two reflections from the oil layer will be observed (from the upper and lower surface) and the film thickness can be obtained by measuring the time difference between the two reflections provided the ultrasound speed in the oil is given. This is commonly known time-of-flight method and it is schematically shown in figure 2.11. Using the ToF method to measure thick film has been reported by Witting et al. (1994) and Kamei and Serizawa (1998).

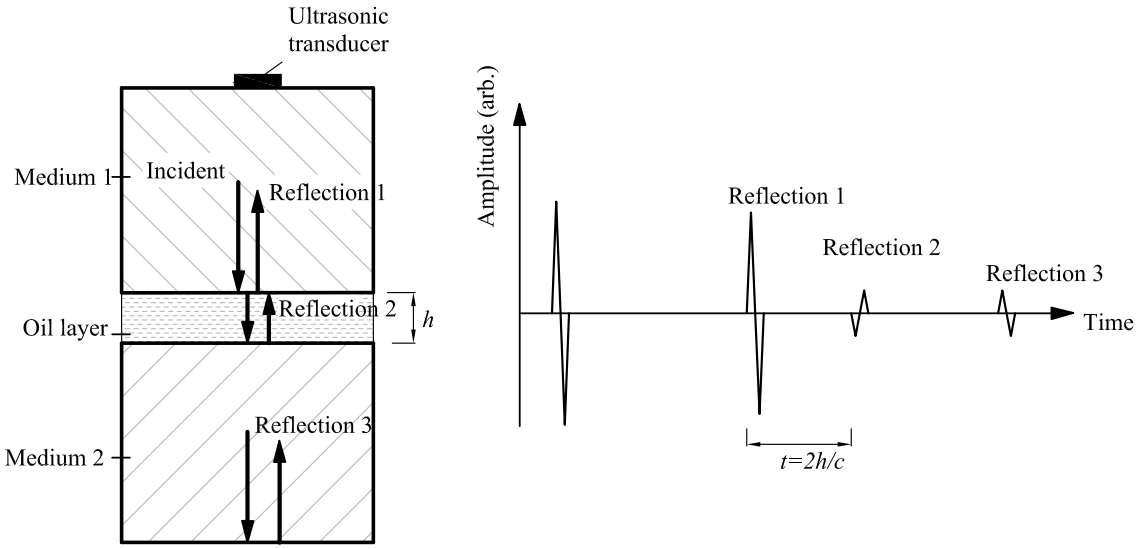


Figure 2.11. *Schematic of oil measurement using the ultrasonic ToF method*

However, when the oil film is not thick enough, which is usually the case in bearing contacts, these two reflections overlap each other in the time domain. As an alternative,

resonance method and spring method are developed for measuring thin film which cannot be measured by using the ToF method. Two corresponding models, the continuum model and the spring model, have been developed for interpreting ultrasonic data to film thickness.

The continuum model is based on using the through-thickness resonances to measure the film thickness and has been described in many reports. Pialucha and Cawley (1994) used a longitudinal ultrasonic wave to detect thin layers in a three layer system (the two surrounding media are considerable thick) and built up a relationship between reflection coefficient and frequency. In the continuum model, the resonance frequency is inversely proportional to the thickness, that means high frequency transducers are necessary to resonant very thin oil film. By way of an example, the first resonance frequency for a oil layer of 20 microns is 36.3 MHz and it rises to 363 MHz for an oil layer of 2 microns. In practice, the frequency of ultrasonic transducers is commonly less than 60 MHz because of excessive cost and increasing attenuation for higher frequency waves in steels.

The spring model is an alternative for the case of very thin film thickness. In fact, the oil film thickness in rolling element bearings under normal operation conditions is in the sub-micron range. This model was initially used by Tattersall (1973) and Schoenberg (1980) to study the ultrasonic response at a perfect and imperfect bonded interface between two elastic media. Later, Hosten (1991); Hosten et al. (1987) studied the ultrasonic wave reflection/transmission at interfaces in a water-composite-water system which depends on the stiffness of the thin composite layer.

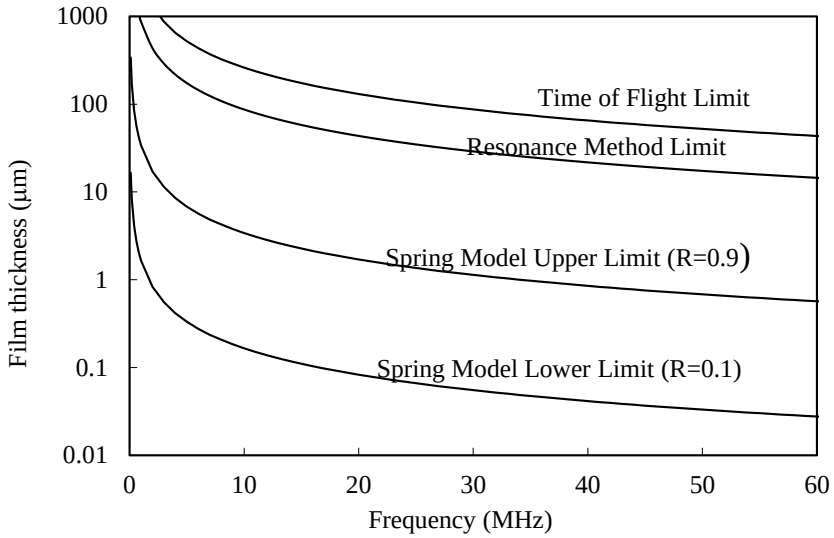


Figure 2.12. *Film thickness measurement limits of three ultrasonic methods: the ToF, the continuum (resonance) and the spring method, reproduced from Dwyer-Joyce et al. (2003)*

The oil layer between the inner/outer ring and rolling element in a rolling element bearing forms a solid-oil-solid system where the oil layer is very thin and behaves like a spring (having extremely high stiffness) when interacts with ultrasound. Dwyer-Joyce and coworkers used ultrasonic reflection method coupled with the spring model for measuring oil film thickness both in hydrodynamic (Dwyer-Joyce et al. (2004); Kasolang et al. (2013); Reddyhoff et al. (2005)) and elastohydrodynamic (Dwyer-Joyce et al. (2003, 2006); Zhang et al. (2006)) lubrication. They compared the reflection coefficient for an oil layer between steels predicted from both the continuum model and the spring model, and found that the results agree well at frequencies lower than the oil resonance frequency (Dwyer-Joyce et al. (2003)). Also, the limitations for film thickness measurement by using the ToF, resonance and spring method are studied and the results are given in figure 2.12.

2.5.6 Load monitoring

Capacitive sensors have been reported with success in measuring the stress of a rolling ring via its deformation by Pottier et al. (2011). A groove was made on the upper of the ring to deposit the capacitive probe, the probe and the groove bottom surface of the ring were the two electrodes of the capacitive sensor. The deformation of the bearing ring was measured by the capacitive sensor which behaved like a displacement sensor. Elastic deformation

took place in the machined ring when it is loaded, the groove bottom surface of the ring constituted the mobile electrode of a capacitor which has a varying gap. The capacitance is a function of the gap distance, the surface area of the electrodes and the permittivity of the dielectric between the electrodes. Therefore, the capacitive probe converts the distance change into a voltage, using a charge amplifier based electronic device. The load applied in the contact was achieved by measuring change in the output voltage which was due to the ring deflection. One important requirement of applying this technique is to groove the bearing ring, weakening the bearing assembly.

Later, Rasolofondraibe et al. (2012) instrumented a capacitive sensor on the bearing housing instead of modifying the bearing ring. The probe and the outer surface of the bearing ring constituted the two electrodes of the sensor, and the varying distance between them was treated as the deformation of the ring. Experiments have been done on a roller bearing under either a static or a dynamic radial load. The measured results were reported to consistent well with the applied load. However, one disadvantage of this method is that the capacitive sensor measured the distance between the bearing outer surface and the probe which is attached on the bearing housing, and hence the bearing vibration plays a significant role in the measured results. In addition, the capacitive sensor gives an average result over the electrodes width which is relative large (12 mm) compared with the actual contact width (<1 mm).

2.6 Summary

Basics of rolling element bearings are introduced in this chapter. The bearing load is shared by a number of rolling element and raceway contacts, and the load in these contacts are determined by the bearing load, the number of the rolling elements, the bearing geometry and properties. For each contact, the shape, size, stress and deformation can be calculated on the basis of the Hertz theory.

A brief introduction of bearing lubrication is provided as lubricants are commonly used to reduce the friction and wear in bearings. The viscosity and density of lubricants are important parameters for the bearing lubrication, both of which change with temperature and pressure. Characterized by the oil film thickness and the surface roughness, there are four different lubrication regimes which are commonly referred as boundary lubrication, mixed lubrication, hydrodynamic lubrication and elastohydrodynamic lubrication.

A wealth of techniques have been developed for bearing performance monitoring and a brief review of these techniques is given in this chapter. Vibration monitoring, oil debris analysis and acoustic emission techniques gains great success in detecting bearing defects

but only after these defects present. However, there is little report about direct measurement of load transmitted by rolling elements to a bearing raceway before bearing failure occurs. Load information is significant to predict bearing remaining service life and to schedule corresponding maintenance service. Therefore, it is necessary to develop a technique to direct measure load transmitted in rolling bearing contacts.

Chapter 3

Ultrasonic Response at Boundaries and Acoustoelastic Effect

This chapter describes the principles that govern how ultrasonic waves interact with a boundary between two different media in terms of the acoustic impedance. The response of an ultrasonic wave at a boundary is determined by the acoustic impedance of the media, the stiffness of the boundary and the frequency of the wave. For a liquid layer between two surfaces, three typical methods used to determine the layer thickness are introduced with the main concern of the spring model, which represents the ultrasonic response from a thin film in bearing contacts in this work. Phase change in the reflected wave is investigated as it affects the ultrasonic ToF determination. The acoustoelastic effect, the variation of ultrasound speed in a material with load, is introduced and it relates the ultrasonic ToF to load in the next chapter.

3.1 Ultrasonic transducer

An ultrasonic transducer is a device that is used to generate and/or receive ultrasonic waves. Figure 3.1 gives a common configuration of an ultrasonic transducer in the market. It includes a piezoelectric element, a curved lens help with focusing, backing material, a wear plate, a connector and a case housed all components inside. Here we mainly considered the piezoelectric element as it is the heart of the transducer which actually converts electrical signal into mechanical vibration and vice verse, other elements are detailed in the reference (Papadakis (1999)).

The piezoelectric elements, cut from quartz crystals or large piezoelectric ceramics, have aligned dipoles after polarization. The elements could be cut into discs, rings, plates or

cylinders, and also in different directions to generate different types of wave (longitudinal or shear wave). The material consists of many randomly oriented grains and has no piezoelectric effect before polarization. The crystal structure phase changes as the material is heated over its Curie temperature. If a strong dc electric field is applied during the cooling process, a net polarization will be obtained. The material is then permanently polarized in the direction of the electric field and piezoelectric effect can be observed.

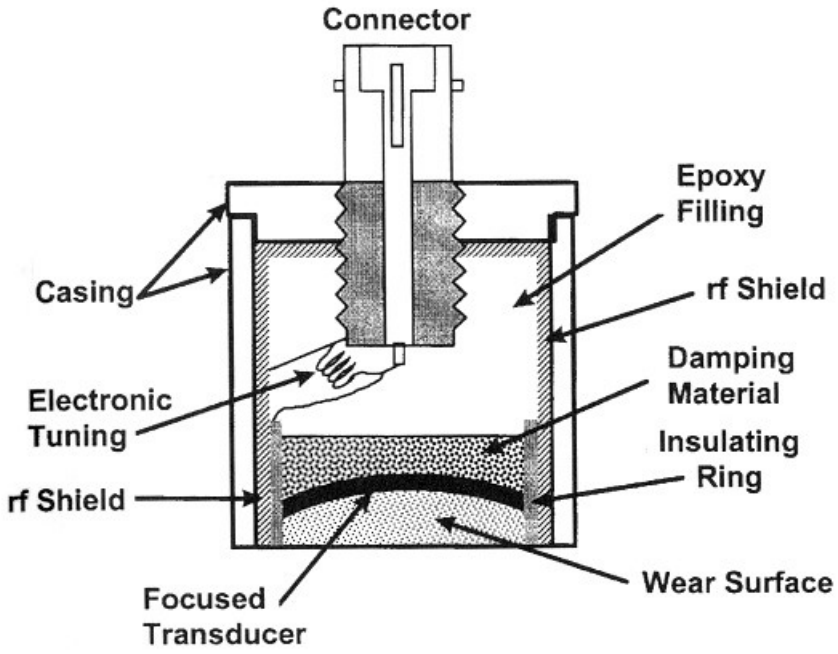


Figure 3.1. *Schematic of a typical commercial focusing ultrasonic transducer (Papadakis (1999))*

A pair of electrodes are formed by coating the piezoelectric elements with conducting materials like gold or silver. These elements can then be excited by applying an electric field through the material. The polarized molecules are aligned by the electric force, resulting in the crystal oscillation. The maximum oscillation is achieved when the frequency of the electric field coincides with one of the natural frequencies of vibration of a particular axis.

The piezoelectric elements can be operated at a resonant frequency in two modes: through thickness resonance mode and through length resonance mode. In the through thickness resonance mode, the resonance take places in the direction perpendicular to the radiating faces; in the through length resonance mode, the resonance occurs paralleling with the radiating faces (Blitz (1963)). In this work, a high frequency was required and the piezoelectric elements were excited to provide through thickness resonance. In this mode, the central frequency of the polarized piezoelectric sensor is inversely proportional to the thickness.

Normally, the wavelength of the resonant wave in a thin element is twice of the thickness. The central frequency is related to the element thickness by the following equation

$$f = \frac{c}{2h}, \tag{3.1}$$

where f is the natural central frequency of the piezoelectric element, c is the ultrasound speed in the material, and h is the thickness of the element.

The piezoelectric elements can be directly bonded on the specimen and function as ultrasonic sensors, although the energy of an ultrasonic pulse is distributed in a narrow range of frequencies. A layer of backing material can be used to improve the response. This is the case for all ultrasonic sensors used in this work.

3.2 Ultrasonic response at a plane boundary between two media

When an ultrasonic wave is normally incident at a plane boundary formed between two elastic bodies, part of the energy is reflected back into the first medium and part of the energy is transmitted into the second medium. Figure 3.2 schematically shows the reflection and transmission of a longitudinal wave normally propagating through a perfect bonded plane interface. The proportion of the incident wave to be reflected and transmitted through the boundary is determined by the acoustic properties of the surrounding media. The acoustic impedance is defined as the product of the material density and the speed of ultrasonic wave in the material (Kräutkramer and Kräutkramer (1977)); similar to the electrical and mechanical impedance, it is a measure of acoustic resistance ability of a material.

In general, the displacement field of a longitudinal plane wave propagating in x-direction with a sinusoidal form in a non-absorbent medium can be expressed as :

$$u(x,t) = A\sin(2\pi ft - k_w x) = Ae^{j(2\pi ft - k_w x)}, \tag{3.2}$$

where u is the displacement filed in the medium in x-direction, A is the amplitude, f is the wave frequency and k_w is the wavenumber ($k_w = \frac{2\pi}{\lambda}$).

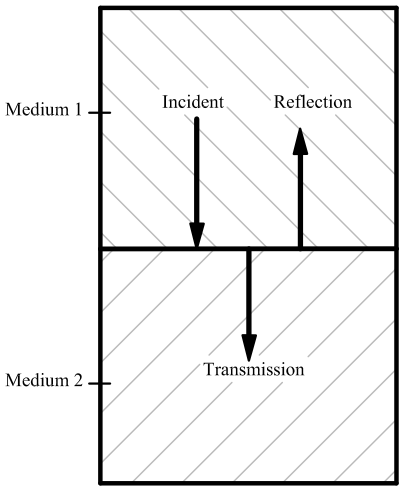


Figure 3.2. *Reflection and transmission of longitudinal waves normally incident to a perfect bonded plane boundary separating two media*

If the displacement of the incident, reflected and transmitted wave at the boundary are designated by u_I, u_R, u_T, and the corresponding amplitudes are A_I, A_R and A_T , then

$$u_I(x,t) = A_I e^{j(2\pi ft - k_{w1}x)}, \tag{3.3}$$

$$u_R(x,t) = A_R e^{j(2\pi ft + k_{w1}x)}, \tag{3.4}$$

$$u_T(x,t) = A_T e^{j(2\pi ft - k_{w2}x)}. \tag{3.5}$$

The ultrasonic wave incidents and transmits in the positive while reflects in the negative direction of x, as shown by figure 3.2, and the displacement field in medium 1 is a superposition field of both the incident and the reflected wave. Therefore, the displacement in medium 1 and medium 2 can be written as :

$$u_1(x,t) = A_I e^{j(2\pi ft - k_{w1}x)} + A_R e^{j(2\pi ft + k_{w1}x)}, \tag{3.6}$$

$$u_2(x,t) = A_T e^{j(2\pi ft - k_{w2}x)}. \tag{3.7}$$

By using the definition of the strain $\varepsilon = \frac{\partial u}{\partial x}$ and the strain-stress relationship $\sigma = E\varepsilon$, the

stress field in media 1 and 2 can be expressed as :

$$\sigma_1(x,t) = -jE_1k_{w1}A_Ie^{j(2\pi ft-k_{w1}x)} + jE_1k_{w1}A_Re^{j(2\pi ft+k_{w1}x)}, \tag{3.8}$$

$$\sigma_2(x,t) = -E_2jk_{w2}A_Te^{j(2\pi ft-k_{w2}x)}. \tag{3.9}$$

Since the material acoustic impedance, Z, is defined as the product of the density and the speed of sound in the material ($Z = \rho c$), and the elastic modulus is related to the acoustic impedance by $E = \rho c^2 = Zc$ (Kräutkramer and Kräutkramer (1977)), equation (3.8) and (3.9) becomes :

$$\sigma_1(x,t) = E_1\frac{\partial u_1(x,t)}{\partial x} = -j2\pi fZ_1A_Ie^{j(2\pi ft-k_{w1}x)} + j2\pi fZ_1A_Re^{j(2\pi ft+k_{w1}x)}, \tag{3.10}$$

$$\sigma_2(x,t) = E_2\frac{\partial u_2(x,t)}{\partial x} = -j2\pi fZ_2A_Te^{j(2\pi ft-k_{w2}x)}, \tag{3.11}$$

where E_1, E_2 are modulus of elasticity for media 1 and 2, respectively.

Provided the boundary is located at $x = 0$, (the medium 1 locates at $x \geqslant 0$ while the medium 2 locates at $x \leqslant 0$), as shown in figure 3.2. For a perfect bonded interface, both the displacement field and stress field are continuous, expressed as :

$$u_1(x,t)|_{x=0} = u_2(x,t)|_{x=0}, \tag{3.12}$$

$$\sigma_1(x,t)|_{x=0} = \sigma_2(x,t)|_{x=0}. \tag{3.13}$$

Substituting equation (3.6) and (3.7) into equation (3.12) gives

$$A_I + A_R = A_T, \tag{3.14}$$

and substituting equation (3.10) and (3.11) into equation (3.13) gives

$$(A_I - A_R)Z_1 = A_TZ_2. \tag{3.15}$$

By combining equations (3.14) and (3.15), amplitude ratios are related to material acoustic impedance by:

$$\frac{A_R}{A_I} = \frac{Z_1 - Z_2}{Z_1 + Z_2}, \tag{3.16}$$

$$\frac{A_T}{A_I} = \frac{2Z_1}{Z_1 + Z_2}. \tag{3.17}$$

The amplitude ratio of the reflected wave to the incident wave is defined as reflection coefficient and the ratio of the transmitted wave to the incident wave is transmission coefficient (Schoenberg (1980)), denoted by R and T, respectively.

$$R = \frac{A_R}{A_I} = \frac{Z_1 - Z_2}{Z_1 + Z_2}, \tag{3.18}$$

$$T = \frac{A_T}{A_I} = \frac{2Z_1}{Z_1 + Z_2}. \tag{3.19}$$

As shown by equations (3.18) and (3.19), the passage of an ultrasonic wave at the perfectly bonded plane boundary is totally determined by the acoustic impedance of contacting media. It is frequency independent which means an ultrasonic wave with either low or high frequency can be used for a perfect plane boundary. Reflection coefficient reaches the minimum value of zero and transmission coefficient becomes to unity if media 1 and 2 are identical in terms of acoustic impedance.

From equation (3.18), it is noted that reflection coefficient is positive when an ultrasonic wave propagates from a medium with large acoustic impedance to a medium with small acoustic impedance, and it is negative when the wave travels in the opposite way.

For the former case where a wave incidents from medium 1 having larger acoustic impedance with respect to medium 2, reflection coefficient approaches its maximum value of unity, which means most energy of the ultrasonic wave reflects back into medium 1. For instance, an ultrasonic wave propagates to a steel-air interface will be totally reflected (the acoustic impedance of steel is 46,600,000 $\mathrm{Pa \cdot s \cdot m^{-3}}$) while air acoustic impedance is 430 $\mathrm{Pa \cdot s \cdot m^{-3}}$). This method is widely used in ultrasonic distance measurements and flaw detections.

For the latter case when a wave incidents from medium 1 having considerable small acoustic impedance to medium 2, reflection coefficient approaches -1, which means most energy of the ultrasonic wave transmits into medium 2. This is due to the fact that a phase change of π radians takes place in the displacement and thus the velocity of the ultrasonic wave. This can be observed for a coupling layer-specimen boundary, such as a water-steel interface.

Figure 3.3 schematically shows an example of utilizing an ultrasonic transducer for distance measurement or flaw detections in solids. Ultrasound travels in the specimen object until meets a boundary and the reflected wave is received by the same transducer. If the specimen have no cracks, the first reflection occurs at the back face of the solid as the pres-

ence of solid-air interface, however, if the object has a crack (as shown by the dashed line), the first reflection at the discontinuity presents in time history before the reflection from the end, as shown by the dashed pulse in the figure. By measuring the ToF, the distance that ultrasound traveled can be easily achieved provided the ultrasound speed in the material is known.

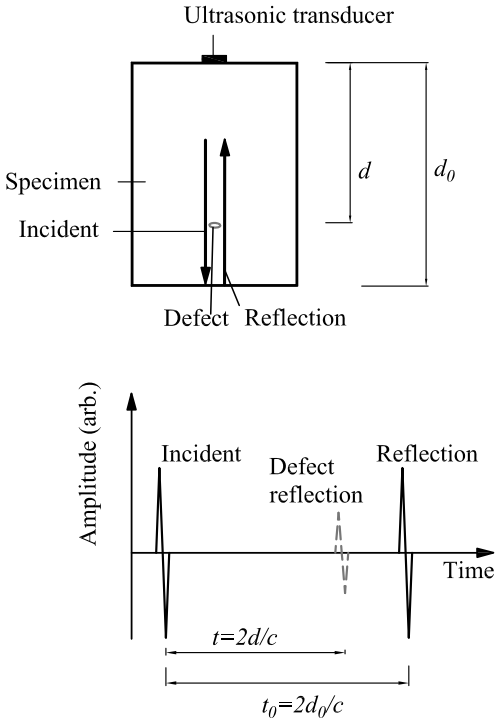


Figure 3.3. *Schematic of thickness measurement or flaw detection in a solid with an ultrasonic transducer.*

The acoustic impedance of some common used materials are given in table 3.1 and the reflection coefficient for a perfect bonded plane boundary between different materials can be calculated from equation (3.18) and the results are given in table 3.2.

Table 3.1: Acoustic impedance of some commonly used materials (Blitz (1963))

Material	Longitudinal wave speed ($m \cdot s^{-1}$)	Density ($kg \cdot m^{-3}$)	Acoustic impedance ($x10^6 kg \cdot m^{-2} \cdot s^{-1}$)
Aluminum	6400	2700	17.28
Cooper	4700	8900	41.83
Brass	3500	2700	9.45
Steel	5900	7900	46.61
Gold	3700	10500	38.85
Silver	3200	19300	61.76
Nylon	2700	1140	3.08
Perspex	2700	1200	3.24
PZT5A1	7750	4000	31.00
Water	1500	1000	1.50
Mineral oil	1450	877	1.27
Air	330	1.3	0.00043

Since air has a sufficient low acoustic impedance relative to solids and liquids, it is impossible to transmit ultrasonic waves from the transducer through air into solids. Thus, a medium is always necessary to couple the transducer to the test specimen. For the boundary between the coupling medium and test specimen, the transmission coefficient can be calculated from equation (3.19). In order to propagate more energy from the transducer, the coupling medium is required to have similar acoustic impedance of the examined subject.

Table 3.2: Reflection for perfect boundaries between some commonly materials

Material	Amplitude of reflection coefficient
Steel-Aluminum	0.46
Steel-Cooper	0.045
Steel-Brass	0.21
Steel-Nylon	0.878
Steel-Perspex	0.87
Steel-Water	0.937
Steel-Oil (Shell T68)	0.945
Steel-Air	0.999981

3.3 Ultrasonic response at a thin layer separating two media

This section is going to describe ultrasonic waves propagate through a three-layer system where two media is separated by a layer of liquid, which is the normal case for tribological contacts with lubricant film. There are three methods available for film thickness determination by using ultrasound: ToF method, resonance method and spring method.

3.3.1 ToF method

When the film layer is sufficiently thick or the ultrasonic wavelength is small enough, an ultrasonic wave is reflected at both the top and the bottom surface of the layer, giving two discrete reflected peaks in time domain. As such a case, reflection coefficient at the two surface of the film is determined by the material acoustic impedance, see equation (3.18). The layer thickness can be obtained from the ultrasonic ToF and the speed in the film layer by using the following equation.

$$h = \frac{ct}{2}. \tag{3.20}$$

As the film becomes thinner, the ToF method becomes less accurate since the reflected signals are too close in the time trace and it becomes difficult to distinguish. For a very thin layer, the two reflected signals may be overlapped in time domain makes the identification becomes quite challenge. The minimum film thickness that the ToF method can measure is determined by the frequency of ultrasonic wave, which has been depicted by figure 2.12 in Chapter 2.

Under such a case that the oil layer is so thin that the ToF method is inappropriate to determine film thickness, resonance method (also known as continuum model) and spring method are utilized as alternatives.

3.3.2 Resonance method

Figure 3.4 schematically shows ultrasonic wave propagation in a typical three-layer system. A plane boundary is formed on each side of the thin layer which is denoted by medium 3 in the figure, the reflection and the transmission take place at each interface. Again, the reflection and transmission at a boundary depends on the mismatch between the acoustic impedance of the material on each side.

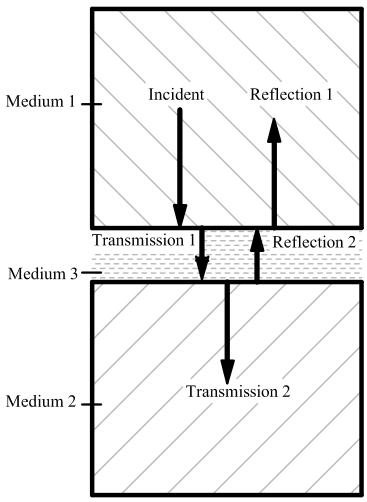


Figure 3.4. *Schematic diagram of the response of the ultrasonic wave in a three-layer system*

Let the first boundary (between media 1 and 3) locates at $x = 0$ and the thickness of the thin layer is h , i.e., the second boundary (between media 3 and 2) locates at $x = h$, then the

displacement field of media 1, 2, 3 can be written as :

$$u_1(x,t) = A_I e^{j(2\pi ft - k_{w1}x)} + A_{R1} e^{j(2\pi ft + k_{w1}x)}, \tag{3.21}$$

$$u_2(x,t) = A_{T2} e^{j(2\pi ft - k_{w2}x)}, \tag{3.22}$$

$$u_3(x,t) = A_{T1} e^{j(2\pi ft - k_{w3}x)} + A_{R2} e^{j(2\pi ft + k_{w3}x)}, \tag{3.23}$$

where k_{w3} is the wave number in the third medium. Using the definition of the strain and the strain-stress relationship, the stress in each medium can be written as :

$$\sigma_1(x,t) = -j2\pi f Z_1 A_I e^{j(2\pi ft - k_{w1}x)} + j2\pi f Z_1 A_{R1} e^{j(2\pi ft + k_{w1}x)}, \tag{3.24}$$

$$\sigma_2(x,t) = -j2\pi f Z_2 A_{T2} e^{j(2\pi ft - k_{w2}x)}, \tag{3.25}$$

$$\sigma_3(x,t) = -j2\pi f Z_3 A_{T1} e^{j(2\pi ft - k_{w3}x)} + j2\pi f Z_3 A_{R2} e^{j(2\pi ft + k_{w3}x)}. \tag{3.26}$$

Using equations (3.23) and (3.26), the displacement and stress at boundary 1 and 2 can be expressed as that of the medium 3. For the upper interface at $x = 0$,

$$u_3(0,t) = A_{T1} e^{j(2\pi ft)} + A_{R2} e^{j(2\pi ft)}, \tag{3.27}$$

$$\sigma_3(0,t) = -j2\pi f Z_3 A_{T1} e^{j(2\pi ft)} + j2\pi f Z_3 A_{R2} e^{j(2\pi ft)}, \tag{3.28}$$

and for the lower interface at $x = h$,

$$u_3(h,t) = A_{T1} e^{j(2\pi ft - k_{w3}h)} + A_{R2} e^{j(2\pi ft + k_{w3}h)}, \tag{3.29}$$

$$\sigma_3(h,t) = -j2\pi f Z_3 A_{T1} e^{j(2\pi ft - k_{w3}h)} + j2\pi f Z_3 A_{R2} e^{j(2\pi ft + k_{w3}h)}. \tag{3.30}$$

Resonance method involves applying an ultrasonic wave, the frequency of which is equal to one of the resonance frequencies of the thin layer, to a multilayer system. The success of using this approach to determine a thin intermediate layer has been reported by many researchers (Hosten (1991); Kinra et al. (1994); Pialucha and Cawley (1994)). Continuum models have been established, with which the ultrasonic wave response from a multilayer system can be illustrated. In this model, the two boundaries are perfectly bonded with sur-

rounding elastic bodies, which means the stress and displacement are continuous at boundaries, which gives :

$$u_1(0,t) = u_3(0,t), \tag{3.31}$$

$$\sigma_1(0,t) = \sigma_3(0,t), \tag{3.32}$$

$$u_3(h,t) = u_2(h,t), \tag{3.33}$$

$$\sigma_3(h,t) = \sigma_2(h,t). \tag{3.34}$$

The resonance frequency is determined by the layer thickness and the wave speed in the layer, which can be expressed as (Pialucha et al. (1989)) :

$$f_{res} = \frac{c_3 k_n}{2h}, \tag{3.35}$$

where k_n is the resonance number.

3.3.2.1 Odd resonant mode

If k_n is an odd number here, $k_{w3}h = k_n\pi$ ($k_{w3} = \frac{2\pi}{\lambda_3}$), then $e^{j(2\pi ft \pm k_{w3}h)} = -e^{j(2\pi ft)}$, equations (3.29) and (3.30) are reduced to :

$$u_3(h,t) = -A_{T1}e^{j(2\pi ft)} - A_{R2}e^{j(2\pi ft)}, \tag{3.36}$$

$$\sigma_3(h,t) = j2\pi f Z_3 A_{T1}e^{j(2\pi ft)} - j2\pi f Z_3 A_{R2}e^{j(2\pi ft)}. \tag{3.37}$$

By comparing equations (3.36), (3.37) with equations (3.27), (3.28), it is found the displacement and stress at two sides of the thin layer are related by :

$$u_3(h,t) = -u_3(0,t), \tag{3.38}$$

$$\sigma_3(h,t) = -\sigma_3(0,t). \tag{3.39}$$

By making use of equation (3.31) and (3.33), and substituting equations (3.21) and (3.22) into (3.38), the following relationship will be obtained when the common exponential

factor is canceled :

$$A_{T2}e^{j(-k_{w2}h)} = -(A_I + A_{R1}). \tag{3.40}$$

By making use of equation (3.32) and (3.34), and substituting and equations (3.24) and (3.25) into (3.39), the following relationship will be obtained (again the common exponential factor is canceled) :

$$Z_2 A_{T2}e^{j(-k_{w2}h)} = Z_1(-A_I + A_{R1}). \tag{3.41}$$

The reflection and transmission coefficient can be obtained from equation (3.40) and (3.41), written as :

$$R_{res} = \frac{A_{R1}}{A_I} = \frac{Z_1 - Z_2}{Z_1 + Z_2}, \tag{3.42}$$

$$T_{res} = \frac{A_{T2}}{A_I} = \frac{-2Z_1}{Z_1 + Z_2} e^{j(k_n \pi \frac{c_3}{c_2})}. \tag{3.43}$$

3.3.2.2 Even resonant mode

If k_n is an even number here, then $e^{j(2\pi ft \pm k_{w3}h)} = e^{j(2\pi ft)}$ and equations (3.29), (3.30) become to :

$$u_3(h,t) = A_{T1}e^{j(2\pi ft)} + A_{R2}e^{j(2\pi ft)}, \tag{3.44}$$

$$\sigma_3(h,t) = -j2\pi f Z_3 A_{T1}e^{j(2\pi ft)} + j2\pi f Z_3 A_{R2}e^{j(2\pi ft)}. \tag{3.45}$$

By comparing equations (3.44), (3.45) with equations (3.27), (3.28), it is found the displacement and stress at two sides of the thin layer are equal.

$$u_3(h,t) = u_3(0,t), \tag{3.46}$$

$$\sigma_3(h,t) = \sigma_3(0,t). \tag{3.47}$$

Reflection and transmission coefficient can be achieved by substituting equations (3.21) and (3.22) into (3.46), and substituting equations (3.24) and (3.25) into (3.47), along with using the boundary condition equations (3.31), (3.33), (3.32) and (3.34). The results are same to equations (3.42) and (3.43).

It is observed that reflection and transmission coefficients are independent of the acoustic impedance of the thin intervening layer when the frequency of the incident wave is equal to one of layer resonant frequency (either odd or even mode). Reflection coefficient achieves a minimum at resonance frequencies and this value approaches zero when the surrounding media have identical acoustic impedance. By using equation (3.35), resonance frequency can be calculated for a thin layer having a specific thickness.

As an example, the reflection coefficient spectrum for a thin oil layer between two steel elements are given in figure 3.5. The thickness of oil layer changes from 0.1 to 20 microns and the frequency range is limited to 60 MHz. It is observed that for a oil film thinner than 15 microns the first resonance frequency is higher than 60 MHz which is the practical upper limit for ultrasonic test with steel (restricted by the ultrasonic energy loss in the hosting medium).

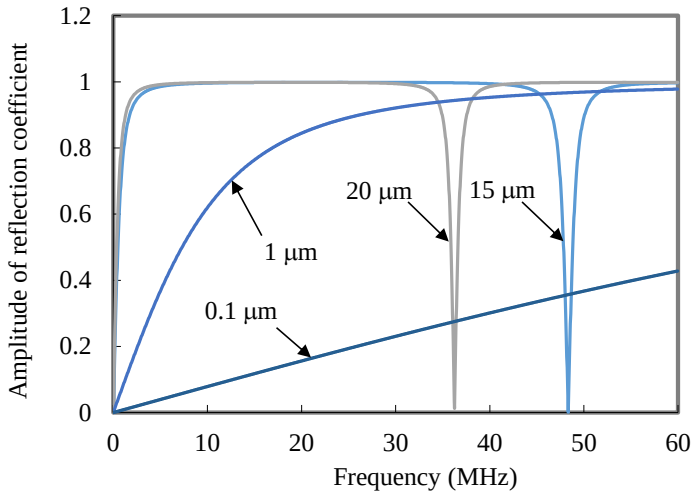
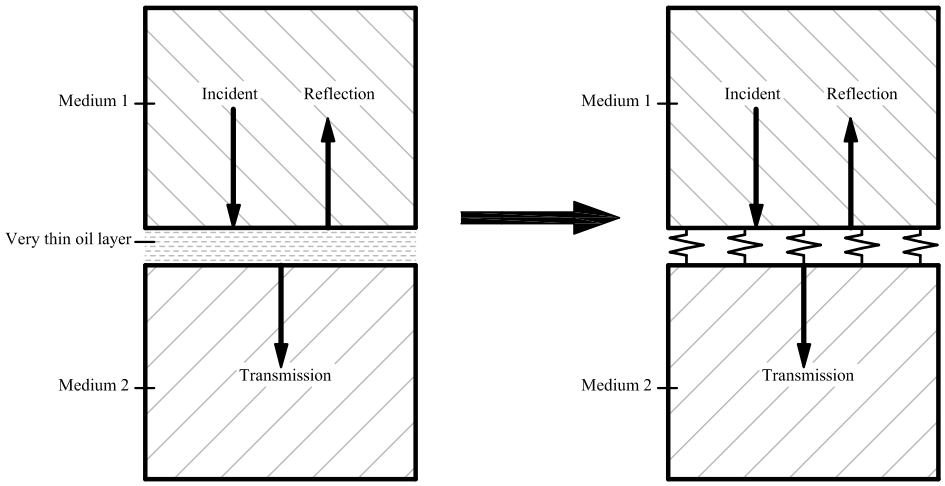


Figure 3.5. *Amplitude of reflection coefficient spectrum for a layer of oil with varying thickness between steel components (predicted from continuum model by Dwyer-Joyce et al. (2003))*

3.3.3 Spring method

When the intermediate layer is very thin, of the order of a few microns, and the wavelength of the ultrasonic wave is large, typically several hundred times larger than the layer thickness, then the thickness of the layer is negligible and this thin layer can be modeled as a series of springs with negligible mass. This is known as spring model (Schoenberg (1980);

Tattersall (1973)). Figure 3.6 schematically illustrates a lubricated tribological contact with ultrasound propagation and the thin oil layer is modeled by springs.



`Figure 3.6.` *A lubricated tribological contact with a very thin oil layer can be represented by a spring model*

Spring model can be used for the following cases:

- A thin liquid layer between two considerable large elastic bodies, such as a lubricant layer separating rolling elements and bearing rings. The success of using spring model to measure oil film thickness (less than 15 microns) has been reported (Dwyer-Joyce et al. (2003)).

- A thin layer of solid, like the adhesive layer in joints. In this case, both the normal and the transverse stiffness of the intermediate layer are related to the layer thickness and its elastic constants (Pialucha (1992)).

- Partially solid-solid contact without liquid, like two rough surfaces pressed together. As the presence of the surface roughness, the two surfaces contact with each other at a wealth number of asperities in the micro level. Drinkwater et al. (1996) investigated the stiffness of the contact interface between two rough surfaces under load, found that spring model can effectively predict the stiffness of the contact interface formed between most machined practical rough surfaces.

- Imperfect bond layer between the adhesive and the adherend with small unbounded areas. In these partial bonds, the size and the density of the disbond area can be statistically described and then related with the stiffness of the interface; this method has been reported by many authors, for instance, Angel and Achenbach (1985).

The displacement in media 1 and 2 are the same as equation (3.21) and (3.22) shown and the stress in media 1 and 2 can be expressed as (3.24) and (3.25). In this model, the very thin oil layer is represented by a number of springs, considered as a non-perfect bonded interface across which the stress is continuous but the displacement is not. The stress (i.e. the spring force) can be expressed in terms of stiffness and the difference in displacement:

$$\sigma_3(t) = K\left(u_2(h,t) - u_1(0,t)\right) \approx K\left(u_2(0,t) - u_1(0,t)\right). \tag{3.48}$$

By making use of the boundary condition $\sigma_1(0,t) = \sigma_3(t) = \sigma_2(0,t)$, the reflection coefficient can be derived as :

$$R = \frac{Z_1 - Z_2 + i\left(\dfrac{2\pi f}{K}\right) Z_1 Z_2}{Z_1 + Z_2 + i\left(\dfrac{2\pi f}{K}\right) Z_1 Z_2}, \tag{3.49}$$

where K represents the stiffness of the intermediate thin layer.

Reflection coefficient for a thin oil layer from the spring model is a complex quantity which consists of both amplitude and phase information, rather than a real quantity as described before for the case of a perfect bonded interface between two media. It is determined by the acoustic impedance of surrounding media, the thin layer stiffness and the wave frequency.

By multiplying the complex conjugate of the denominator, the equation (3.49) can be rewritten as

$$R = \frac{\left(Z_1^2 - Z_2^2\right) + \left(\dfrac{2\pi f}{K} Z_1 Z_2\right)^2}{\left(Z_1 + Z_2\right)^2 + \left(\dfrac{2\pi f}{K} Z_1 Z_2\right)^2} + i \frac{\dfrac{4\pi f}{K} Z_1 Z_2^2}{\left(Z_1 + Z_2\right)^2 + \left(\dfrac{2\pi f}{K} Z_1 Z_2\right)^2}. \tag{3.50}$$

Then the amplitude of reflection coefficient can be expressed as

$$|R| = \frac{\sqrt{(Z_1 - Z_2)^2 + \left(\frac{2\pi f}{K} Z_1 Z_2\right)^2}}{\sqrt{(Z_1 + Z_2)^2 + \left(\frac{2\pi f}{K} Z_1 Z_2\right)^2}}, \tag{3.51}$$

and the phase of the reflection coefficient can be expressed as

$$\phi_R = \arctan \frac{\frac{4\pi f}{K} Z_1 Z_2^2}{\left(Z_1^2 - Z_2^2\right) + \left(\frac{2\pi f}{K} Z_1 Z_2\right)^2}. \tag{3.52}$$

From equation (3.51) and (3.52), it can be observed that not only the amplitude of reflection coefficient but also the phase of reflection coefficient varies with the mediate layer stiffness.

3.3.3.1 Stiffness of a liquid layer

The stiffness of an intermediate layer describes the resistance to the compression of the layer with applying pressure. For a liquid layer, it represents the nominal pressure required to cause a unit compression in the layer (Dwyer-Joyce (2005)), expressed as :

$$K = -\frac{dp}{du}, \tag{3.53}$$

where u is the compression displacement, and the stiffness of a liquid is related to the bulk modulus and the layer thickness, which is expressed as :

$$K = \frac{B}{h}, \tag{3.54}$$

where h is the thickness and B is the bulk modulus.

The speed of the ultrasonic wave in a liquid is determined by the modulus and the density of the liquid, described by Brekhovskikh (1960) :

$$c = \sqrt{\frac{B}{\rho}}, \tag{3.55}$$

Substituting equation (3.55) into equation (3.54), then the stiffness of the liquid layer can

be written as

$$K = \frac{\rho c^2}{h}. \tag{3.56}$$

The stiffness tends to zero for the extreme case when the liquid thickness is sufficiently thick, whilst it increases remarkably for the case that the liquid film becomes vanishing thin.

3.3.3.2 Amplitude of reflection coefficient

The amplitude of reflection coefficient is related to the stiffness of the thin oil layer by equation (3.51) while its stiffness depends on the density, the speed of sound and the thickness of the oil. By substituting equation (3.56) into equation (3.51), the amplitude can be given in terms of thickness as :

$$|R| = \sqrt{\frac{(Z_1 - Z_2)^2 B^2 + (2\pi f h Z_1 Z_2)^2}{(Z_1 + Z_2)^2 B^2 + (2\pi f h Z_1 Z_2)^2}}. \tag{3.57}$$

Clearly, the amplitude of reflection coefficient depends on the acoustic impedance of the surrounding media 1 and 2, the wave frequency, the bulk modulus and the thickness of the thin liquid layer. The responses of ultrasonic waves with different frequency from a thin layer oil (having a thickness of 0.5 microns), separating steel components, steel and aluminum components, steel and perspex components, are given in figure 3.8, respectively. The acoustic properties of the media is given in the table 3.1. The magnitude of reflection coefficient is 0.04 at a frequency of 1 MHz and rises to 0.89 at a frequency of 50 MHz for a steel-oil-steel system; it increases from 0.46 to 0.80 as the frequency rises from 1 to 50 MHz for a steel-oil-aluminum system. The amplitude of reflection coefficient increases significantly with the wave frequency for these two cases, but, it rarely varies with frequency for the steel-oil-perspex case, remaining stable at 0.87 for the frequency from 1 to 50 MHz.

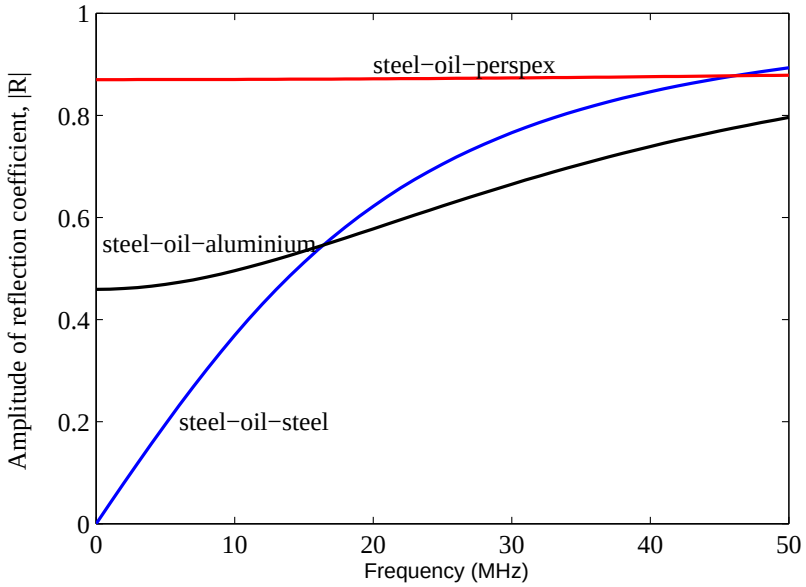


Figure 3.7. *The reflection coefficient amplitude spectra for a mineral oil layer separating steel-steel, steel-aluminium and steel-perspex elements. The thickness of oil (Shell T 68) is 0.5 microns.*

Rearranging the equation (3.57), the thickness of the thin liquid layer can be expressed in terms of amplitude of reflection coefficient that can be obtained from experiments.

$$h = \frac{B}{2\pi f Z_1 Z_2}\sqrt{\frac{(Z_1+Z_2)^2|R|^2-(Z_1-Z_2)^2}{1-|R|^2}}. \tag{3.58}$$

Once the reflection coefficient is measured, the thickness of the thin liquid layer can be achieved provided that the acoustic impedance of surrounding media and the bulk modulus of the liquid are given. If the separated media have the same acoustic impedance Z, the equation (3.58) then reduces to

$$h = \frac{B}{\pi f Z}\sqrt{\frac{|R|^2}{1-|R|^2}}. \tag{3.59}$$

Figure 3.8 shows the reflection coefficient spectra for a thin oil layer between two identical elements. For the three conditions, the amplitude of reflection coefficient shows largest dependency on the wave frequency in the steel-oil-steel case, followed by the aluminum-oil-aluminum case, and the least dependency for the perspex-oil-perspex case.

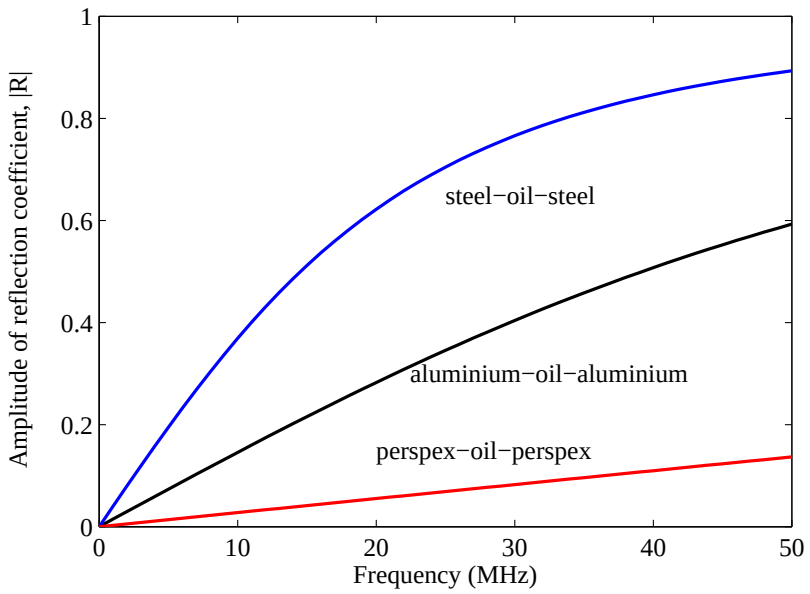


Figure 3.8. *The reflection coefficient spectra for a mineral oil layer between two identical elements. The thickness of oil (Shell T 68) is 0.5 microns.*

As steel is the most common material for bearings, the response of ultrasonic wave from steel-oil-steel is the major interest in this study. Usually, the thickness of oil layer in bearings depends on the bearing geometry, the applied load and the operational speed. Figure 3.9 describes the reflection coefficient spectra for a oil layer between steel components with varying thickness, which can be functioned as a guideline for choosing transducer frequency for bearing film thickness measurements. As stated before, the spring model is more suitable for oil layer thinner than several microns, which is clearly depicted by the figure. For an intermediate oil layer having thickness between 1 and 5 microns, ultrasonic waves with low frequency, say less than 20 MHz, are required, while higher frequency is required for a layer thinner than 1 micron. Need to bear in mind that ultrasonic waves with high frequency are associated with high attenuation, limiting the distance that they can penetrate.

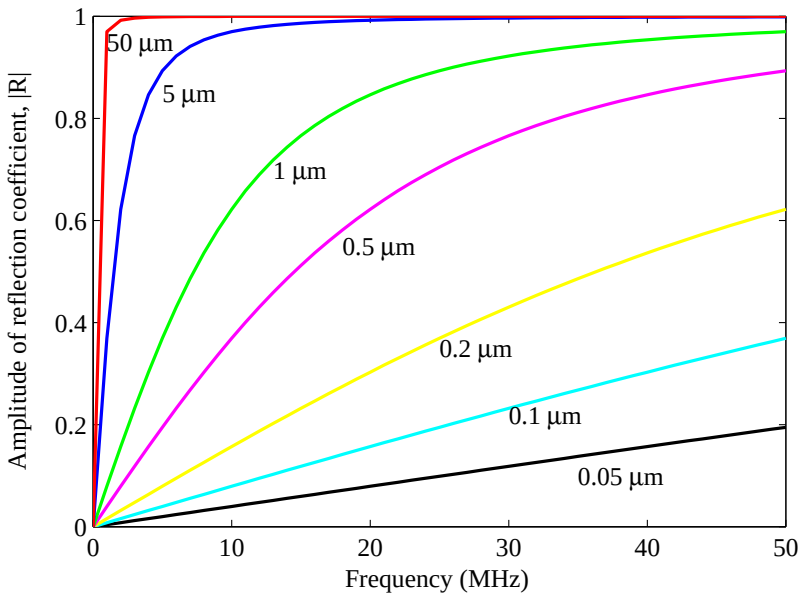


Figure 3.9. *The reflection coefficient spectra for a mineral oil layer (Shell T68) between steel components. The oil thickness varies from 0.05 to 50 microns.*

3.3.3.3 Phase change

The phase of the reflection coefficient is a measure of the phase difference between the reflected and the incident wave. It is related to the thin layer stiffness, given by equation (3.52) where the stiffness is related to the thickness by equation (3.54). As a result, the phase of the reflection coefficient can be written in terms of the thin layer thickness (Reddyhoff et al. (2005)):

$$\phi_R = \arctan \frac{4\pi f B h Z_1 Z_2^2}{\left(Z_1^2 - Z_2^2\right) B^2 + \left(2\pi f h Z_1 Z_2\right)^2}. \tag{3.60}$$

When the surrounding media have the same acoustic impedance Z, equation (3.60) is then reduced to :

$$\phi_R = \arctan \frac{B}{\pi f h Z}. \tag{3.61}$$

Figure 3.10 gives a reflection coefficient phase spectra for an oil layer between steel elements with varying thickness. It can be observed that the phase of the reflection coefficient for a thick oil layer is quite small and it increases as the oil thickness decreases. For a very thin layer, less than 0.05 microns, the stiffness approaches infinite and the phase shift between the reflected and the incident wave approaches $\frac{\pi}{2}$.

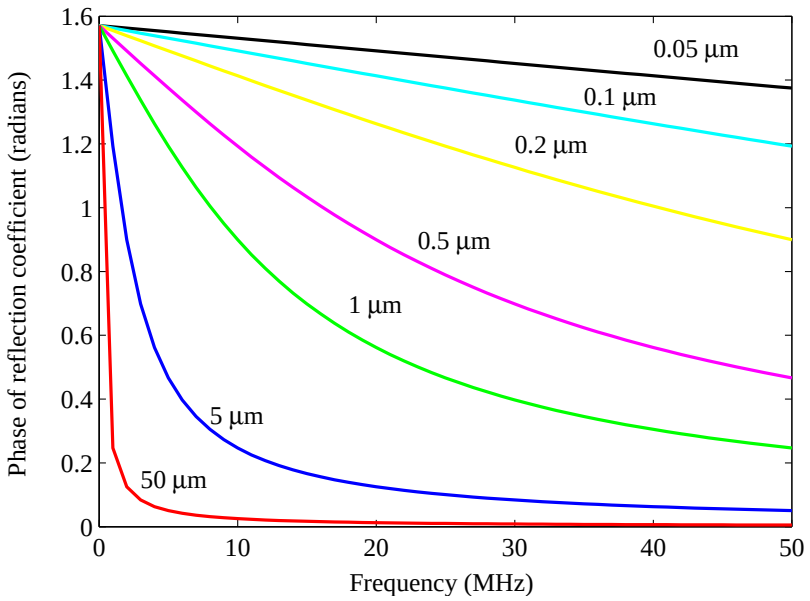


Figure 3.10. *The phase of reflection coefficient spectra for a mineral oil (Shell T68) layer between steels. The oil thickness varies from 0.05 to 50 microns.*

The film thickness can be achieved once the phase difference between reflected and incident signal is measured as :

$$h = B\frac{Z_1Z_2^2 \pm \sqrt{\left(Z_1Z_2^2\right)^2 - \left(Z_1Z_2\right)^2\left(Z_1^2 - Z_2^2\right)\tan^2\phi_R}}{2\pi f\left(Z_1Z_2\right)^2\tan\phi_R}. \tag{3.62}$$

Equation (3.58) and (3.62) show that both the amplitude and the phase of the reflection coefficient can be used to interpret the film thickness by using the spring model. Oil film thickness in both hydrodynamic lubrication (Reddyhoff et al. (2005)) and elastohydrodynamic lubrication (Gasni (2012)) has been measured, using both the amplitude and the phase of the reflection coefficient. For the simplest case, a thin oil layer separating two media having identical acoustic impedance, the amplitude of the reflection coefficient is a simple cosine of the phase.

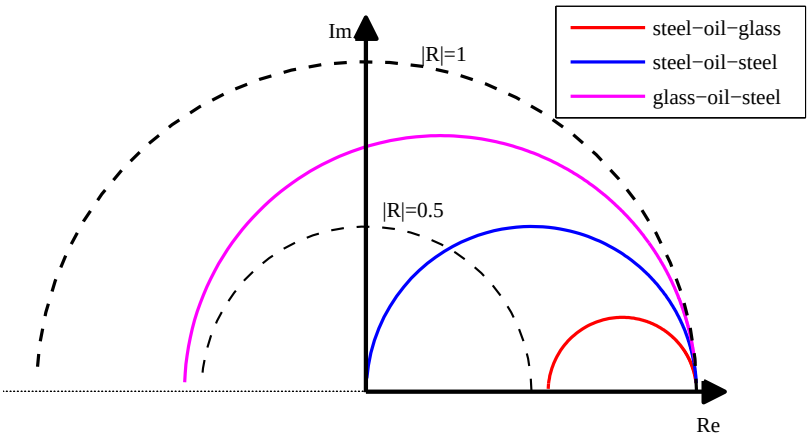
$$\mid R \mid = \cos\phi_R. \tag{3.63}$$

For a given ultrasonic wave, the upper and the lower limit of the phase change depends on both the dissimilarity of acoustic impedance and the stiffness. They are given for three typical conditions for surrounding materials in table 3.3.

Table 3.3: Limits of phase change for three acoustic impedance mismatch conditions (Reddyhoff (2007))

$\phi_R(radians)$	Case 1	Case 2	Case 3
	$Z_1 > Z_2$	$Z_1 = Z_2$	$Z_1 < Z_2$
Maximum	$< \frac{\pi}{2}$	$\frac{\pi}{2}$	π
Minimum	0	0	0

The Argand diagram of the reflection coefficient is given in figure 3.11, showing the effect of acoustic impedance mismatch between contact media and the effect of layer stiffness. The reflection coefficient for ultrasonic waves reflected from an oil layer between different materials are calculated from equation (3.49). The frequency of the wave is 10 MHz while the thickness of oil varies from 0.01 to 50 micrometers.



`Figure 3.11.` *Argand diagram for the reflection coefficient from an oil layer with varying thickness between different materials: a) steel-oil-glass; b) steel-oil-steel; c) glass-oil-steel (Reddyhoff (2007)).*

In the first case, an ultrasonic wave travels from a steel component to an oil layer and is reflected back with part of wave energy propagating into a glass element. For a thick oil layer, the stiffness of the layer tends to zero and almost total reflection is obtained. As oil thickness drops, the film stiffness increases and the amplitude of reflection coefficient reduces. The process is further demonstrated by figure 3.12.

- Initially the film is of 50 micrometers thick and the stiffness is negligible; and the reflection coefficient can be represented by a vector from the origin to the point ' A '

on the real axis. The amplitude of the reflection coefficient achieves the maximum of unity and the phase angle is zero.

- With the reduction in thickness, the layer stiffness increases and the reflection coefficient can be represented by a vector from the origin to the point ' B ' on the right semi-circle. The amplitude of reflection coefficient decreases while the phase angle increase slightly. For instance, the amplitude of reflection coefficient reduces to 0.73 while the phase achieves the maximum of 0.29 radians when the oil thickness reduces to 2 micrometers.

- Followed by further reduce in the film thickness, the point B moves to the left semi-circle in the figure. The amplitude of reflection coefficient decreases gradually while the phase angle falls quickly.

- When the oil layer becomes vanishingly thin, 10 nanometers, the film becomes quite stiff and the reflection coefficient can be represented by a vector from the origin to the point ' C ' on the real axis. The amplitude reduces to its minimum of 0.55 and the phase angle reduces to zero.

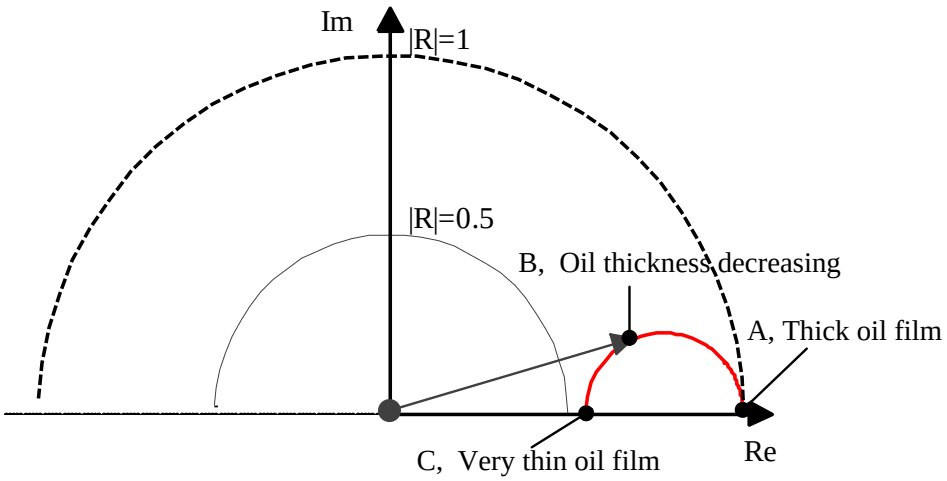


Figure 3.12. *Reflection coefficient from oil layer between steel and glass. Ultrasonic waves is normally incident from steel to the oil layer.*

In the second case, the oil layer is formed between steel components. The reflection coefficient goes through all the stages as it does in the first case. Initially, the amplitude of the reflection coefficient is unity and the phase angle is zero for a thick layer, 50 micrometers. Followed by gradual reduce in the amplitude and remarkable increase in the phase change.

Finally, the amplitude of reflection coefficient reduce to zero while the phase achieves the maximum of $\frac{\pi}{2}$ when the oil layer becomes vanishing thin, resulting in the point ' C ' approaches the origin.

In the third case, an ultrasonic wave travels from a glass component to an oil layer and is reflected back with part of wave energy propagating into a steel surface. For a thick oil layer, the stiffness of the layer tends to zero and almost total reflection is obtained. As the thickness of oil layer decreases, the stiffness becomes stiffer and stiffer. The amplitude of reflection coefficient remains large and the phase angle increase gradually. Finally the amplitude reduces gradually to 0.5 and the phase angle increase to π for a vanishing thin oil layer. In this case, as the acoustic impedance of glass is less than that of steel, the real part is negative, resulting in the point ' C ' lies on the real axis to the left of the origin.

3.4 Acoustoelastic effect

The speed of ultrasonic wave in a material depends on the state of the stress and this is known as the acoustoelastic effect (Egle and Bray (1976)). Stress measurements with ultrasonic techniques are based on the stress induced anisotropic behavior which is result from the non-linearity in the strain-displacement and constitutive relations.

Two types of bulk waves are most widely used: longitudinal waves and shear waves. A wave propagating in the direction of the particle motion is referred to as a longitudinal wave, while a wave propagating transverse to the direction of the particle motion is referred to as a shear wave. Figure 3.13 gives the velocity of ultrasonic waves and applied stress in orthogonal directions, where c_{xx}, c_{yy} and c_{zz} are the velocity of longitudinal waves and c_{xy}, c_{xz},c_{yx} c_{yz}, c_{zx},c_{zy} are the velocity of shear waves.

It is found that when particles of an ultrasonic wave vibrates in the direction of the stress field, it exhibits more sensitivity to stress; while the wave has less sensitivity to stress when the movement of particles is perpendicular to the stress (Bray and Tang (2001)). Thus, in figure 3.13 the velocity of ultrasonic waves c_{zz}, c_{yz} and c_{xz} are more sensitive to applied stress than c_{xx}, c_{xy} , c_{yx} , c_{yy}, c_{zx} ,c_{zy}.

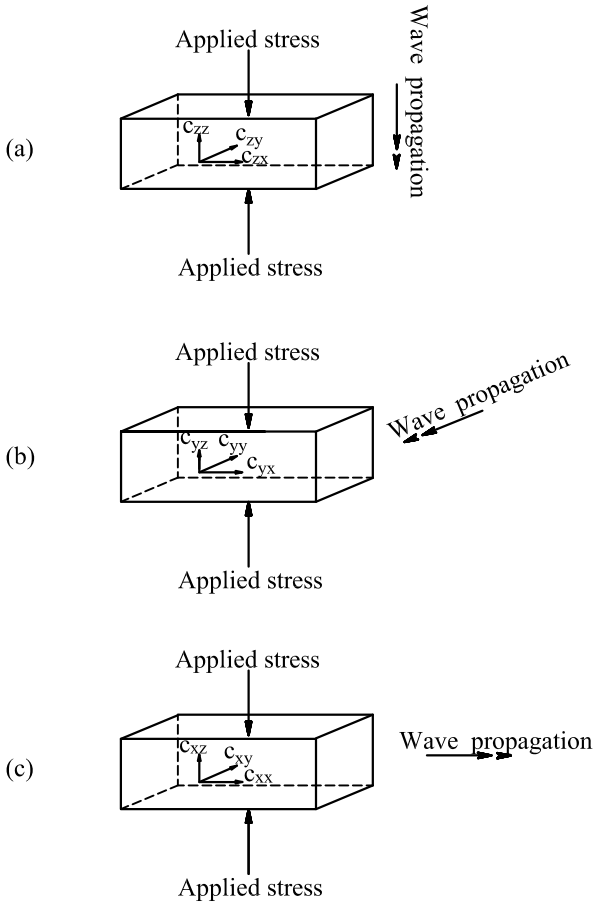


Figure 3.13. *Velocity of ultrasonic waves and stress field in orthogonal directions (Bray and Tang (2001)).*

The speed of the ultrasonic waves in a solid with applied stress depends on the second and the third order elastic constants, which has been described by Hughes and Kelly (1953), as :

$$\rho_0 (c_{zz})_P^2 = \lambda + 2\mu + (2l + \lambda)\,\theta + (4m + 4\lambda + 10\mu)\,(\varepsilon_z)_P\,, \tag{3.64}$$

$$\rho_0 (c_{zx})_P^2 = \mu + (\lambda + m)\,\theta + 4\mu\,(\varepsilon_z)_P + 2\mu\,(\varepsilon_x)_P - \frac{1}{2} n \left(\varepsilon_y\right)_P\,, \tag{3.65}$$

$$\rho_0 \left(c_{zy}\right)_P^2 = \mu + (\lambda + m)\,\theta + 4\mu\,(\varepsilon_z)_P + 2\mu\,\left(\varepsilon_y\right)_P - \frac{1}{2} n\,(\varepsilon_x)_P\,, \tag{3.66}$$

$$\theta = (\varepsilon_x)_P + \left(\varepsilon_y\right)_P + (\varepsilon_z)_P\,. \tag{3.67}$$

where ρ_0 is the initial density in unstressed state; c_{zz} is the speed of longitudinal waves

propagating along the z axis, and c_{zx} and c_{zy} are the speed of shear waves propagating along the z axis. λ and μ are the second order elastic constants while l , m and μ are the Murnghan's third elastic constants. ε_x, ε_y and ε_z are the strain components in x,y,z axes, respectively. The subscript 0 and P represent parameters in unloaded and loaded state, respectively. The second order elastic constant and Possion's ratio are given by :

$$\lambda = \frac{3B(3B-E)}{9B-E}, \tag{3.68}$$

$$\mu = \frac{3BE}{9B-E}, \tag{3.69}$$

$$\nu = \frac{\lambda}{2(\lambda+\mu)}, \tag{3.70}$$

where ν is the Poisson's ratio, B is the bulk modulus and E is the elasticity modulus.

For the simple state of uni-axial stress where the stress is only applied in the z direction as shown by figure 3.13a, the components of strain in the x, y axes are related to the strain in the z axis by $(\varepsilon_x)_P = \left(\varepsilon_y\right)_P = -\nu\,(\varepsilon_z)_P$. Equations (3.64), (3.65) and (3.66) become :

$$\rho_0\,(c_{zz})_P^2 = \lambda + 2\mu + \left[4(\lambda+2\mu) + 2(\mu+2m) + 2\nu\mu\left(1+2\frac{l}{\lambda}\right)\right](\varepsilon_z)_P, \tag{3.71}$$

$$\rho_0\,(c_{zx})_P^2 = \rho_0\,(c_{zx})_P^2 = \mu + \left[4\mu + \nu\frac{n}{2} + m(1-2\nu)\right](\varepsilon_z)_P. \tag{3.72}$$

Longitudinal waves generated by a piezoelectric sensor propagating parallel to the stress were used in this study as they exhibit most sensitivity to applied stress. Expressions of ultrasonic waves speed propagating perpendicular to the applied stress as shown in figure 3.13b and c Hughes and Kelly (1953), are given here for the reader's interest.

$$\rho_0\,(c_{xx})_P^2 = \lambda + 2\mu + \left[2l(1-2\nu) - 4\nu(m+\lambda+4\mu)\right](\varepsilon_z)_P, \tag{3.73}$$

$$\rho_0\,(c_{xz})_P^2 = \rho_0\left(c_{yz}\right)_P^2 = \mu + \left[(\lambda+m+2\mu)(1-2\nu) + \nu\frac{n}{2}\right](\varepsilon_z)_P, \tag{3.74}$$

$$\rho_0\left(c_{xy}\right)_P^2 = \rho_0\left(c_{yz}\right)_P^2 = \mu + \left[(\lambda+m)(1-2\nu) - 6\nu\mu - \frac{n}{2}\right](\varepsilon_z)_P. \tag{3.75}$$

From equation (3.64) the ultrasound speed without stress is only determined by the

second order elasticity of the material, as :

$$\rho_0 (c_{zz})_0^2 = \lambda + 2\mu, \tag{3.76}$$

where $(c_{zz})_0$ is the ultrasonic wave speed in unloaded state. By differentiating equation (3.71), the relationship between the wave speed and the strain can be written as :

$$\rho_0 (c_{zz})_P d(c_{zz})_P = \left[2(\lambda + 2\mu) + (\mu + 2m) + \nu\mu \left(1 + 2\frac{l}{\lambda} \right) \right] d(\varepsilon_z)_P . \tag{3.77}$$

Provided the speed change is small :

$$\rho_0 (c_{zz})_P (c_{zz})_0 \approx \rho_0 (c_{zz})_0^2 = \lambda + 2\mu, \tag{3.78}$$

Then the relative change in the speed of sound is related to the strain by dividing equations (3.77) and (3.78) as :

$$\frac{d(c_{zz})_P}{(c_{zz})_0} = \left[2 + \frac{(\mu + 2m) + \nu\mu \left(1 + 2\frac{l}{\lambda} \right)}{\lambda + 2\mu} \right] d(\varepsilon_z)_P . \tag{3.79}$$

If the acoustoelastic constant is defined as the relative change in the ultrasonic speed with corresponding strain, designated by L_{zz} in the $z-$direction, the effect of strain that introduced by applied load on the change of ultrasound speed can be expressed as (Egle and Bray (1976)) :

$$\frac{d(c_{zz})_P}{(c_{zz})_0} = L_{zz} d(\varepsilon_z)_P , \tag{3.80}$$

where

$$L_{zz} = 2 + \frac{(\mu + 2m) + \nu\mu \left(1 + 2\frac{l}{\lambda} \right)}{\lambda + 2\mu}. \tag{3.81}$$

The acoustoelastic constant illustrates the ultrasound speed variability in the material under different load states, which is determined by the material second and third order elastic constants.

3.5 Summary

In this chapter, the response of ultrasonic waves at different boundaries are investigated. Reflection and transmission of ultrasonic waves have been quantified by the reflection coefficient and the transmission coefficient, respectively. For a perfect bonded interface, reflection coefficient is a real number and only determined by the acoustic properties of the media in contact. This is the case for an oil layer formed between two elastic bodies, the thickness of which can be determined by using either the ToF method or the resonance method. However, for a very thin layer separating two media where spring model is utilized, the reflection coefficient is a complex quantity. In this case, the reflection coefficient depends on acoustic properties of all involved media, the wave frequency and the thin layer stiffness. The phase change between the incident and the reflected wave, has been studied. It can be concluded that the magnitude of the phase change is determined by the acoustic impedance mismatch between contact media and the intermediate layer stiffness. The ultrasound speed in a loaded material varies linearly with load and this behavior can be illustrated by the material acoustoelastic constant.

Chapter 4

Modeling ToF Change with Contact Load

After introducing the interaction of ultrasonic waves with boundaries, this chapter develops an analytical model for the variation of ultrasonic ToF in one of the contact bodies with load via the surface deflection and change in speed of sound. However, in practical ultrasonic measurements, the phase of the reflected wave also changes with contact condition and this phase change appears as an apparent time shift in simple peak to peak time determination. The contributions of the surface deflection, the change in speed of sound and the phase change effect on the ToF change are investigated. The relationship between the ToF change and contact load developed in this chapter is used to interpret ultrasonic measurement data as contact load in Chapter 6, 7 and 8.

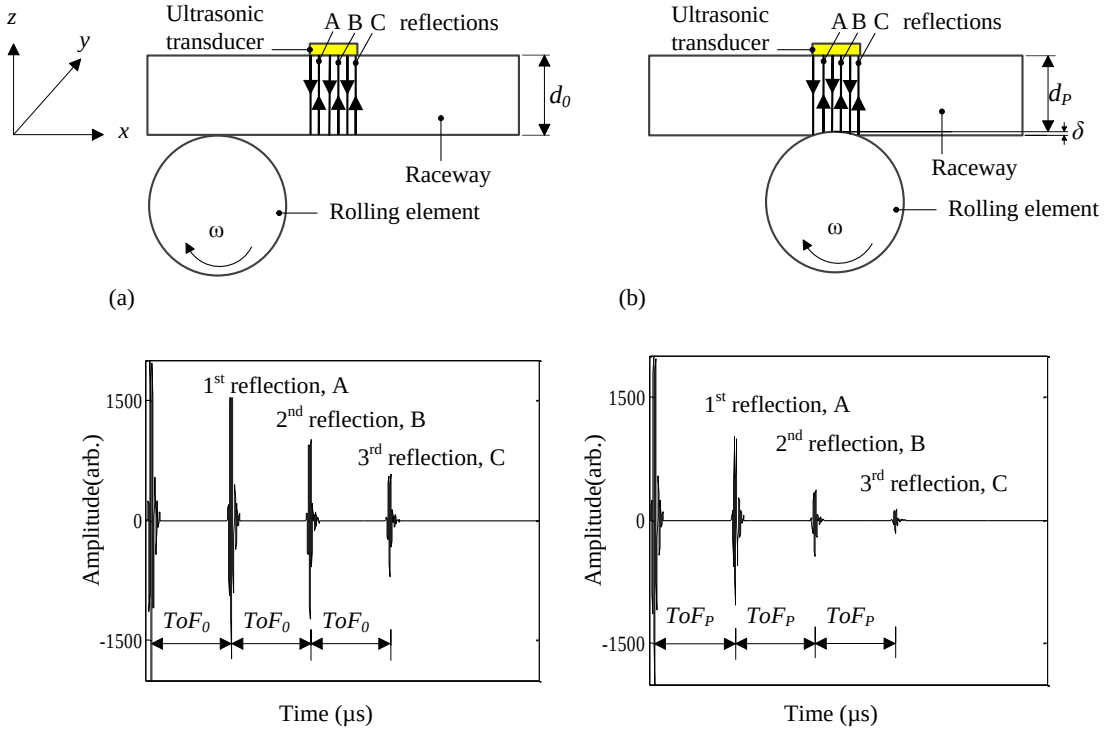
4.1 Introduction

The ultrasonic ToF in a component is determined by the distance that ultrasound travels through and the wave speed, both of which depend on the load state. This chapter aims to model the relationship between the ultrasonic ToF and the contact load.

When an ultrasonic transducer is coupled to a component, the ToF of ultrasound in the element depends on its thickness and the wave speed, as shown by equation (3.20) (here h represents the component thickness). By measuring the ultrasonic ToF, the thickness can be obtained with given ultrasound speed, which is the basics of ultrasonic distance measurements. Using this method, changes in the dimensions of a component can be measured. For most thickness measurements, the speed of ultrasound in a given material is treated as a constant. This is the case for a free surface, but not the case for a surface contact with

another under high load.

For contact surfaces under high load, this approach may leads to inaccurate results since the ultrasonic wave speed in a material is not a constant but varies with the load/stress state, which is known as acoustoelastic effect. In tribological contacts, such as a rolling bearing contact, gear contact, cam/tappet contact, piston ring/liner contact and wheel/ rail contact, the elastic deformation is in the order of micrometers and the corresponding ToF change is in the nanosecond range. Meanwhile, the ultrasound speed changes due to the high contact stress (in the GPa range), gives the ToF change of the same order. Therefore, effects of both the surface deflection and the ultrasound speed change on the ToF need to be considered.



`Figure 4.1.` *Typical ultrasonic waveform from (a) an unloaded raceway and (b) a loaded raceway-roller contact. The waveform consists of the initial pulse and three consecutive reflections from the contact surface of the raceway (A, B, and C).*

Figure 4.1 shows typical longitudinal ultrasonic waveforms reflected from the back surface of a raceway when (a) it is free and unloaded, and (b) it is loaded by a rolling element (ball or roller). The ultrasonic sensor is directly coupled to the raceway, worked in the pulse-echo mode where ultrasonic waves traveling from the sensor to the contact face, being reflected back and received by the same transducer. The first pulse picked up by the sensor is the reflection from the front face of the raceway (transducer - raceway interface),

followed by the first reflected pulse (peak A) from the back face of the raceway (i.e. contact interface in unloaded and loaded state). As the wave reflects back and forth in the raceway, the sensor picks up these subsequent pulses, which are observed in the time domain trace. This is illustrated by peak B, the 2nd reflection, and peak C, the 3rd reflection from the back face of the raceway in figure 4.1. The time interval between adjacent peaks (ToF_0 or ToF_P) is governed by the wave path ($2d_0$ or $2d_P$) and the wave speed of in the material (c_0 or c_P). As the raceway is loaded by the passage of the rolling element, the wave path length reduces ($2d_P < 2d_0$). Also, the speed of the ultrasound in the stressed material increases ($c_P > c_0$) and therefore the ToF reduces ($ToF_P < ToF_0$). With the wave propagation, energy will be lost due to transmission at the interface (which is much greater in the loaded case), wave dispersion and attenuation in the raceway. Therefore, the amplitude of the signal reduces with the wave path.

As described in Chapter 3, there is a phase change in the reflected pulses from the contact interface, which depends on the local contact stiffness. In rolling element contacts, the contact stiffness can be considerably high and results in a large phase change. This phase change alters the arrival time of the reflected pulses and it appears as an apparent time shift in simple peak to peak time determination. The magnitude of the ToF caused by the phase change effect is also in the order of nanoseconds, which is significant compared to the ToF caused by the deflection and the wave speed and cannot be neglected.

Therefore, the ultrasonic ToF change caused by the contact load via the surface deflection and the acoustoelastic effect, and that caused by the phase change effect are considered in the following sections.

4.2 ToF change caused by the surface deflection

4.2.1 Surface deflection

Load on a bearing is shared by a number of contacts between races and rolling elements, see figure 4.2. For each contact, the elastic deformation occurs and the magnitude is determined by the material property, the bearing geometry and load upon the contact. As described in Chapter 2, load upon each contact can be calculated. However, the principal interest of this study is given to the most loaded contact in a ball or a roller bearing without loss of generality. The load and the deflection equations are summarized and represented here.

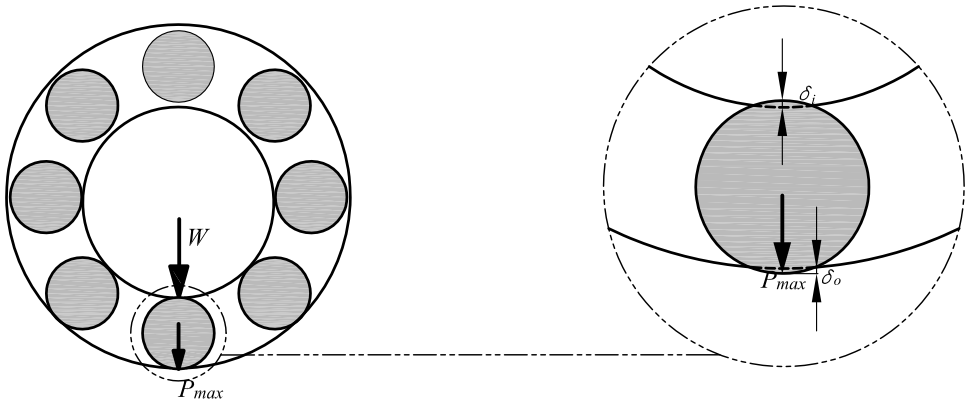


Figure 4.2. *The contact deformation in the most loaded rolling element and the bearing races.*

For the most loaded ball-raceway contact, the load in contact can be expressed as:

$$P_{max} = \frac{WZ_w}{N}, \tag{4.1}$$

where Z_w is a load parameter which has been given by equation (2.12), N is the number of rolling elements.

The maximum deformation in the most heavily loaded raceway-ball contact, taking place at the elliptical contact centre, which can be obtained using equation (2.46) which is presented here again for convenience.

$$\delta_{max} = F\left[\frac{9}{2\varepsilon R'}\left(\frac{P_{max}}{\pi k E'}\right)^2\right]^{1/3}. \tag{4.2}$$

For the most loaded roller-race contact, the load in contact can be calculated from:

$$P_{max} = \frac{W\left(1-\dfrac{c_d}{2\delta_r}\right)}{\dfrac{N}{2\pi}\left(\psi_l - \dfrac{c_d}{2\delta_r}\sin\psi_l\right)} \tag{4.3}$$

The maximum deflection in the line contact between the raceway and the roller (on the load line) , can be obtained by using equation (2.64) which is represented here.

$$\delta_{max} = 3.84\times10^{-8}\frac{P_{max}^{0.9}}{(l_0\times1000)^{0.8}}. \tag{4.4}$$

4.2.2 ToF change caused by the surface deflection

The propagation of an ultrasonic wave through a component when it is in unloaded and loaded states is investigated. The ToF in the unloaded component, is determined by the thickness of the race and the ultrasound speed which are denoted by d_0 and c_0 in the unloaded race, respectively; the relationship can be expressed as:

$$ToF_0 = \frac{2d_0}{c_0}. \tag{4.5}$$

As the deformation occurs in the race when it is subjected to load, the thickness changes from d_0 to d_P and the wave path changes from $2d_0$ to $2d_P$. The ultrasonic ToF in the loaded component which is designated by ToF_P, can be obtained by

$$ToF_P = \frac{2d_P}{c_P}, \tag{4.6}$$

where c_P is ultrasonic wave speed in the race in loaded state.

The ToF change caused by the surface deflection alone, denoted by Δt_δ, can be calculated by subtracting the ToF under load from that in unloaded state, expressed as:

$$\Delta t_\delta = ToF_0 - ToF_P. \tag{4.7}$$

If the deformation in the raceway is denoted by δ, then the ultrasonic wave path change is $2\delta = 2d_0 - 2d_P$. Here the principal interest is the effect of the surface deflection on the ToF change alone, the ultrasound speed in the race in loaded state is assumed to be as the same as that in unloaded state, i.e. $c_P = c_0$. Therefore, the ToF change caused by the surface deflection can be directly related to the race deformation by the following equation:

$$\Delta t_\delta = \frac{2d_0}{c_0} - \frac{2d_P}{c_P} = \frac{2\delta}{c_P}. \tag{4.8}$$

The deformation of mating surfaces can be obtained from finite element analysis, Hertzian contact analysis, elastic theory, numerical method, or empirical load/deflection relations.

4.3 ToF change caused by the acoustoelastic effect

As described in Section 3.4, the ultrasound speed in a material changes with the load/stress state, which results in a change in the ToF. The magnitude of speed change can be predicted from the strain with given acoustoelastic constant, as shown by equation (3.80). The material acoustoelastic constant depends on its second and third order elastic constants. The

material second order elastic constants can be obtained from its bulk modulus and Young's modulus which are commonly available. However, the third order elastic constants are usually obtained from experiments (Egle and Bray (1976); Smith (1963); Tanala et al. (1995)). In this work the acoustoelastic constant for bearing steel was measured from the ultrasonic ToF.

4.3.1 Acoustoelastic constant measurement

Experiments have been carried out to measure the acoustoelastic constant for a typical bearing steel (100Cr6). A photograph of the test apparatus for a cylindrical roller is given in figure 4.3. The roller has a diameter of 30 mm and a length of 30 mm.

`Figure 4.3.` *Photography of the acoustoelastic constant test apparatus for a cylindrical roller.*

An ultrasonic sensor having the centre frequency of 10 MHz was coupled directly to a bearing steel roller and longitudinal waves were passed through the roller from end to end. The roller was compressed in the axial direction in a standard hydraulic loading machine with a capacity of 250 kN. A schematic diagram of the mechanical apparatus is given in figure 4.4. Strain gauges upon the roller were used to measure the strain while a thermocouple

was used to monitor the temperature during the test. Therefore, the roller length was measured for each load case. Ultrasonic reflections from the roller end were recorded and the ToF in the roller was measured for each load case; the method used to measure the ToF will be detailed in Section 5.3. Once the roller length and the ToF are achieved, the ultrasound speed in the roller under different load can be determined from equation (4.6). The change in wave speed can be easily obtained by subtracting the wave speed in the loaded roller from the wave speed in the unloaded roller.

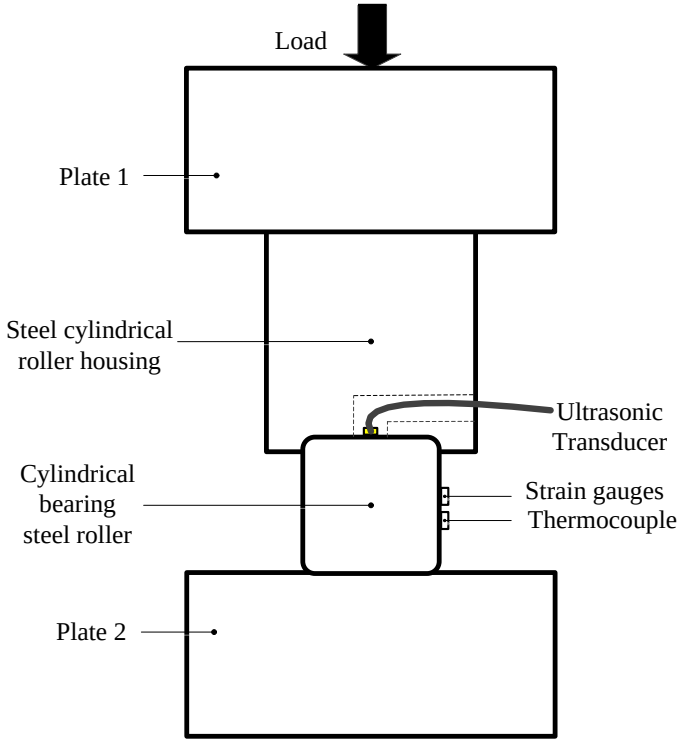


Figure 4.4. *Schematic of the acoustoelastic constant measurement apparatus. A 10 MHz ultrasonic transducer was used to generate longitudinal wave which propagated through the roller.*

Figure 4.5 gives a plot of measured relative change in the speed of sound (the ratio of the load induced speed change to the initial speed) against the measured strain for three repeated loading cycles. In the figure, several features are observed. Firstly, the relative speed change linearly varies with the roller strain. In the experiment, the roller was subjected to compression and thus the strain gets a sign of minus. Secondly, little hysteresis and some plastic deformation are observed in the first loading and unloading cycle. The measured relative speed changes in the first loading are slightly higher than those obtained in the first

unloading test, especially at the low load/strain end. In addition, the relative speed change is small. For instance, the maximum ratio is 0.0018 when the roller is under the maximum compression strain of 740 $\mu\mathrm{m}/\mathrm{m}$.

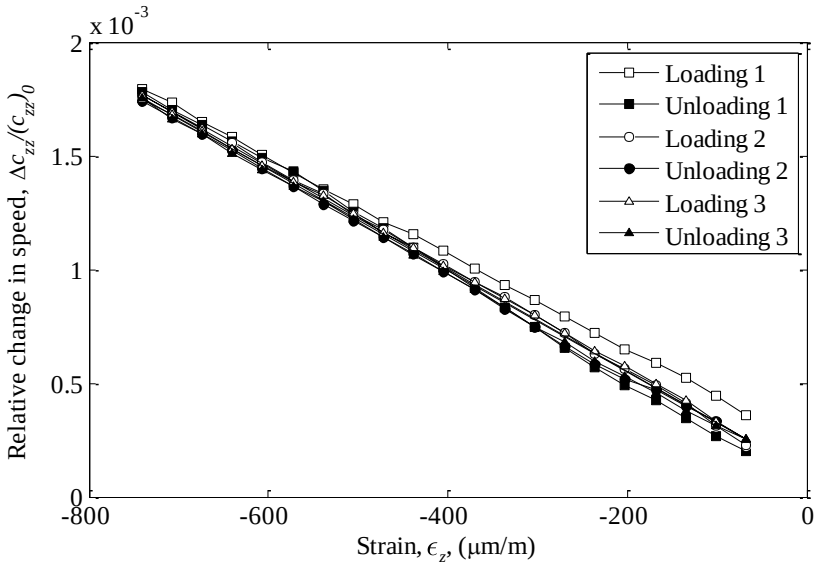


Figure 4.5. *Measured relative speed change of ultrasound in a bearing steel with strain*

According to the definition (see equation (3.80)), the acoustoelastic constant is equal to the slope of lines in the graph. Slight variations in the measured acoustoelastic constant can be observed in the figure. To be specific, in the first loading and unloading tests, the acoustoelastic constant was measured at -2.12 and -2.41, respectively. However, the acoustoelastic constant remains at -2.24 in the subsequent load cycles. A best fit was applied to all test data in figure 4.5, giving the acoustoelastic constant for this bearing steel of -2.24. The measured acoustoelastic constant for longitudinal waves in this bearing steel was compared with that for longitudinal waves in engineering materials measured by other researchers in table 4.1. It can be observed that the acoustoelastic constant measured from the novel ultrasonic ToF technique in this experiment agrees well with those reported in the literature.

Table 4.1: Acoustoelastic constant for longitudinal waves in engineering materials

Material	Load	L_{zz}
Bearing steel	Compression	-2.24
Aluminium (Tanala et al. (1995))	Tension(rolling direction)	-2.7
	Tension(transverse direction)	-3.1
316L Stainless steel (Tanala et al. (1995))	Tension	-2.1
4140 Steel (Tang and Bray (Montreal, PQ, 1996))	Tension(2.25 MHz)	-2.2
	Tension(5 MHz)	-2.36
Rail steel (1080) (Egle and Bray (1976))	Tension	-2.38
	Compression	-2.45

4.3.2 ToF change caused by the acoustoelastic effect

The strain in the normal direction is not constant but varies from a maximum at the contact surface reducing with depth through the component. The ultrasonic speed in the normal direction will therefore also vary through the component depth. The cumulative effect of the varying strain field must be used to determine the change in the ultrasonic wave speed and hence the ToF change. As such a case, the average strain can be expressed as:

$$(\bar{\varepsilon_z})_P = \frac{\int (\varepsilon_z)_P dz}{d_P}. \tag{4.9}$$

Note that the sign of $(\varepsilon_z)_P$ is negative because the bearing race is in compression under load. In the unstressed state, if the initial strain is zero, i.e. $(\bar{\varepsilon_z})_0 = 0$, then the strain change caused by load can be expressed as:

$$\Delta\varepsilon_z = (\bar{\varepsilon_z})_0 - (\bar{\varepsilon_z})_P = -\frac{\int (\varepsilon_z)_P dz}{d_P}. \tag{4.10}$$

The ultrasonic speed change caused by load can be obtained from the strain change in the material once the acoustoelastic constant for the material is known, and equation (3.80)

becomes:

$$\Delta c_{zz} = L_{zz}\,(c_{zz})_0\,\Delta\varepsilon_z, \tag{4.11}$$

where $\Delta c_{zz} = (c_{zz})_0 - (c_{zz})_P$. Then the speed of ultrasonic waves propagating in the compressed element can be obtained by

$$(c_{zz})_P = (c_{zz})_0\,(1 - L_{zz}\Delta\varepsilon_z)\,. \tag{4.12}$$

It can be observed that the ultrasound speed in a component increases with stress in the compression case while it decreases with the stress in the tension case. The speed change is governed by the strain (or stress), the speed in the unloaded state and the acoustoelastic constant of the material.

If the principal attention is given to the effect of the speed change alone, i.e, the surface deflection is not considered at this stage, the ToF that the ultrasonic wave taken in the loaded medium ToF_P becomes

$$ToF_P = \frac{2d_0}{(c_{zz})_P}, \tag{4.13}$$

and the difference between the ToF in the unloaded and loaded state can be obtained as:

$$\Delta t_c = \frac{2d_0}{(c_{zz})_0} - \frac{2d_0}{(c_{zz})_P}. \tag{4.14}$$

The deformation in the contacting media is defined by the strain as:

$$\delta = d_0\Delta\varepsilon_z, \tag{4.15}$$

Substituting equations (4.15) and (4.12) into equation (4.14), then the change in the ultrasonic ToF caused by the acoustoelastic effect can be expressed in terms of the deflection as:

$$\Delta t_c = \frac{-2d_0L_{zz}\delta}{(c_{zz})_0\,(d_0 - L_{zz}\delta)}, \tag{4.16}$$

Since the elastic deformation is relative small with respect to the component thickness, $L_{zz} = -2.24$ and $d_0 - L_{zz}\delta \approx d_0$, equation (4.16) simplifies to:

$$\Delta t_c = \frac{-2L_{zz}\delta}{(c_{zz})_0}. \tag{4.17}$$

From equation (4.17), it is clear that the ultrasonic ToF change is determined by the

initial speed in the medium in the unloaded state, the acoustoelastic constant and the deformation of the medium. Since the deformation of the medium is load dependent, the ultrasonic ToF change resulted from the acoustoelastic effect is directly related to load.

4.4 Apparent ToF change caused by the phase change effect

4.4.1 Phase change between the incident and the reflected wave

As stated in Section 3.3.3, the response of ultrasonic waves can be described by the spring model when using ultrasonic waves to investigate a thin liquid layer in a three layered system. The reflection coefficient for the system is governed by the acoustic properties of contact media, the oil layer stiffness and the wave frequency. The phase of the reflection coefficient is a measure of phase change between the incident and the reflected wave, which presents as an apparent time shift as shown in figure 4.6.

Figure 4.7 gives a typical reflected waveform from a steel-air interface, where the first peak is the initial pulse from the sensor and the second is the first reflection from the back face of the steel (steel-air interface). In this case, the reflection coefficient is a real quantity as shown by equation (3.18). The amplitude of the reflection coefficient approaches to unity because of the large difference between steel and air acoustic impedance, see table 3.1, which means nearly all of ultrasonic waves is reflected back at the steel-air interface. Therefore, the reflected waved from a steel-air interface can be thought as an accurate approximation of the incident wave at the boundary.

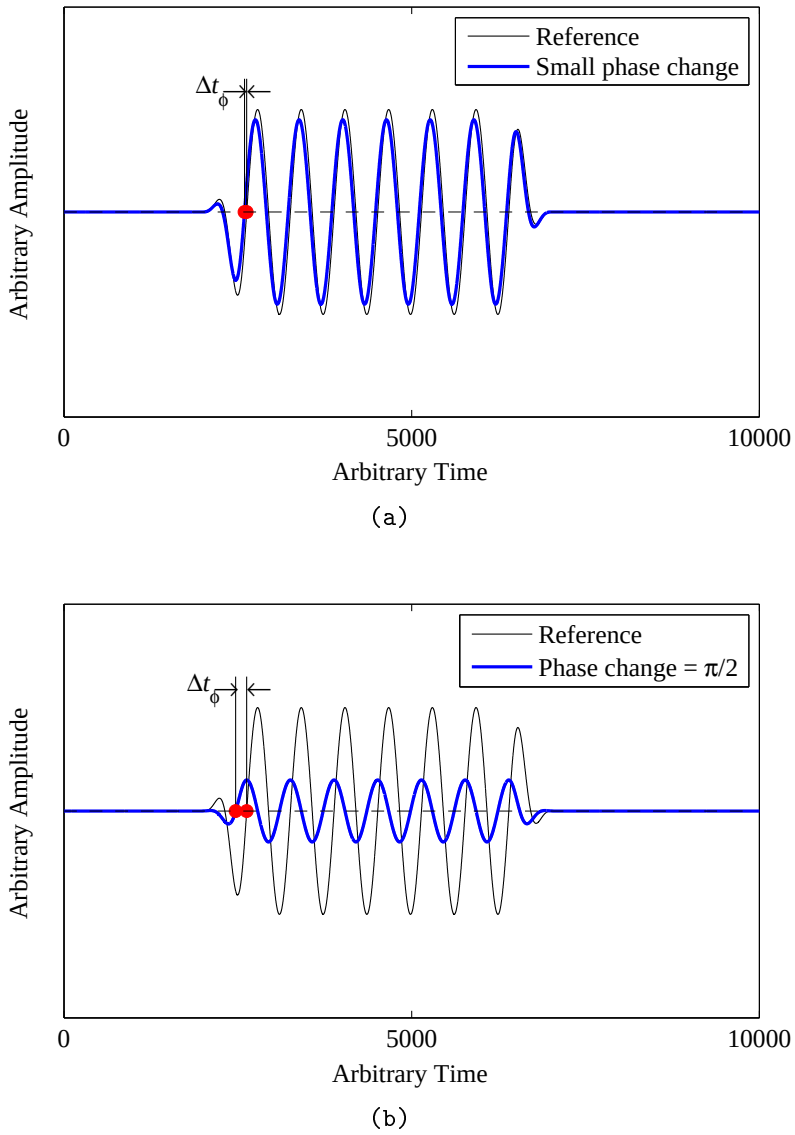


Figure 4.6. *Schematic diagram showing the ToF change between two ultrasonic signals determined by a zero-crossing approach. The ToF change caused by: a) a small phase change and b) a phase change of* $\frac{\pi}{2}$.

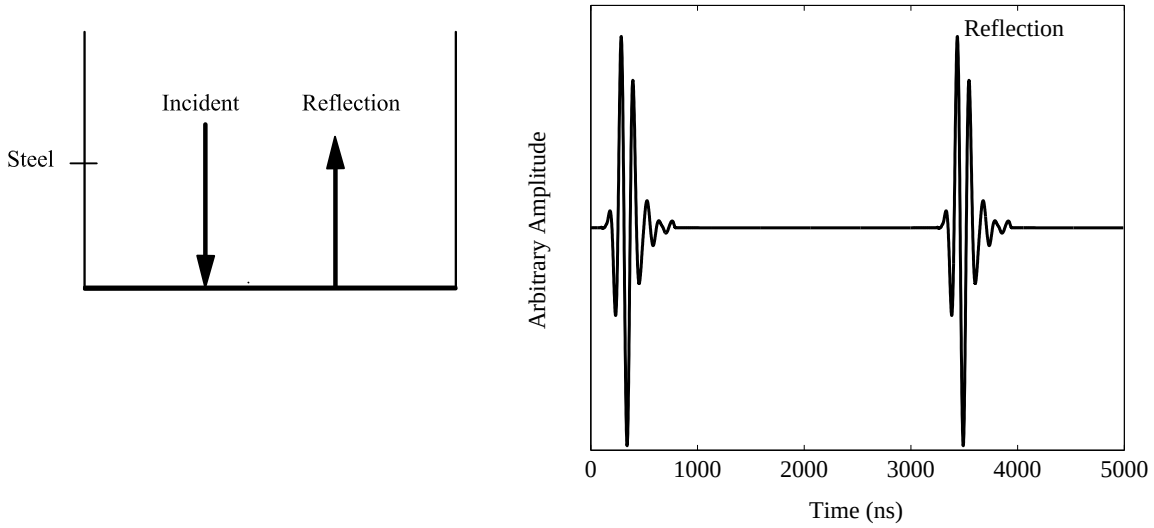


Figure 4.7. *A reflected waveform from a perfectly bonded boundary - steel-air interface.*

By contrast, figure 4.8 gives a typical reflected waveform from the steel-oil-steel case, where the second echo is the first reflection from the back face of steel (the steel-oil interface).

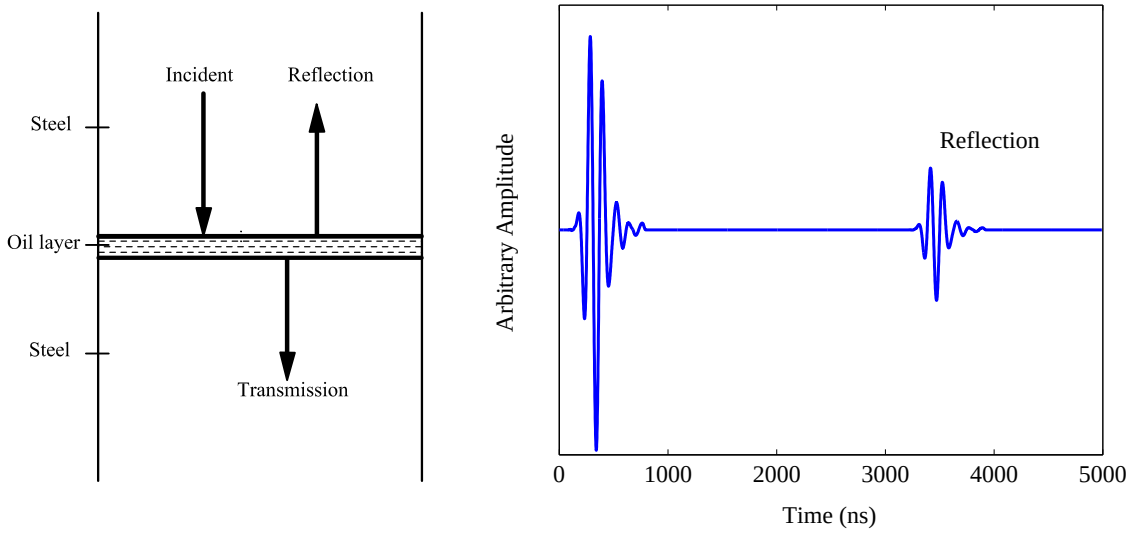


Figure 4.8. *A reflected waveform from an imperfectly bonded boundary - solid-oil-solid interface.*

To compare the first reflection from the back face of steel, located from 3000 to 4000 nanoseconds (ns) in the time history, they are extracted and represented in figure 4.9. Clearly,

the steel-oil interface has much smaller amplitude of the reflected pulse than the steel-air interface. Moreover, the reflected pulse from the steel-oil interface shifts slightly to the left of the reflected pulse from the steel-air interface.

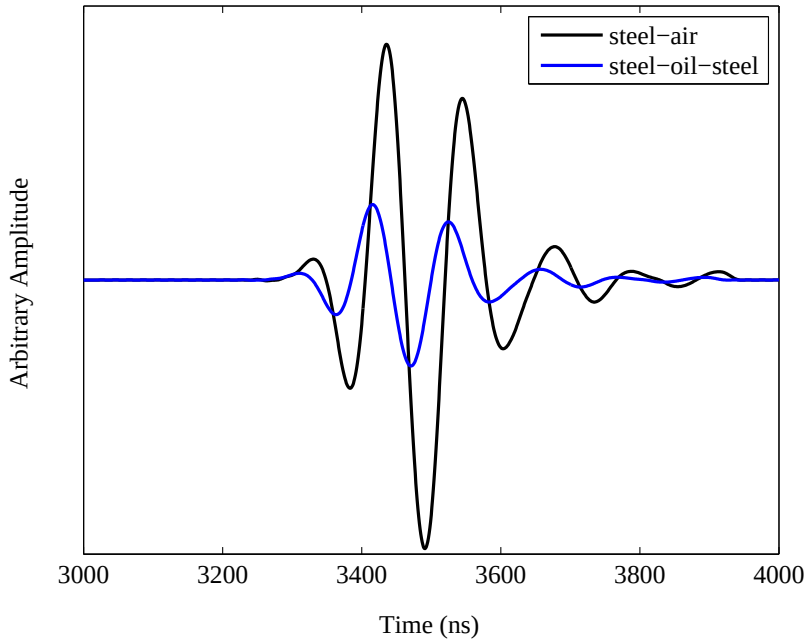


Figure 4.9. *The first reflected pulse from a steel-oil-steel interface compared with that from a steel-air interface.*

According to the definition in Chapter 3, the incident pulse is required to obtain the reflection coefficient. However, with the absence of direct record of the incident signal, a reference signal from a known interface is used instead in practical measurements. Then, the amplitude of the reflection coefficient is described by equation (4.18), and the phase of the reflection coefficient is illustrated by equation (4.19).

$$\mid R \mid = \frac{A(f)}{A_r(f)} \mid R_r \mid, \tag{4.18}$$

$$\phi_R = \phi - \phi_r. \tag{4.19}$$

where $\mid R \mid$ is the amplitude of the reflection coefficient, $A(f)$ and ϕ are the amplitude and the phase of the reflected wave from the given interface, respectively. The subscript r represents for the reference. In ultrasonic tests, the reflected signal from a specimen-air interface is always taken as the reference signal, since the reflected pulse is in-phase with the incident pulse and they have same amplitude.

4.4.2 ToF change caused by the phase change effect

Ultrasonic waveform is commonly given in the time domain trace. The easiest practical way to determine the ultrasonic ToF is to measure the time between the relevant peaks in the waveform using a zero-crossing method. Figure 4.10 illustrates a simple zero-crossing approach to determine the time shift between a reference signal and a intended signal. For the case of a reflection from a free surface, the reflected wave is in-phase with the incident wave and a simple zero-crossing approach can be used to obtain the ToF. This method is commonly used, for example to determine the tension in a threaded fastener from the change in ToF under load. There is no phase change between the unloaded reference and the loaded signal in figure 4.10. The time difference, Δt, between the zero point of the reference signal and the corresponding zero point of the measured signal, represents the change in ToF caused by the component length change and the ultrasound speed change in the stressed fastener.

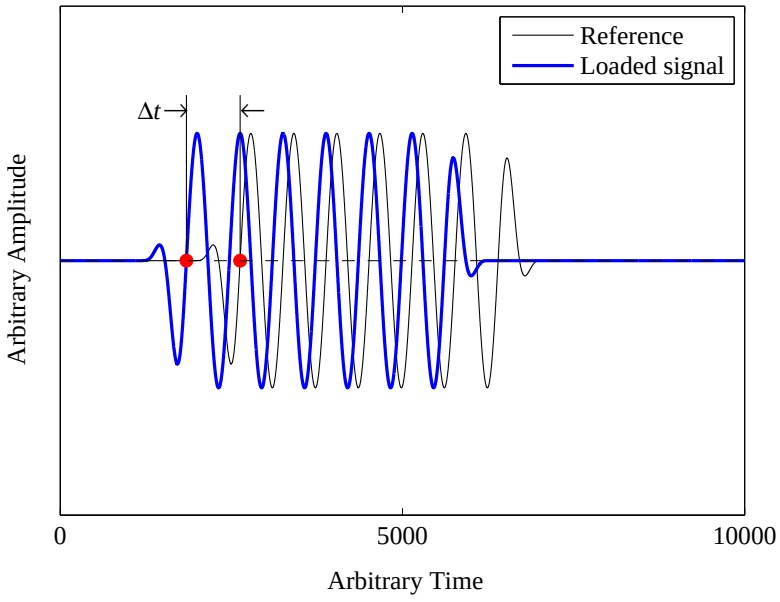


Figure 4.10. *Schematic diagram showing the ToF change between two ultrasonic signals determined by a zero-crossing approach. The ToF change is caused by a change of wave path and a wave speed change.*

However, for the case of a reflection from an interface with a contact present, the phase can vary between 0 and π. The magnitude of the phase change (and indeed the amplitude) depends on the contact condition (Reddyhoff et al. (2005)). For a given ultrasonic wave to an oil layer between two identical media, the phase change varies from zero to $\frac{\pi}{2}$ for an interface stiffness from zero to infinite. But the stiffness of the contact cannot be known

a priori as the bearing lubrication regime depends on applied load, bearing speed, material properties and the roughness of contact surfaces. As an effect of the phase change, the timing of the zero-crossing points of the reflected wave changes in addition to the change caused by the load dependent wave path and speed. Therefore, in this case using a zero-crossing technique results in an inaccurate measurement of the ToF.

The effect of the phase change on altering the zero-crossing points of the reflect wave is schematically shown in figure 4.6. Three ultrasonic signals from a surface in different conditions are presented. The reference is a reflected signal from a surface in unloaded and free (out of contact with other components) state, and the signal with small phase change is reflected from the surface in a lightly loaded mating condition in which a thick oil film layer is formed. The signal with a phase change of $\frac{\pi}{2}$ is a reflection from the surface in a heavily loaded mating condition in which the oil film layer becomes vanishingly thin. In the schematic figure 4.6a and b, the principal interest is given to the phase shift and it is assumed that there is no length change in the wave path, which means only the ToF change resulted from the phase change is shown.

The phase change results in an apparent change in the ToF, which is denoted by Δt_ϕ. It depends on the phase difference between the reference signal (in the unloaded state) and the reflected signal in the loaded state, ϕ_R, and the frequency of the wave, as:

$$\Delta t_\phi = \frac{\phi_R}{2\pi f}. \tag{4.20}$$

In rolling bearing contacts, the oil film is rather thin and the phase change is considerable, which can be predicted from equation (3.61). As an illustration of the magnitude of the phase effect, if the oil film between steels was 0.5 μm, the phase change would be 1.12 radians. When using an ultrasound transducer with nominal central frequency of 10 MHz, the ToF change from this phase shift effect would be 17.8 ns. This is significant compared to the ToF change from load and cannot be neglected.

Therefore, a technique is required to extract the ToF change caused by the phase shift from that caused by load via the change in the component thickness and the change in speed of sound. Despite the phase of the reflected pulse depends on the contact stiffness, the envelope of ultrasonic energy is not altered by the contact dependent phase shift. The phase shift from the contact stiffness can be removed by applying the ToF determination technique on the envelopes of the reference and the loaded signals.

The approach taken is to analyse the envelope of the reflected signal by applying the fast Hilbert transform. The Hilbert transform of a function $f(t)$ is defined by equation (4.21), the divergence at $t = \tau$ is treated by taking the Cauchy principal value of the integral (Bracewell

(1978)).

$$\mathscr{H}[f(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(\tau)}{t-\tau} \, \mathrm{d}\tau \tag{4.21}$$

It can be seen that the Hilbert transform is a linear function of f (t) and can be obtained by convolution with $1/\pi t$. Therefore, the Hilbert transform of a function can be calculated from equation (4.22) where the $*$ denotes the convolution operation.

$$\mathscr{H}[f(t)] = \frac{-1}{\pi t} * f(t) \tag{4.22}$$

An analytic signal is a complex time signal whose real part is the original signal and the imaginary part is the Hilbert transform of the original signal. It can be represented by equation (4.23). The signal measured by the ultrasonic transducer is the 'real' component of the ultrasonic wave. The ultrasonic wave, however, has an additional 'imaginary' component which can be obtained by applying the fast Hilbert transform. The result of this process is to obtain the time dependent complex vector of the ultrasonic wave. By calculating the time dependent magnitude of this complex waveform vector, the envelope of the wave, which can be considered analogous to a time based energy profile of the reflected wave, is obtained.

$$x(t) = f(t) + i\mathscr{H}[f(t)] \tag{4.23}$$

The Hilbert transform has been widely used in signal analysis. Adhikari et al. (1998) applied the cross-correlation approach to the envelope of local field potentials to obtain the time lag between brain areas, they found the results obtained corresponded well with the time lag obtained from the time lag calculated from the phase locking approach. The Hilbert-Huang transform (Huang et al. (1998)) has been successful in the non-linear and non-stationary signal analysis. This approach is a combination of the empirical mode decomposition and the Hilbert transform. The time series data first is decomposed into a finite and often small number of intrinsic mode functions (IMFs), based on the local characteristic time scale of the data. By applying the Hilbert transform to each IMF, the energy-time-frequency distribution can be calculated. This is referred to as the Hilbert spectrum and then the system characteristics can be identified. In this work, the envelope time trace has been used instead of the energy-time-frequency distribution.

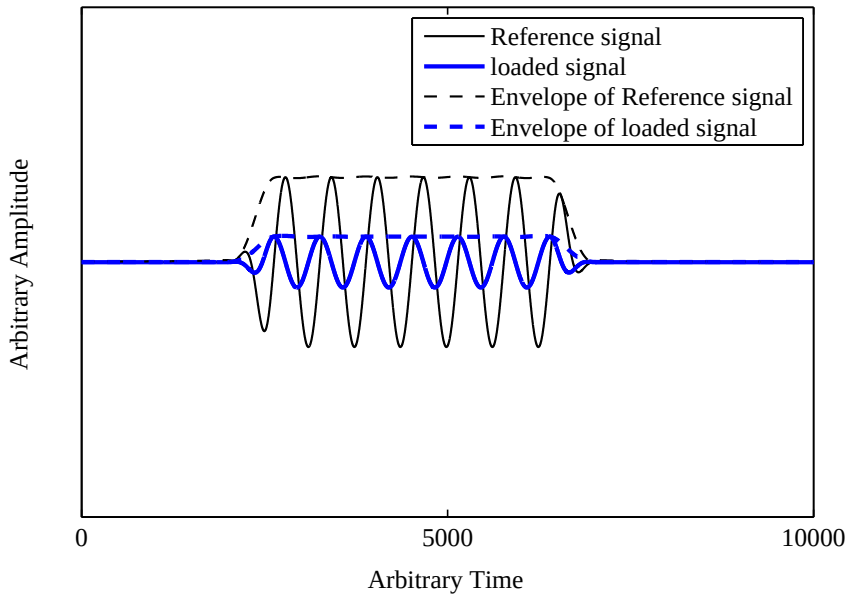


Figure 4.11. *Simulation of the phase shift between reference signal and reflected signal under load. The phase shift between these two signals is evident but there is no phase shift between their envelopes obtained through the Hilbert transform.*

As an example, figure 4.11 shows the effect of phase shift on the envelopes of the two simulated signals in figure 4.6b. The reference signal is a simulated reflection from a free surface, and the loaded signal is a simulated reflection from the same surface under highly loaded contact having a phase shift approaching to $\frac{\pi}{2}$. From the figure, the phase shift between two signals is clear and appears as a modification of the ToF: the solid blue line shifts to the left relative to the reference. Despite the phase change in the reference and the loaded signal, there is no time shift between the envelopes. The extent of the two envelopes are independent of the phase of the signal, though the amplitude will vary depending on the proportion of ultrasonic energy which has been transmitted through the contact into the mating component.

By cross correlating the reference signal envelope with the measured signal envelope, the time shift of the ' energy packet ' of the wave is achieved, which corresponds to the change in the ToF as a result of the geometric deflection and the acoustoelastic effects alone. The time shift of the ' energy packet ' is the net change in the ToF which excludes the effect of the phase shift.

4.5 Total ToF change

The total change in the ToF between the reflected ultrasonic waves is thus a summation of the ToF change from the deflection, the change in ultrasound speed and the phase change, given by :

$$\Delta t = \Delta t_\delta + \Delta t_c + \Delta t_\phi, \tag{4.24}$$

where Δt_δ and Δt_c are directly load dependent, and Δt_ϕ depends on the contact condition which indirectly depends on load. If the ToF change resulted from load is denoted by Δt_t, which is the sum of Δt_δ and Δt_c, by making use of equations (4.8) and (4.17) it can be expressed as :

$$\Delta t_t = \frac{2(1-L_{zz})\delta}{(c_{zz})_0}. \tag{4.25}$$

Substituting equations (4.25) and (4.20) into equation (4.24), then the total time difference between the reflected ultrasonic pulse from a free surface and the reflected pulse from a loaded contact surface becomes:

$$\Delta t = \frac{2(1-L_{zz})\delta}{(c_{zz})_0} + \frac{\phi_R}{2\pi f}. \tag{4.26}$$

By way of an example, the predicted ToF change from a cylindrical roller bearing contact is calculated. Table 4.2. shows the bearing geometry and properties used for the calculation of the raceway deflection, which can be obtained from the empirical equation (2.64).

Table 4.2: Cylindrical roller bearing geometry and properties

Symbol	Description	Value
OD	Diameter of the inner ring, m	0.367
ID	Diameter of the outer ring, m	0.259
D	Diameter of roller, m	0.054
l_0	Effective length of roller, m	0.079
N	Number of rollers	15
E	Modulus of Elasticity, GPa	207
ν	Poisson's ratio	0.3
d_0	Thickness of inner raceway, m	0.0195

The deflection from equation (4.4) has been used in equation (4.8) and (4.17) to obtain the ToF change contributed by the deflection and the ultrasound speed change, respectively. The oil film thickness in the roller bearing has been predicted from the well-known Dowson and Higginson equation (Dowson and Higginson (1966)) and then used to obtain the stiffness. This in turn has been used to calculate the apparent ToF change caused by the phase shift from the contact stiffness. Figure 4.12 shows the contribution of the three components as load on the most heavily loaded roller-race contact increases from 27.2 to 272 kN, corresponding bearing load increases from 100 to 1000 kN. The bearing is operated at 100 revolutions per minute (rpm) and lubricated by a Hyspin oil VG32.

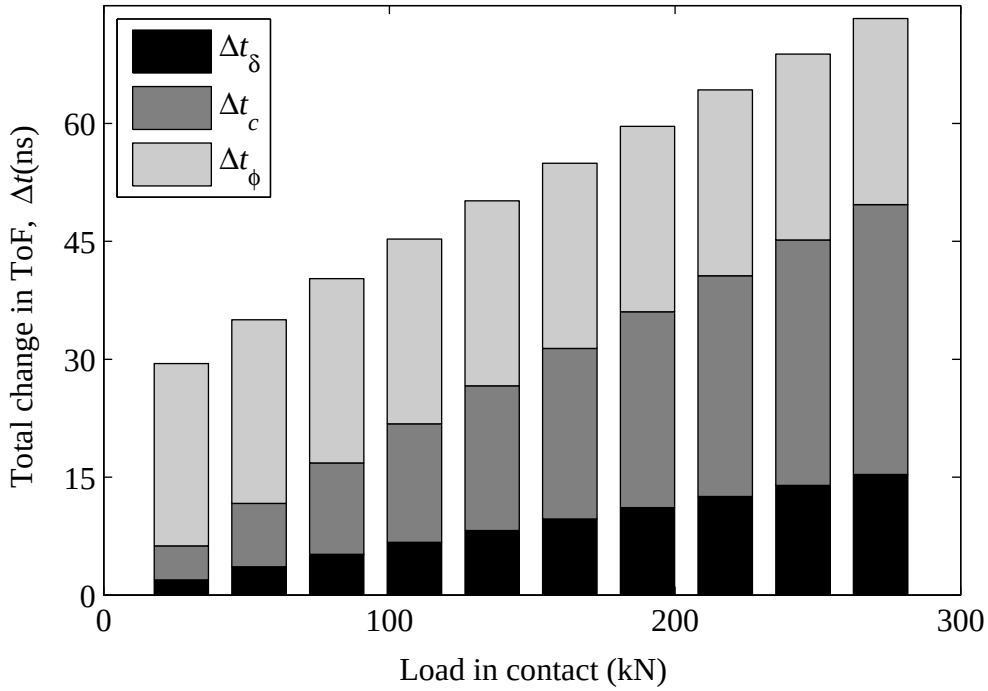


Figure 4.12. *Predicted change in the ToF for a roller bearing contact under various loads. For each load, the change in the ToF contributed by the surface deflection, the acoustoelastic effect and the phase shift are given.*

The change in the ultrasonic ToF is small reaching a maximum of ~ 70 ns at the highest bearing contact load of 272 kN. It is clear that all three components are significant and cannot be neglected. At low load, the apparent the ToF change due to the phase change dominates. This remains almost constant as the load increases (the oil film thickness shows comparatively weak dependency on load but a large dependency on speed and lubricant viscosity) whilst the contribution from the deflection and the change in the ultrasonic speed increase proportionally.

4.6 Conclusion

In this chapter, an analytical model for the variation of the ultrasonic ToF in one of contact bodies with load has been developed. The ToF of ultrasound in the contact element is determined by load upon the contact via the surface deflection and the acoustoelastic effect. In practical ultrasonic measurements, the phase shift occurs in the reflected pulses, which distorts the signal and gives an apparent time shift, making simple peak to peak time determination technique unsuitable. Although the phase shift between a signal and a reference

is apparent, there is no phase shift in their envelopes. Therefore, the Hilbert transform can be used to calculate the envelope of signal and thus the phase change effect is eliminated. The time difference between the envelope of loaded signal and that of the reference, is then only caused by the surface deflection and the acoustoelastic effect. The net change in the ultrasonic ToF is related to load and this relationship is used to interpret measured ultrasonic data as load in the following chapters.

Chapter 5

Development of Ultrasonic Measurement System for Load Monitoring

Chapter 4 describes an ultrasonic ToF change model for tribological contacts on the basis of the load-dependent surface deflection and acoustoelastic effect as well as the contact dependent phase change effect. This chapter covers the development of the measurement system. The first section describes the ultrasonic apparatus, including the detailed development of transducers. The next section describes the software used for data acquisition, followed by the description of the signal processing techniques to measure the change in the ultrasonic ToF. The ultrasonic measurement system will be used in Chapter 6 for monitoring load in a model line contact, and in Chapter 7 and 8 for monitoring load in real bearing contacts.

5.1 Ultrasonic apparatus

The ultrasonic apparatus is shown schematically in figure 5.1. It includes an ultrasonic pulser-receiver unit, a digitizer and an ultrasonic transducer, and a data acquisition system. The heart of the measurement system is the ultrasonic pulser-receiver unit as it plays a significant role to communicate with the transducer.

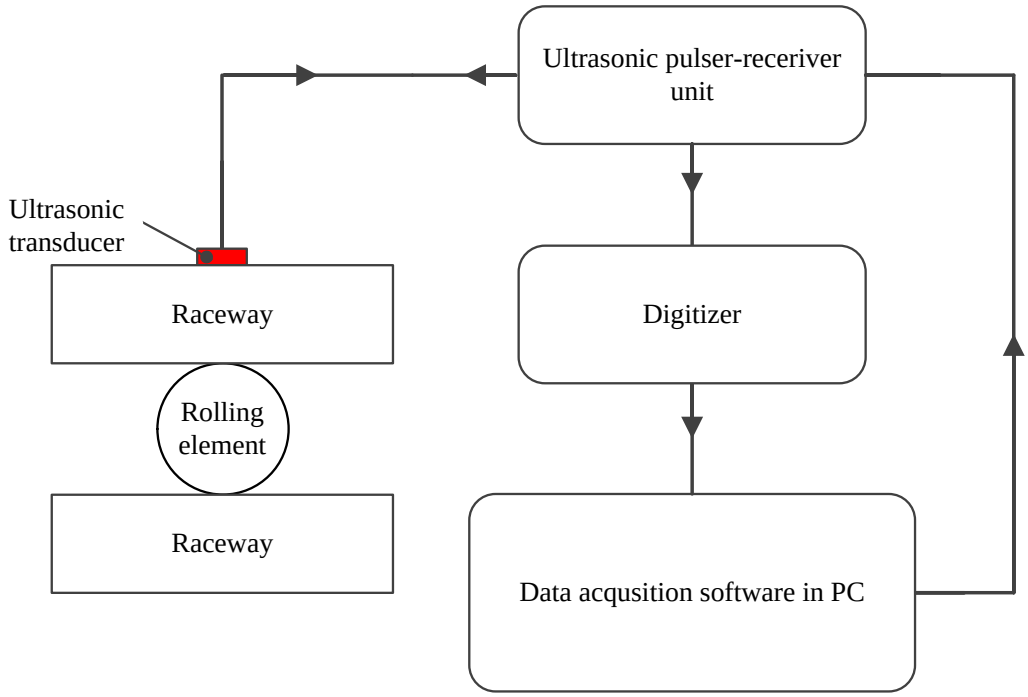


Figure 5.1. *Schematic of the ultrasonic measurement apparatus*

5.1.1 Pulser

A pulser unit is required to excite the piezoelectric element. A short duration voltage pulse from the pulser is sent to excite the transducer, and the ultrasonic wave generated by the transducer is determined by the configuration of the activating pulse.

The frequency of ultrasound is inversely proportional to the pulse duration time. To obtain maximum acoustic energies, the piezoelectric element should be excited at its fundamental natural frequency which is determined by the thickness in the through thickness resonance mode. For the case of applications requiring high frequencies, it is possible to excite the element at its natural frequency and use the upper harmonic frequencies.

The amplitude of the displacement in the incident wave is determined by the voltage of the excitation pulse. The wave reflected from the contact has a much smaller amplitude compared with the incident wave due to the wave diffraction and the material attenuation. The voltage of the pulse should be less than the element's poling voltage but large enough to have clear and distinct reflected signal from the contact surface.

The repetition rate of ultrasonic pulse needs to be carefully considered for dynamic tests, especially those involved with high speed rotation. The number of measurements from a small contact area is determined by the rotation speed and the pulse repetition rate.

The pulser unit can be a function generator, a dedicated pulser unit or a Peripheral Com-

ponent Interconnect (PCI) card. A PCI card pulser was used in this study and it was housed directly in a customized integrated PC made by Tribosonics. The unit have 8 channels in total and the voltage can be provided from 5 to 300 volts. The shape of the pulse shape is predefined in the form of inverse 'top-hat' while the pulse width and the pulse rate for each channel can be configured individually. The pulse width can be be changed from 50 to 1000 ns to excite ultrasonic transducers with central frequency from 1 to 20 MHz. The pulse repetition rate for all 8 channels changes from 160 to 80 kHz, and the pulse rate for a specific channel can be obtained by dividing the defined pulse rate by the number of used channel.

5.1.2 Receiver

In this study, the ultrasonic sensor works in the pulse-echo mode in which the receiver is on the same line with the pulser. A switch circuit between the pulser and receiver was used. Once the pulsing voltage signal was sent to the sensor, the circuit switched to detect the reflected signal from the sensor. The receiver unit used in this study is also a PCI card from Tribosonics, which is housed in the integrated PC.

Piezoelectric sensors generate ultrasound when it receives a short duration electrical signal from the pulser. Ultrasound travels in the component and part of it is reflected back when meets a boundary. The received ultrasonic wave is converted by the transducer into an electrical signal and sent to the receiver.

5.1.3 Digitizer

The received electrical signal is an analog signal and need to convert into digital signal so that the computer can process. It can be done by a digital oscilloscope or a PCI card. To properly represent the analogue wave, the sampling rate is required to be 10 times higher than the signal frequency. For instance, the digitizing rate need to be over 100 MHz to properly approximate a 10 MHz signal. The digitizer used in this work is also a PCI card housed in the integrated PC, having a digitizing rate of 100 MHz at a resolution of 12 bits. For higher accuracy, a digital oscilloscope with higher sampling rate can be used.

A personal computer was used for controlling the pulser and receiver unit and for data acquisition. A PCI cards based PC system, housing the customized pulser and receiver cards together with the digitizer card, from Tribosonics was used in this study. A LabVIEW program was developed to control the pulser receiver unit and the data acquisition system. The pulse voltage, the pulse repetition rate, the pulse width, the variable gain, the length of the signal can be easily configured in the program.

To maximize the overall speed of the integrated measurement system, the pulse rate and the length of signal for digitization need to be considered. The acquired data can be processed either on-line or off-line, depends on the measurement speed required. Simple data process, such as the determination of the ultrasonic speed or the component thickness, can be done on-line. For the case of dynamic test which required high pulse rate and sophisticate data analysis, it is more suitable to store the data for post processing as it reduces the whole measurement system rate.

5.1.4 Development of ultrasonic transducers

Bare piezoelectric elements are used in this work because of their comparably low cost. A typical ultrasonic transducer costs several hundred pounds while a bare piezoelectric element only costs several pounds. Another advantage of using bare piezoelectric sensors is their low profile. In real bearing applications, access to bearing contacts is quite challenging. The thickness of piezoelectric elements can be less than a millimeter assuring they can be used for bearings without disrupting the involved system. All ultrasonic measurements in the following chapters have been carried out with bare piezoelectric elements.

The selection of piezoelectric elements needs to consider several aspects, such as the central frequency, the operating temperature,the configuration of electrodes, and the dimension.

The principle for choosing the sensor central frequency is to have good penetration in the interested material and also clear distinct reflected signals from the interested interface. For all work carried out in this study, the component in our interest was steel and the nominal center frequency of the piezoelectric elements was chosen at 10 MHz. The piezoelectric elements were formed from lead-zirconate-titinate, soft PZT navy type II, providing with high sensitivity, good thermal and time stability. This material can be operated at a temperature up to 185°C.

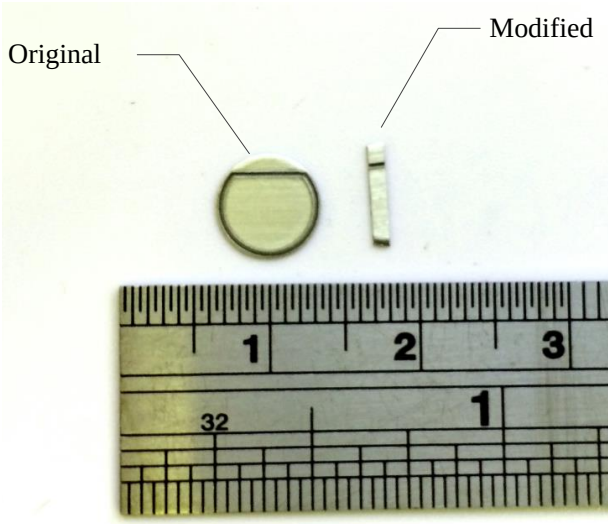


Figure 5.2. *Photograph of the piezoelectric element before and after modification: the left is the original disc form and the right is the modified narrow strip form.*

Figure 5.2 shows a photograph of the element before and after the modification. The piezoelectric elements used in this work are in a disc form and the positive electrode is in a D-shape configuration. The negative electrode is the remainder of the disc. The advantage of this configuration, having two electrodes on the same surface of the element, is to provide a more even bond between the piezoelectric element and the component which avoids wave energy at the bond layer.

From equation (3.1), the thickness of a 10 MHz piezoelectric element is about 0.2 mm. The diameter of the element is 7 mm but rolling bearing contacts are quite small. The contact width of the elliptical contact in ball bearings can be calculated by using equation (2.18) or the contact width of the line contact in roller bearings can be calculated from equation (2.50); either of them is less than 2 mm in this study. As such a case, the measured signal consists of not only reflected signals from the interested contact but also reflected signals from the area around the contact. To improve the spatial resolution, the piezoelectric elements are modified.

In this study, the rolling contact is in the form of either ellipse or narrow strip so the elements were cut to form a narrow strips having a width of 1 mm in the rolling direction. Figure 5.3 shows a schematic diagram of the modified element. After the modification, the active region of the element was determined by the positive electrode, 5 mm in length and 1 mm in width.

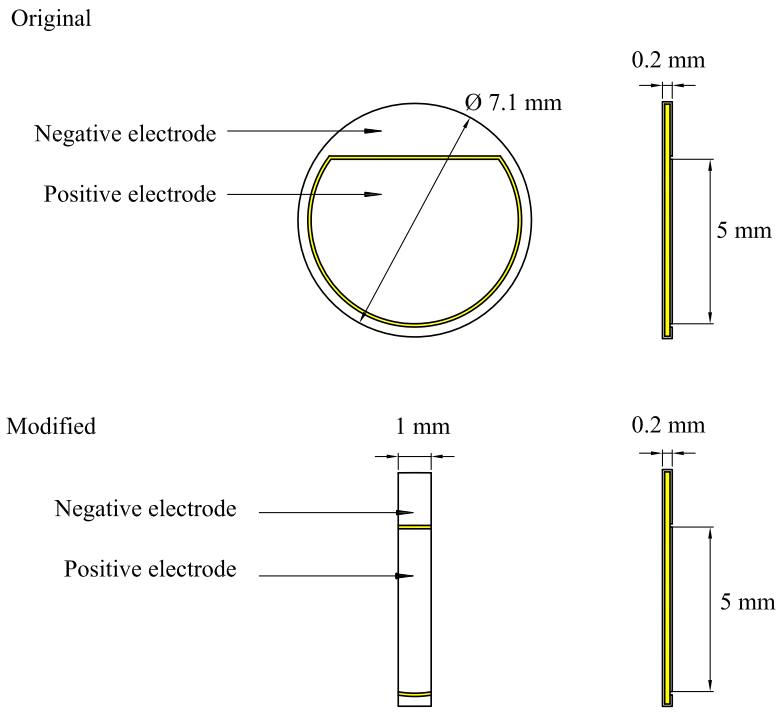


Figure 5.3. *Schematic of the piezoelectric element before and after modification.*

As stated in Chapter 3, the response of ultrasonic waves from the thin layer sandwiched between two media is governed by the mechanical and acoustic properties. The properties of the coupling layer between the sensor and the specimen determines the ultrasonic wave transmission between them. Good coupling is essential to transmit as much energy as possible from the sensor to the component, and vice verse. Ideally, the adhesive layer thickness is equal to an odd number of the quarter of the wavelength, which ensures 100 per cent transmission from the sensor to the subject.

An adhesive from Micro-Measurements (M-Bond 610), was used to attach the modified piezoelectric element on the subject surface (bearing race surface in this work). It is characterized with the low viscosity and the capability to form an extremely thin layer, as thin as 5 μm. Another advantage of the use of this adhesive is able to bear high temperature up to 230°C after proper Curing treatment (Vishay (2010)). The Curing process of the adhesive was done in an oven for 4 hours at a temperature of 150°C with a clamp to provide required pressure. Piezoelectric elements are very fragile so a flexible silicone rubber was used to prevent damage from the clamp. Under the applied clamp pressure, the low viscosity of the adhesive ensured the glue was squeezed to a very thin layer.

Cables are required to build up an electrical circuit. A coaxial cable was wired at the positive and negative part of the element which connects the piezoelectric sensor with the

data acquisition system. To protect the sensor and the wire connection, a high temperature epoxy resin was used. The additional benefit of the epoxy layer is to provide effective damping for the sensor, resulted in a wider bandwidth and hence a higher sensitivity. The epoxy layer also fixes the cable on the subject, reducing the signal variation caused by the cable vibration.

5.2 Data acquisition

Figure 5.4a schematically shows a raceway-rolling element contact with a fixed ultrasonic transducer on the raceway. When the rolling element is away from the sensor, ultrasound is totally reflected back and a typical waveform is given in figure 5.4b.

The first peak is the sensor initial pulse and the second peak is the first reflected pulse from the bottom surface of the race. Followed by a series of reverberations in the raceway. Therefore, the time interval between two consecutive echoes is the ToF of ultrasound in the raceway which depends on the thickness and the ultrasound speed in the race. As the length of the signal to be digitized and stored affects the pulse repetition rate, the interested part of the signal is isolated and stored for data analysis later.

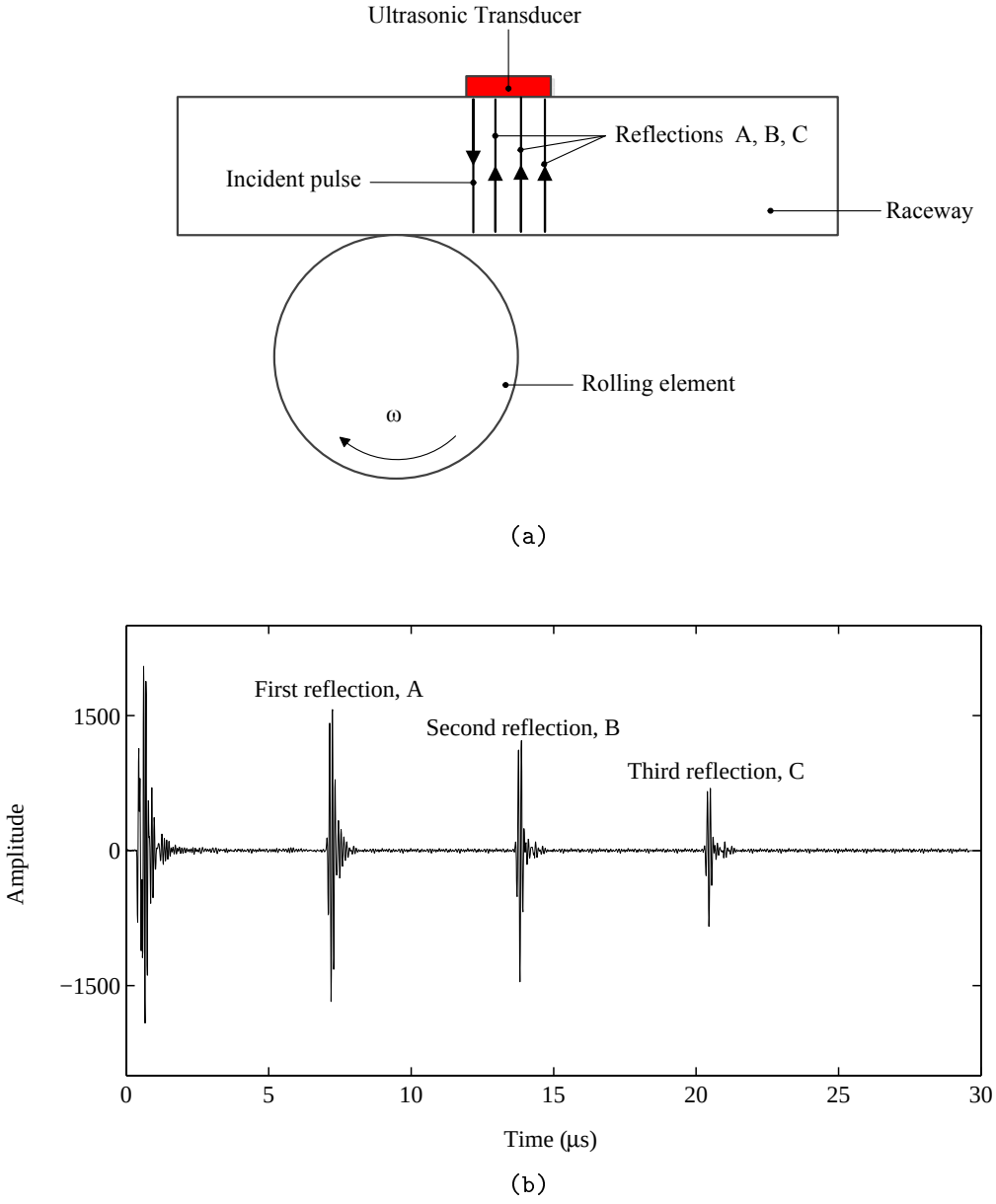


Figure 5.4. *(a) Schematic of the raceway/rolling element contact equipped with an ultrasonic transducer and (b) typical waveform from a 10 MHz sensor bonded on the raceway*

A data acquisition software has been developed in the LabVIEW environment, schematically shown in figure 5.5. The software has been incorporated with the configuration of the pulse-receiver unit, so that the setting for the pulser-receiver in each test can be properly recorded. The next step is to identify the interested part of the reflected signal. A signal window is used for isolating the interested part from the waveform. It can be moved in the time trace by changing the delay of the signal window. The length of the window (data range)

can be changed according to the desired part of the signal. Then the extracted part of the waveform is stored in a first-in-first-out buffer and sequenced into an array. The sequenced arrays and configuration details can be recorded simultaneously by using the data type of cluster in the LabVIEW. A series of clusters are generated during the acquisition period and they are formatted into a binary file to be stored for post data analysis.

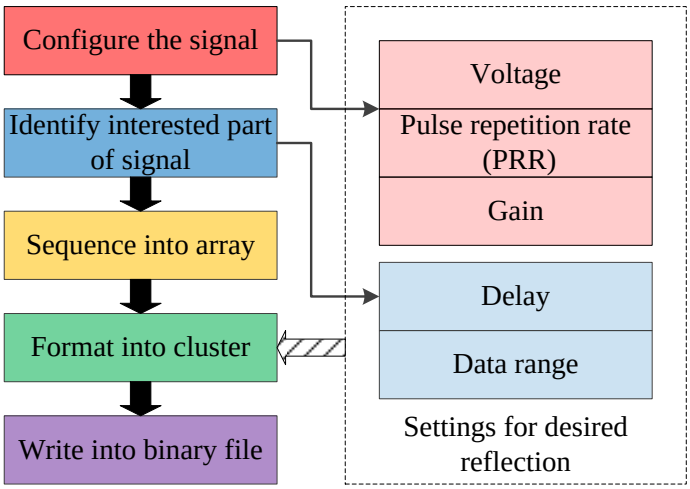


Figure 5.5. *Schematic of data acquisition software.*

Figure 5.6 gives an example of data acquisition for ultrasonic measurement in the cylindrical roller bearing. Since the ToF in one of the mating components is of interest, the first two echoes, the initial pulse and the first reflection from the contact face of the component, having maximum signal to noise ratios, are acquired.

When the contact was far away from the sensor, i.e. ultrasound traveled in the unloaded race, the first reflection had the maximum amplitude and the time interval between first two peaks achieved the maximum. When the rolling contact entered into the sensor window, part of ultrasound was transmitted into roller and the amplitude of the first reflection was largely reduced. Meanwhile, contact stress was introduced in the mating surfaces under the sensor window, deformation and change in ultrasound speed took place, resulting in a reduction in ToF (the first reflection shifted slightly to the left). When the contact exited the sensor window, the amplitude of the first reflection gained its maximum again and the ToF increased back to the original level. The change in ToF caused by the passage of the contact made the first two peaks behaved like compressing a spring, where the first peak corresponds to the fixed end of the spring.

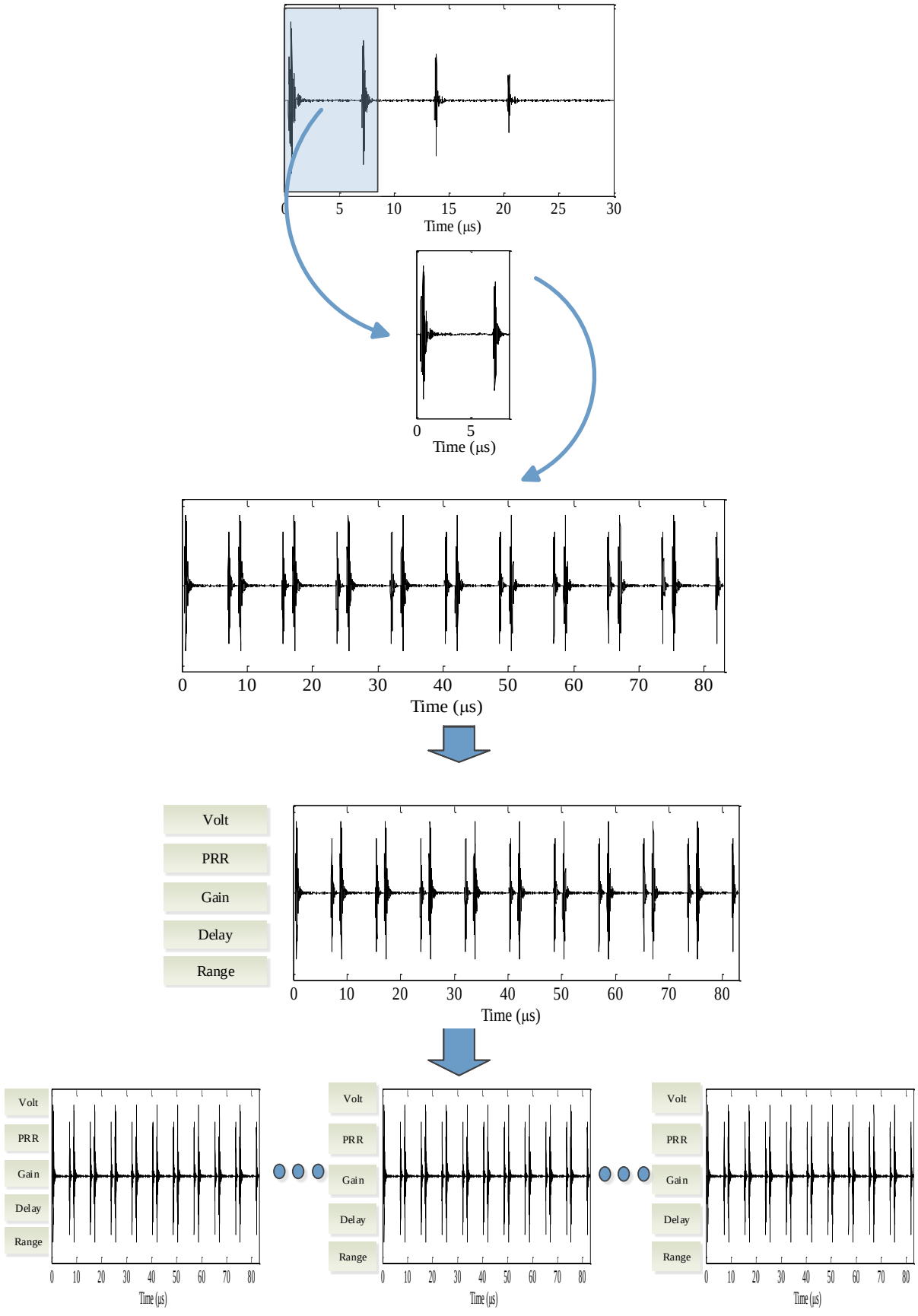


Figure 5.6. *An example showing the data acquisition process (Note: the location and the length of the signal window can be configured).*

5.3 Signal processing

After the signal is digitized and passed to the PC, the data analysis can be performed instantaneously or in post process determined by the operator. As stated in Section 5.1.3, the signal processing for all experiments in this work was carried out in post process.

5.3.1 Signal processing routine

Figure 5.7 schematically shows steps required to obtain ultrasonic ToF, followed by a detailed signal analysis routine.

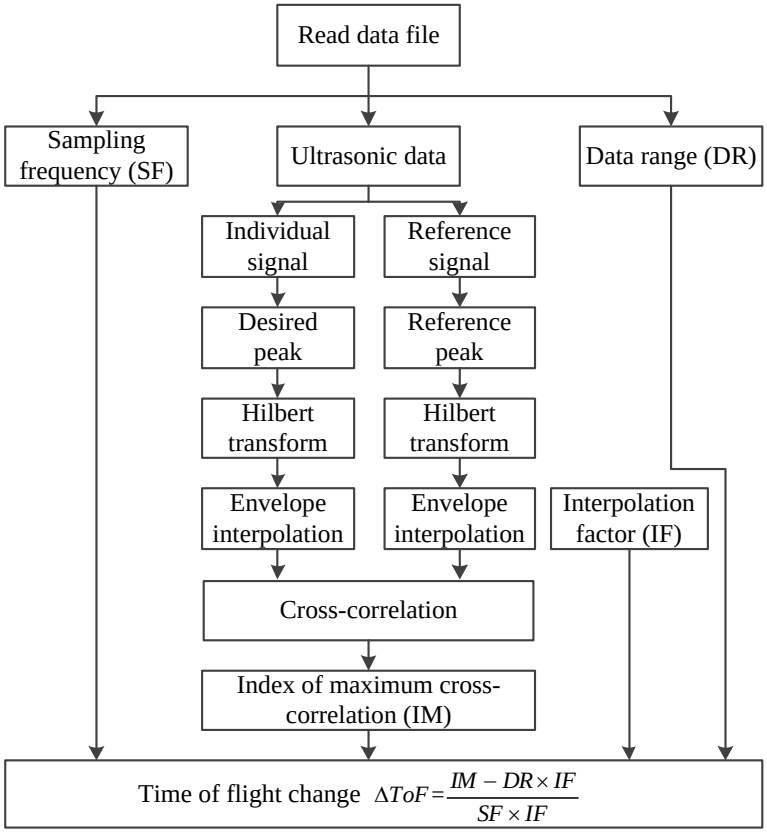


Figure 5.7. *Schematic diagram showing the signal processing routine.*

- The process starts with reading the ultrasonic data from binary files. As both the configuration information and the ultrasonic data are recorded, the first thing is to extract the ultrasonic data. The data can be extracted and stored in an array or many arrays depending on the size. The reflected signals for three typical stages in the rolling process are given in the figure 5.8 : a) when the contact was far away from the

sensor location; b) when the contact approached closely to the sensor ; c) and when the contact was right underneath the sensor window. In the figure, the first reflection from the back face of the component is represented by A for case (a), A' for case (b) and A'' for case (c).

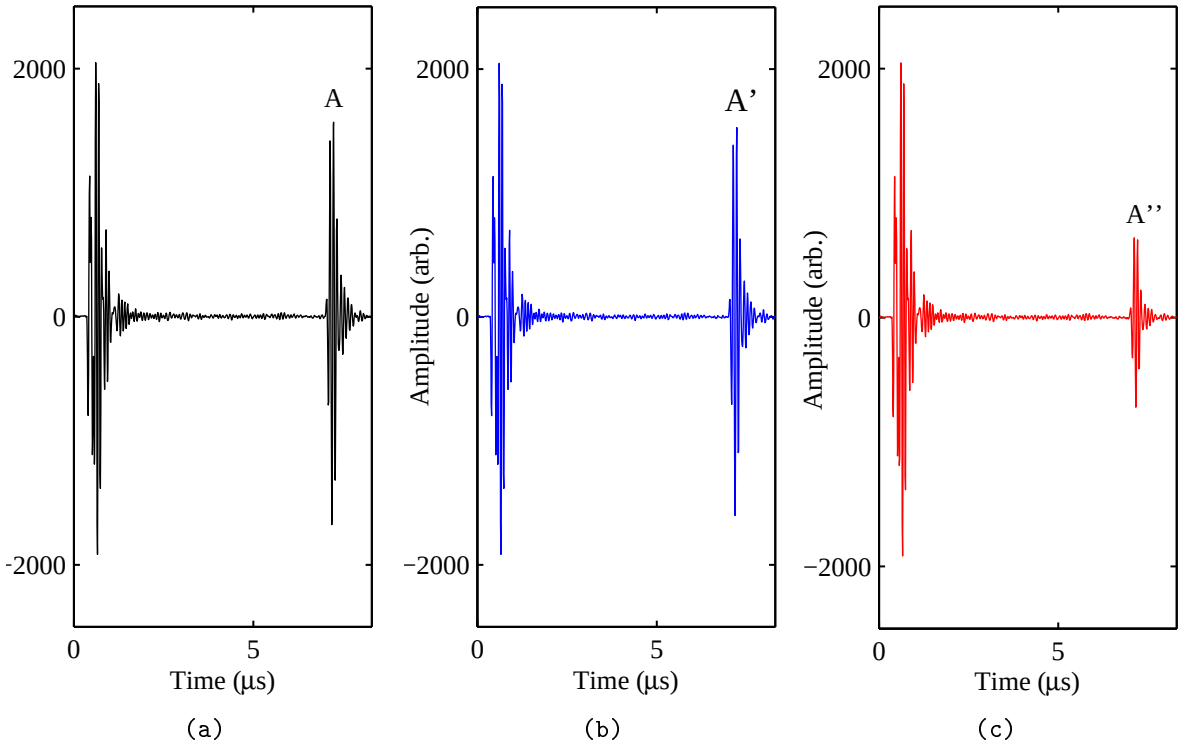


Figure 5.8. *Signals for three typical phases when a rolling contact passes the sensor location: a) the line contact is away from the sensor; b) the line contact is approaches closely to the sensor; c) the line contact is right underneath the sensor.*

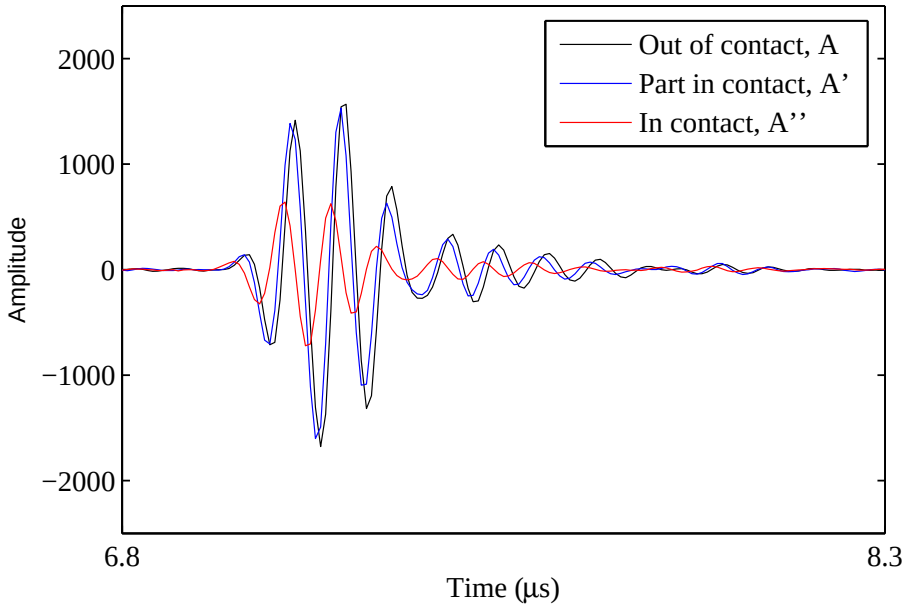


Figure 5.9. *The first reflections when the roller at three different positions under baring load of 980 kN.*

- The next step is to find the signal reflected from the unloaded race which can be taken as the reference signal in the following analysis. The reference signal can be acquired when the inner race is disassembled or can be obtained from the running test when the contact is far away from the sensor. In practical, it is more convenient to take reference when the contact is away from the sensor. The rolling elements are circumferentially, evenly distributed between races and the distance between two elements is considerably large compared with the contact patch in bearings. Therefore, the majority of the acquired signals correspond to the signals reflected from the race-oil interface, and they can be used as the reference. There is a phase change between reflection from race-air interface and the reflection from the race-oil interface, but it will be eliminated in the next steps.

- The change in the ultrasonic ToF presents as the modification of the arrival time of the first reflection, thus it can be obtained by comparing the first reflection in loaded and unloaded states as shown by figure 5.9. However, as discussed in the previous chapter, the contact condition especially the contact stiffness introduce phase shift in the reflected wave, which represents as a change in ToF combing with that from deformation and acoustoelastic effect.

- Although the phase change affects the signal, it does not affect the signal envelope.

As described in Chapter 4, the envelope of the signal can be used instead to eliminate the effect of phase shift on the ToF, and the envelope of signal can be calculated by using the fast Hilbert transform. The envelopes of signal in figure 5.9 are given in figure 5.10.

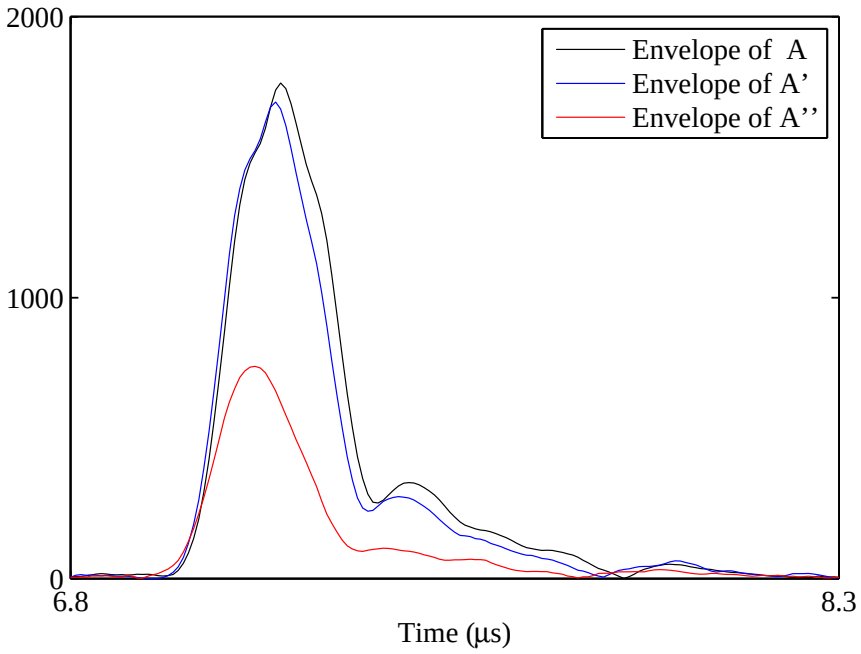


Figure 5.10. *Envelope of signal from out of contact, part in contact and in contact from the cylindrical roller bearing*

- During the signal analysis procedure, one common method to determine the time delay between two signals is to compute their cross correlation function, which is defined by equation (5.1) (Knapp and Carter (1976)). The time delay can be obtained from where the cross correlation function achieves the maximum. By cross correlating the envelope of the reference signal with the envelope of the reflected signal from the loaded contact, the change in ToF as a result of the geometric deflection and the acoustoelastic effect was obtained. A cross-correlation of the envelope of reference signal (A) and the envelope of loaded signal in contact (A") is given in figure 5.11. The lag between the envelopes can be measured by subtracting the length of envelope of A" (number of data points) from the index where the cross-correlation achieves its maximum. As an example, the data range for the envelopes in figure 5.10 is 151 (for 1.5 μs) , the cross-correlation achieves its maximum at 155, thus, the lag between the envelopes is 4 data points.

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)y(t-\tau)\,\mathrm{d}t \tag{5.1}$$

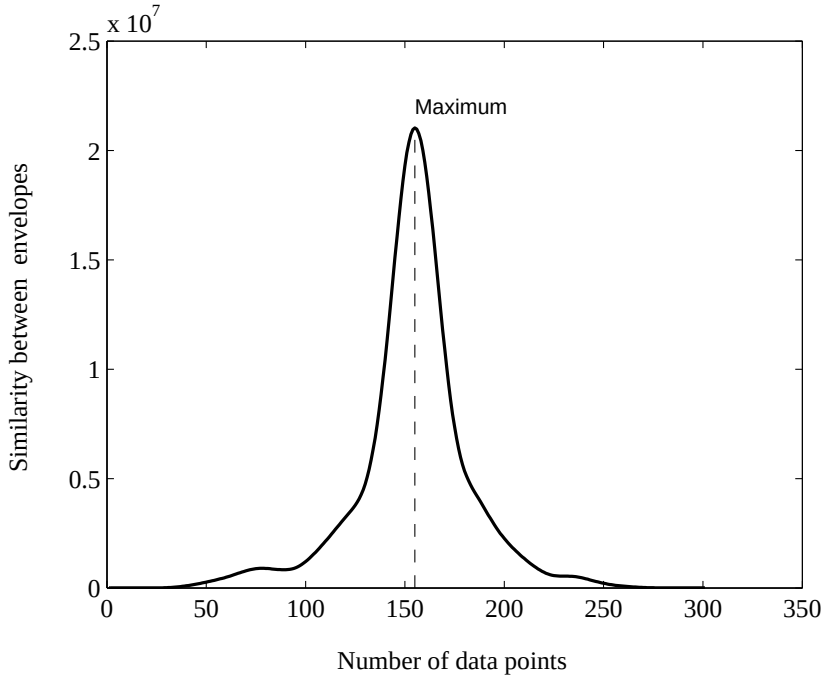


Figure 5.11. *Cross-correlating the envelope of the loaded signal with the envelope of the reference signal*

- By multiplying the time resolution between data points, which depends on the digitizer sampling rate, the change in ToF then can be determined. It can be seen that the accuracy of measured results is determined by the time resolution. There are two ways that can be used to improve the time resolution. The first is to utilize a digitizer with higher sampling frequency and the time resolution can be improved at a higher cost. Another way to improve the time resolution is to use an interpolation technique for the digitized data, which is more cost effective. The later was adopted in this work.

- The time difference between the envelope of the loaded signal and that of the reference is only caused by load since the phase shift has been excluded via the Hilbert transform. Once the ToF change for each load case has been measured, the deflection of the mating surfaces or load on the contact can be obtained using equations in Chapter 4.

5.3.2 Time resolution

In the previous section, the ToF change between the envelopes of the loaded signal and the reference is determined by using a cross-correlation technique, which is equal to the number of data points lag between two envelopes multiplying by the time resolution. For a sampling rate of 100M samples per second, the time resolution between two consecutive data points is 10 ns. As shown in Chapter 4, the total ToF change in rolling bearing contacts was small, 10's of nanoseconds, which requires better time resolution less than 10 ns.

As an illustration, the effect of the time resolution on the measured ToF change is shown in figure 5.12. In the top graph, the signals are from the roller bearing when it is under load of 980 kN; the reference signal is taken when the roller is away from the sensor and the loaded signal is taken when the roller is directly under the sensor position. By using the Hilbert transform, the envelope of the signals are obtained, as shown in the middle, left graph. The envelope of the loaded signal (the red line) advances to that of the reference signal (the dark line) about 4 data points by cross-correlating them, as shown in figure 5.11. As the time resolution is 10 ns, the time shift is measured at 40 ns.

The time resolution can be improved to 1 nanosecond when the envelopes are interpolated by a factor of 10. By cross correlating these interpolated envelopes, the lag is measured at 44 data points, which gives a measured time shift of 44 ns. The measured ToF change increases by 4 ns, through improving the time resolution to 1 ns with the interpolation (with a factor of 10).

The time resolution can be improved further to 0.1 ns when the interpolation factor increases to 100. The envelope of the loaded signal and that of the reference signal are given in the middle, right graph in figure 5.12; the lag between them are measured at 441 data points, illustrated by the bottom graph, giving a measured ToF change of 44.1 ns. The measured ToF change increases by 4.1 ns when the time revolution changes from 10 to 0.1 ns. This corresponds to a deflection of 12.09 microns in the race.

When the time resolution further increases from 0.05 to 0.005 ns (corresponding to a sampling frequency up to 200 GHz), by changing the interpolation factor from 200 to 2000, the measured ToF maintains at 44.08 ns. That means further improvement in the time resolution rarely change the measurement accuracy.

Currently, advanced oscilloscopes with high sampling frequency up to 200 GHz are available in the market but they are quite expensive (about several hundreds of thousands pounds). In this work, an interpolation technique was applied to the envelope of signals aiming at improving the time resolution. This was done in the signal processing routine in the LabVIEW program by using the interpolate 1D Fourier VI. In the interpolate1D Fourier VI, the signal envelope was transformed into the frequency domain and scaled according

to the desired interpolation factor(IF), and then transformed back into the time domain. Therefore, the length of the interpolated envelope was increased by IF times and the time resolution became $^{10}/_{IF}$ ns. Figure 5.12 gives a screen print of the front panel of the VI that was used to measure ToF change from ultrasonic reflections. It includes the calculation of the envelope of the loaded signal and that of the reference signal, the interpolation of the envelopes, and determination the time shift between the envelopes.

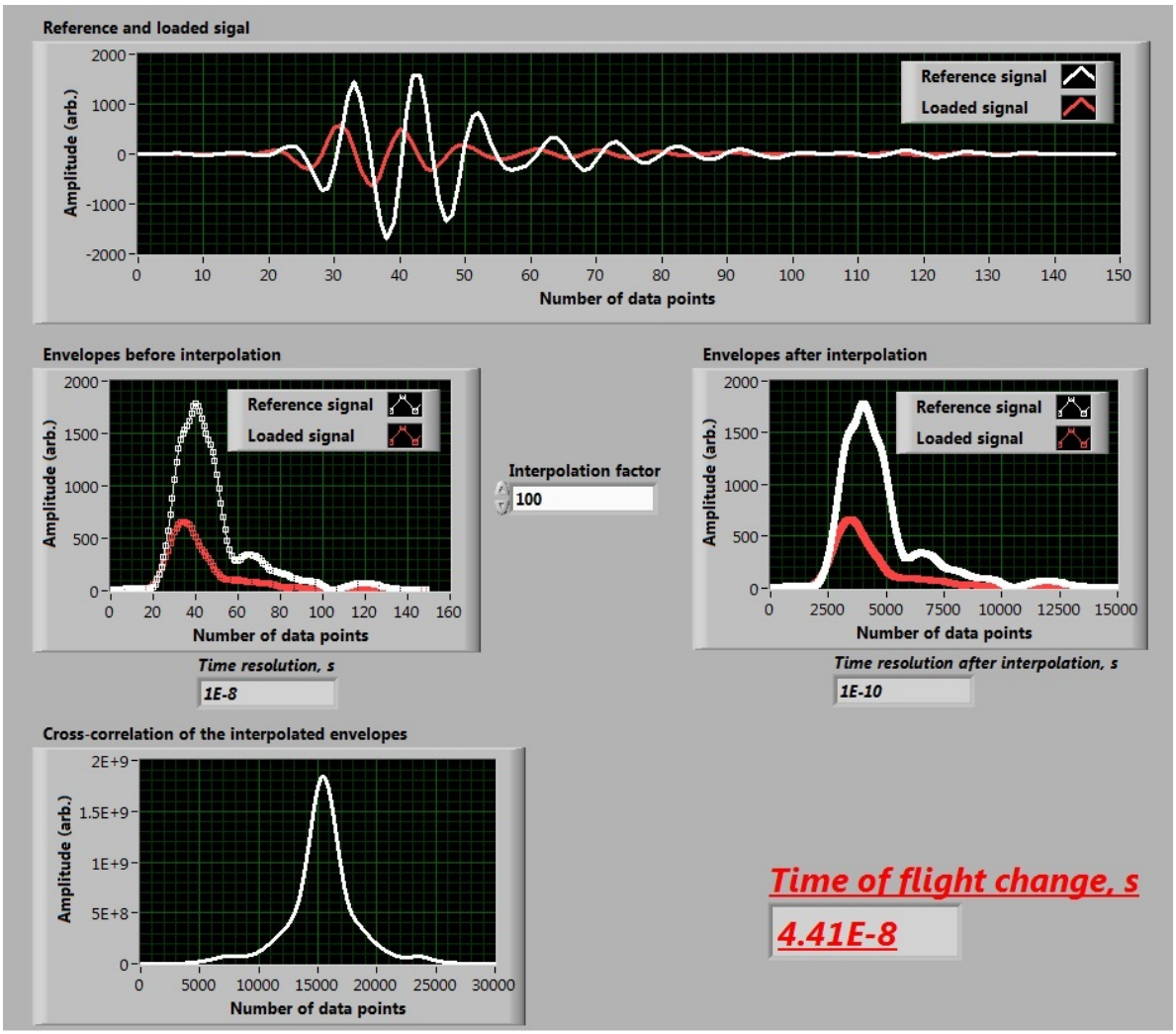


Figure 5.12. *A screen print of the front panel of an example VI used to measure the ToF change between two signals.*

It can be observed that the measurement accuracy is significantly improved when the time resolution varied from 10 to 1 ns. However, further increase in the time resolution has little effect in the accuracy of the measured ToF. In this work the interpolation factor was chosen to be 100 to provide time resolution at 0.1 ns.

The ultrasonic measurement system is limited to measure the ToF change larger than 1 ns although the time resolution can be 0.1 ns. This is determined by the ultrasonic transducer, the pulser-receiver unit and the data acquisition system. Figure 5.13 shows a measured ToF change from the inner raceway (the steel-air interface) over 2 seconds, which should be zero as the surface is unloaded. The change in the ToF up to 1 ns is observed; this is attributed to noise in the measurement system. The ToF technique shows better promise for application to large bearings when both the contact dimension and the raceway deflection are large and hence lead to readily measurable ToF change.

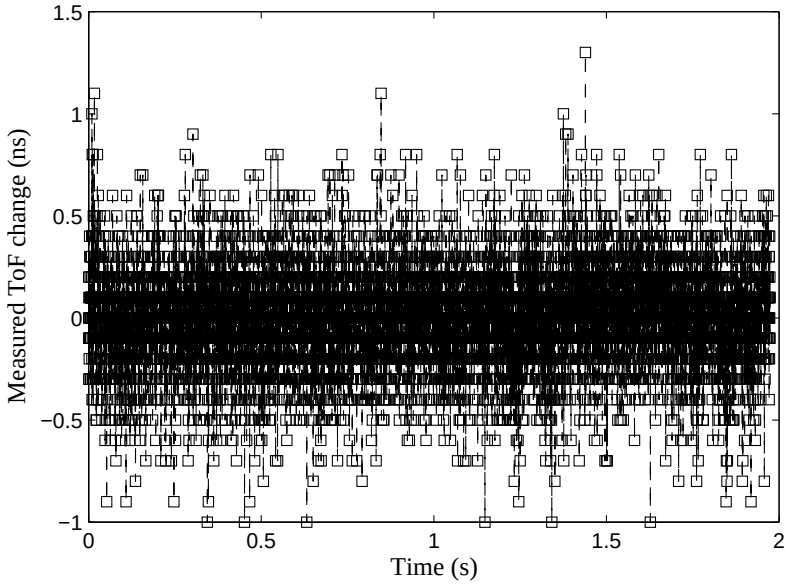


Figure 5.13. *Measured ToF change from the inner race of the roller bearing (steel-air interface) due to noise in the measurement system.*

5.4 Conclusion

This chapter described the ultrasonic apparatus for load measurement in rolling contacts, the data acquisition software and the signal processing routine. Bare piezoelectric elements were used as ultrasonic transducers due to their cost-effective and low profile advantages. They were modified to narrow strips with respect to considerable small contact patches in rolling bearings to provide high spatial resolution. The modified piezoelectric sensors were bonded directly on the specimen by using a high temperature adhesive. An ultrasonic pulser-receiver unit was used to excite the transducer and receive the reflected signal. The received signal was digitized at a high sampling rate before recorded in a PC by using a

data acquisition software developed in the LabVIEW environment. In the signal process, the Hilbert transform was used to calculate the signal envelope to remove the phase change effect, and the change in the ToF caused by load alone was measured. The measured change in the ToF can be converted to deflection or load by using the elastic model developed in Chapter 4.

Chapter 6

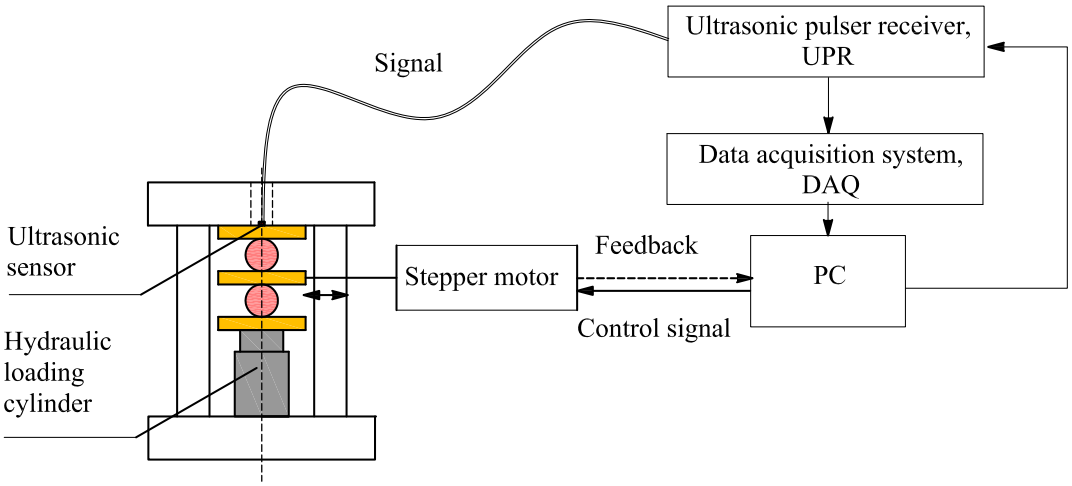
Load Measurements on a Model Line Contact

The ultrasonic measurement system developed in the previous chapter has been used for measuring load on a model line contact which is described in this chapter. The first section is devoted to the development of the mechanical instrumentation and the description of the ultrasonic apparatus. It follows by a series of ultrasonic reflections from the model line contact under different loads, from which ToF changes have been measured. The measured ToF change, caused by the deflection and the change in speed of sound, was related to contact load by using the elastic model developed in Chapter 4. Discussions for features observed in the experiment are given before the conclusion for this chapter.

6.1 Apparatus

6.1.1 Model line contact

Figure 6.1 schematically shows the model line contact apparatus and the ultrasonic instrumentation. Line contacts were formed between cylindrical steel rollers and steel plates under a normal load applied by a hydraulic cylinder. The dimensions of the plates and rollers are given in table 6.1. An ultrasonic transducer was bonded directly on the upper plate and connected to the ultrasonic pulser-receiver, aiming at investigating the line contact between the roller and the upper plate. A small window on the loading frame allowed for the sensor location and the wire connection.



`Figure 6.1.` *Schematic of the model line contact apparatus with ultrasonic instrumentation*

The maximum load capacity of the hydraulic cylinder was 100 kN, which was used for all load measurements in this chapter and load measurements in a ball bearing in Chapter 8. The cylinder was driven by a hand pump through a safety hose and the load applied by the cylinder was indicated by a gauge. In the experiment, the load reading from the gauge was calibrated with a high resolution load cell (Futek FSH00373). A shut off valve was used to provide constant hydraulic load for each load test.

Table 6.1: Dimensions of the plates and rollers

Specimen	Length(m)	Width(m)	Thickness(m)
Upper plate	0.12	0.075	0.0123
Middle plate	0.12	0.075	0.0123
Bottom plate	0.12	0.075	0.0123
Rollers	0.03	0.03(diameter)	

The model line contact in the figure can be indexed while the sensor had a fixed location.

The upper and the lower plate were fixed while the rollers were moved by the central plate. A stepper motor controlled by a LabVIEW routine was used to control the movement of the central plate as shown. In this way, the top roller could be rolled underneath the upper plate and then returned back. A photograph of the test rig is given in figure 6.2.

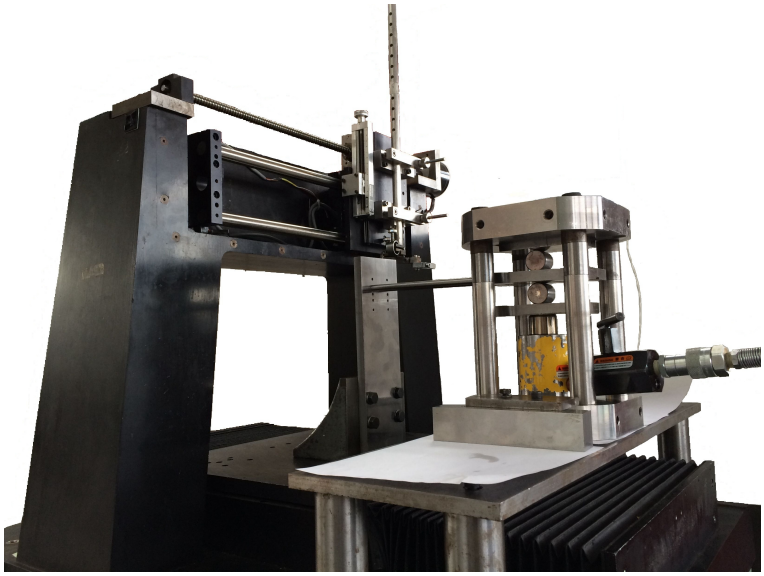


Figure 6.2. *A photograph of the model line contact test rig*

6.1.2 Sensor design

Piezoelectric elements with nominal centre frequency of 10 MHz were used as they provide good penetration into the steel and a clear distinct signal. As the model line contact is quite small in width, to improve the spatial resolution, piezoelectric elements were modified from the original disk forms into narrow strips having width of 1 mm, as shown in figure 5.1. The sensor was directly bonded on the upper surface of the upper plate using a high temperature adhesive. A layer of the epoxy was placed over the sensor to protect it and partially suppress the sensor ringing.

6.1.3 Instrumentation

The ultrasonic pulser receiver was used to generate short duration voltage pulses and also to receive the reflected signals. A 30V 'top hat' signal with duration of 100 ns was employed. An ultrasonic pulse was generated when the transducer was excited by the short duration voltage signal. Then ultrasonic pulse propagated through the upper plate, fully or partly reflected back at the contact face when there is no or a contact presents and received by the

same transducer. The reflected pulse was digitized at 100M samples per second by the data acquisition system and streamed directly to the computer for storage.

6.1.4 Experimental method

Experiments started with recording the reflected signal from the unloaded upper plate alone (steel-air interface), which was taken as the reference signal. The cylindrical rollers were then sandwiched between the steel plates and placed in the desired loading frame. Light normal load was first applied so that the upper plate was slightly loaded against the roller and an initial line contact was made. Then a number of loads were applied on the line contact and ultrasonic reflections from the contact face for each load case were recorded.

Figure 6.3a gives a schematic of the ultrasonic test when the upper plate is unloaded and free, and figure 6.3b shows the corresponding waveform. The reflected waveform consists of three consecutive reflected pulses from the back face of the upper plate, as the ultrasonic wave bounces three times in the plate before disappears and the ultrasonic transducer detects all these reflected pulses. This is illustrated by peaks A, B and C in figure 6.3b, the interval between each two consecutive peaks in the time trace is the ultrasonic ToF in the unloaded plate. In this case, the ToF can be easily obtained using simple time determination techniques.

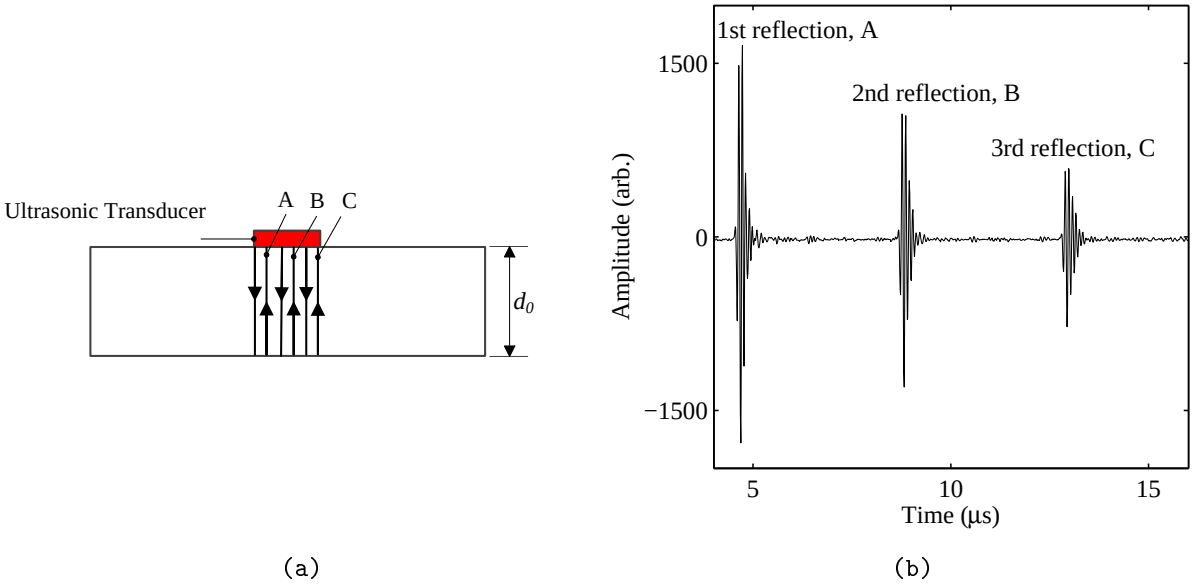


Figure 6.3. *(a) Schematic of the unloaded upper plate instrumented with a modified piezoelectric sensor (10MHz), (b) the reflected waveform consists of first three reflections from the back face of the plate.*

When the contact was directly underneath the sensor position and subjected to load, as shown in figure 6.4a, the corresponding reflected waveform is given in figure 6.4b. Again, it consists the first three reflections from the back face of the upper plate and the interval between each two consecutive peaks is ultrasonic ToF in the loaded plate. However, in this case, the time between two consecutive peaks cannot be measured by using simple time determination techniques. As described in Section 4.4.2, the loaded contact introduces phase change in the reflected pulses, which alters the location of the reflected pulse and appears as an apparent time shift in simple time determination process.

With the propagation, the wave energy will be lost as a result of the wave dispersion, the attenuation in the material and the transition at the interface. As a consequence, the amplitude of the ultrasonic wave reduces with the wave path. This can be clearly seen from figure 6.3b and figure 6.4b.

The amplitude of the first reflection A (A") is the largest as the wave goes through twice the thickness, followed by the amplitude of the second reflection B (B") which goes four times the thickness. The amplitude of the third reflection C (C") is the smallest as the wave travels six times the thickness. For the case of no contact presents, from Chapter 3 it is clear that the wave transition at the steel-air interface is so small that can be neglected. Thus, the reduce in the amplitude between pulse A and pulse B is due to attenuation in the material. For the case of a contact present, part of the ultrasonic energy is transmitted from the plate into the roller and so the amplitude of reflections reduce greater. This is illustrated by the amplitude of reflections A", B" and C" which are much lower than the amplitude of reflections A, B and C.

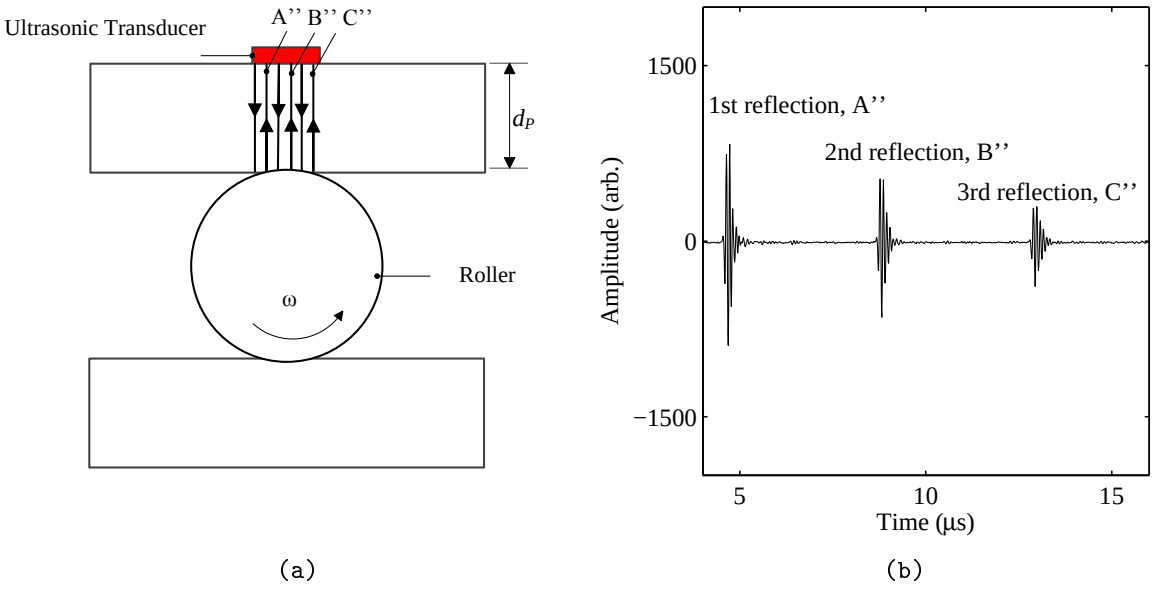


Figure 6.4. *(a) Schematic of the upper plate loaded by the roller, (b) the corresponding reflected waveform.*

Therefore, the reflection coefficient can be obtained by comparing the corresponding peaks figure 6.3 and figure 6.4. The amplitude of the reflection coefficient is equal to the ratio of the amplitude of peak A" to that of peak A. For the second reflected pulses, reflection happens twice at the back face of the plate, then the amplitude ratio between peak B" and B is a square of the reflection coefficient.

The ultrasonic ToF is determined by the wave path and the wave speed. As the surface deflection took place in the line contact when the upper plate is loaded, the thickness of the plate reduces from d_0 to d_P and the wave path varies from $2d_0$ to $2d_P$. Meanwhile, the ultrasound speed in the plate increases from c_0 to c_P due to the compression of the plate. As a consequence, the ToF ultrasound traveling through the plate under applied load is less than that in the unloaded plate ($ToF_P < ToF_0$). The time interval between peak A" and peak B" (ToF_P) in figure 6.3 is shorter than the time interval between peak A and peak B (ToF_0) figure 6.3, although it is not very clear as the difference is in the order of nanoseconds. If only the first reflection is considered, peak A" advances to the left relative to peak A.

6.2 Results

6.2.1 Static load measurement

Initially, the ultrasonic ToF approach was applied to the model line contact when it was positioned right underneath the sensor. From the Hertz theory, the contact width of the model line contact is comparative small and the deflection only distributes in this contact zone. As such a case, it is important to make sure the line contact is aligned with the ultrasonic transducer.

The first reflection from the contact face was extracted from the whole waveform and recorded. A series of time trace plots of the reflected pulse from the back surface of the upper plate (the contact interface in loaded cases) under different loads are given in figure 6.5. The reference signal in the figure refers to the reflection from the free back face of the upper plate and the reflected pulse from the contact interface under different loads are marked by the load.

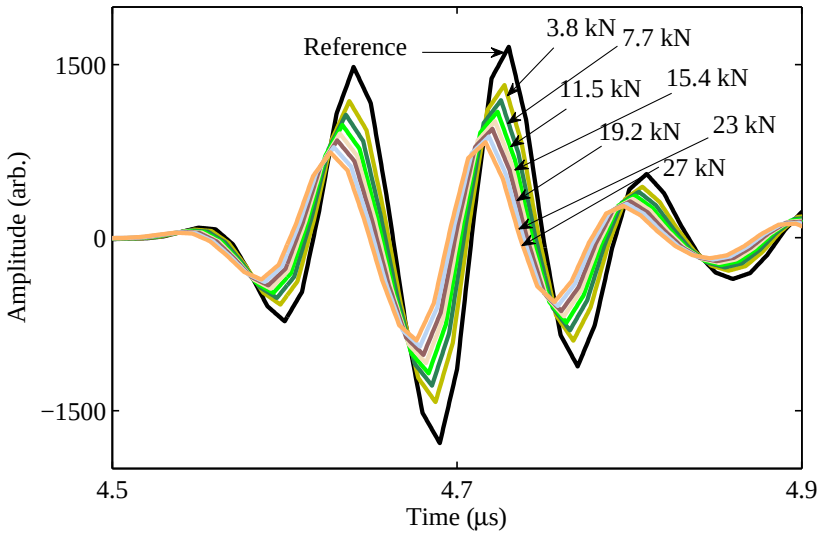
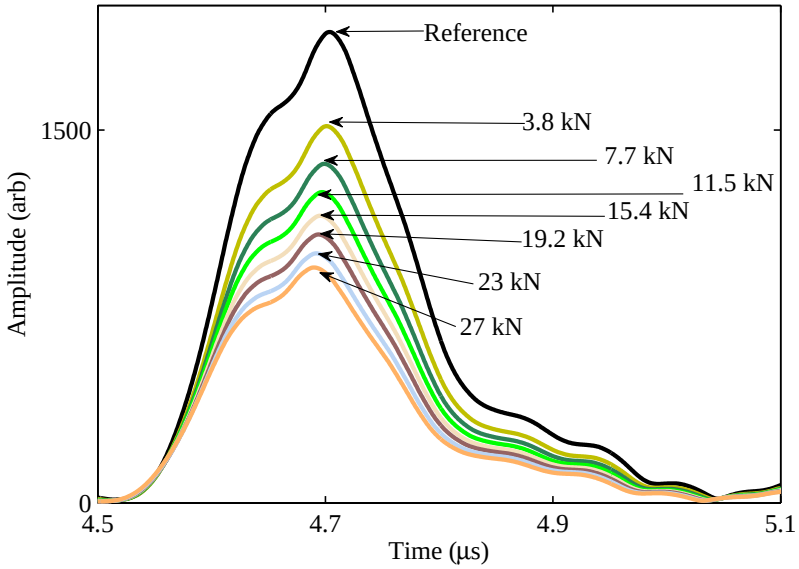


Figure 6.5. *The first reflection from the contact face of the plate under different conditions. The first reflection from the free surface is marked by reference and that from loaded contact interface are marked by the load.*

As stated in Chapter 4, the ultrasonic ToF is determined by the thickness and ultrasound speed in the plate, both of which depend on the load state and thus the ToF. In ultrasonic measurements, the contact interface causes the phase change in the reflected signals which appears as an apparent change in the ToF in simple peak to peak time determination ap-

proaches. The Hilbert transform has been used to obtain the envelope of signals to exclude this phase change effect.

Envelopes of the signals in figure 6.5 were calculated, they are given in figure 6.6. The reduction of the reflected wave energy (the amplitude of the envelope) with increasing load is clear in the figure. The change in the arrival time of the envelope of the first reflection can be obtained by comparing the envelope of loaded signal with that of the reference, as described in Section 5.3.1. This time shift is then caused by the geometric deflection and the acoustoelastic effect.



`Figure 6.6.` *Envelopes of the first reflected pulse from the back face of the plate under different conditions.*

According to each contact load, the ToF change caused by the plate deflection can be predicted from equation (4.8) in Section 4.2.2 and the ToF change caused by the acoustoelastic effect can be predicted using equation (4.17) in Section 4.3.2. The total change in ToF was then related to contact load by equation (4.25) which was reproduced here for the reader's convenience.

$$\Delta t_t = \frac{2\delta\,(1-L_{zz})}{(c_{zz})_0}. \tag{6.1}$$

The deflection of the model line contact was calculated from equation (4.4), which was in turn used to predict the ToF change under different loads. By substituting equation (4.4) into equation (6.1), the relationship between the theoretical ToF change and the contact load

can be expressed as:

$$\Delta t_t = 3.84 \times 10^{-8} \frac{P_{max}^{0.9}}{(l_0 \times 1000)^{0.8}} \frac{(1 - L_{zz})}{(c_{zz})_0}, \tag{6.2}$$

where $L_{zz} = -2.24$, measured from experiment. Figure 6.7 shows the prediction of ToF change under different contact loads: the ToF change resulted from the deflection and the speed change increases with contact load.

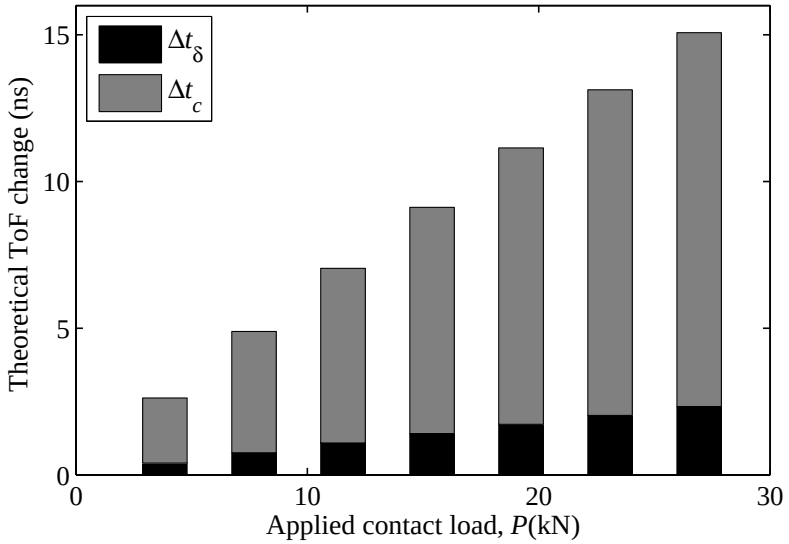


Figure 6.7. *Theoretical ToF change caused by the deflection and the acoustoelastic effect for seven different loads on the model line contact.*

By using equation (6.2), the measured ToF change can be converted into contact load. Figure 6.8 compares the load measured from the ToF change in ultrasonic reflections and the applied load on the line contact. It can be observed that reasonable good agreement has been achieved.

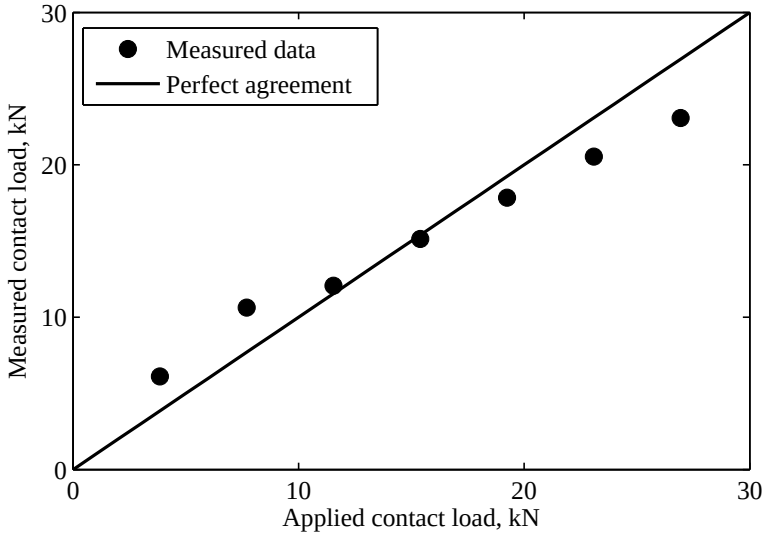


Figure 6.8. *Comparison between the load measured from the ultrasonic ToF method and the applied contact load.*

6.2.2 Change in the ultrasonic waveform as the contact moves

Load measurements were then carried out on the model line contact when it was indexed using a stepper motor as shown in figure 6.2. As the central plate was driven by the stepper motor, the line contact between the roller and the plate approached the sensor, moved underneath it and then receded. The position of the central plate can be indexed by the stepper motor. For each load condition, ultrasonic reflections from the back face of the upper plate when the line contact was at different positions were recorded. Again, the first reflection from the contact face was extracted from the whole waveform.

As an example, figure 6.9 shows the first reflection from the model line contact when (a) the line contact was away from the sensor, (b) the line contact was right underneath the sensor, and (c) the line contact passed through the sensor location. The amplitude of the reflected wave was considerable when the sensor away from the sensor, it reduced greatly when the sensor contact was under the sensor, and it increased back when the sensor moved away from the sensor. Therefore, the reduction in the amplitude can be used to index the passage of the line contact underneath the sensor location.

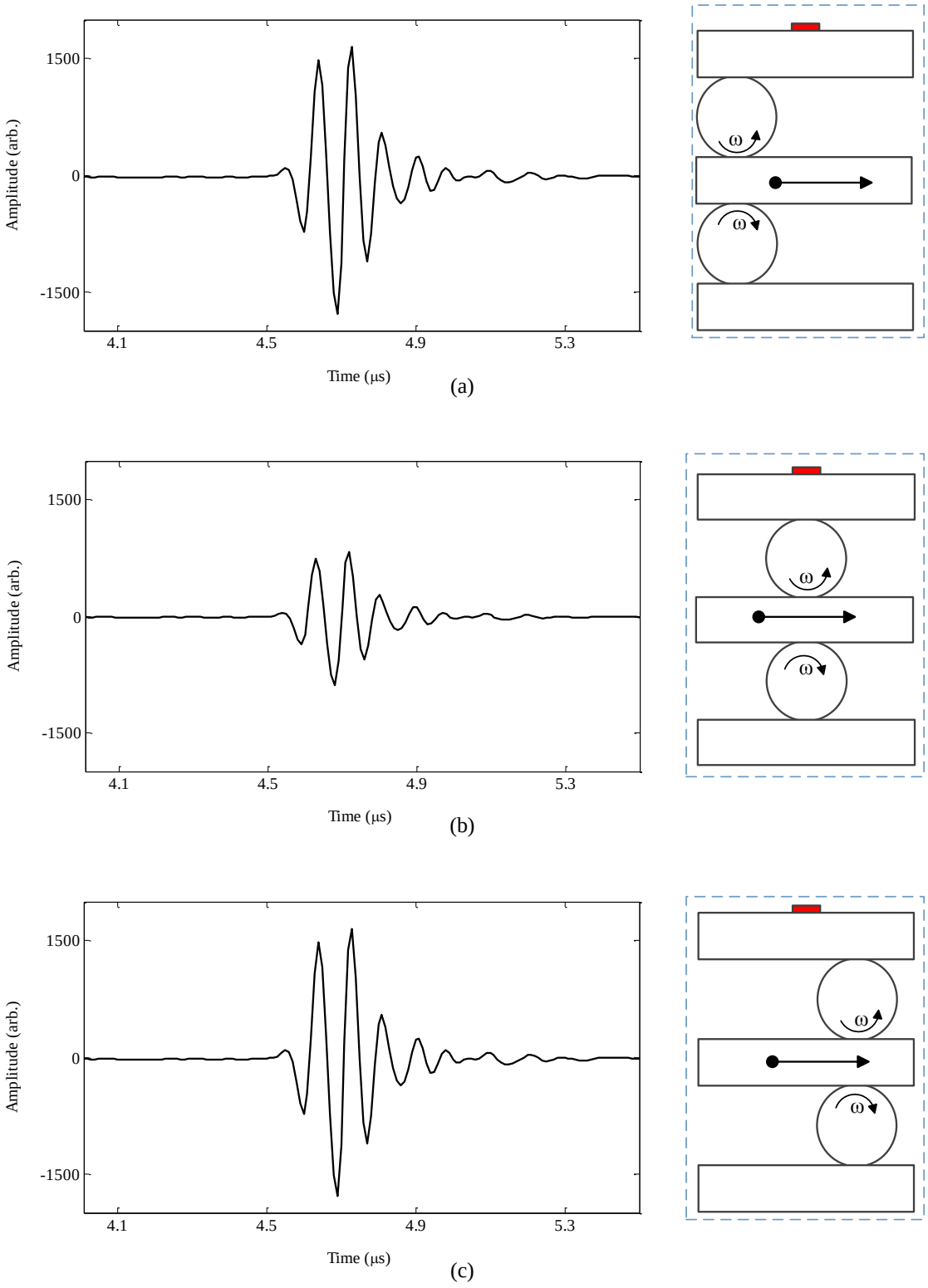


Figure 6.9. *The first reflection from a 10 MHz piezoelectric sensor on the upper plate when the line contact was at different positions under load*

In figure 6.9, when the line contact was directly underneath the sensor, the amplitude change in the signal was apparent while the change in ToF was not. For the model line contact, the ToF change was relative small, approximately 10 ns; the time shift would be more clear if signals in case (a) and (b) were plotted together, as shown in figure 6.10.

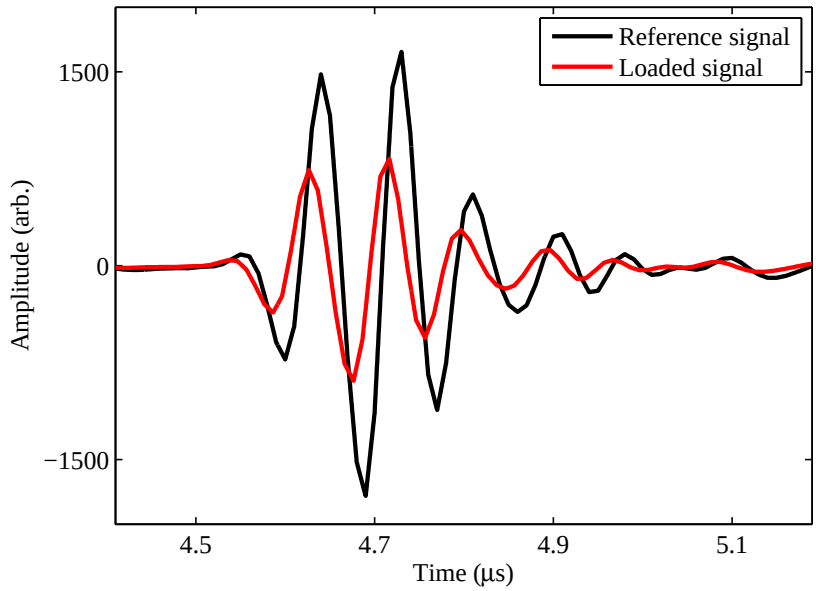


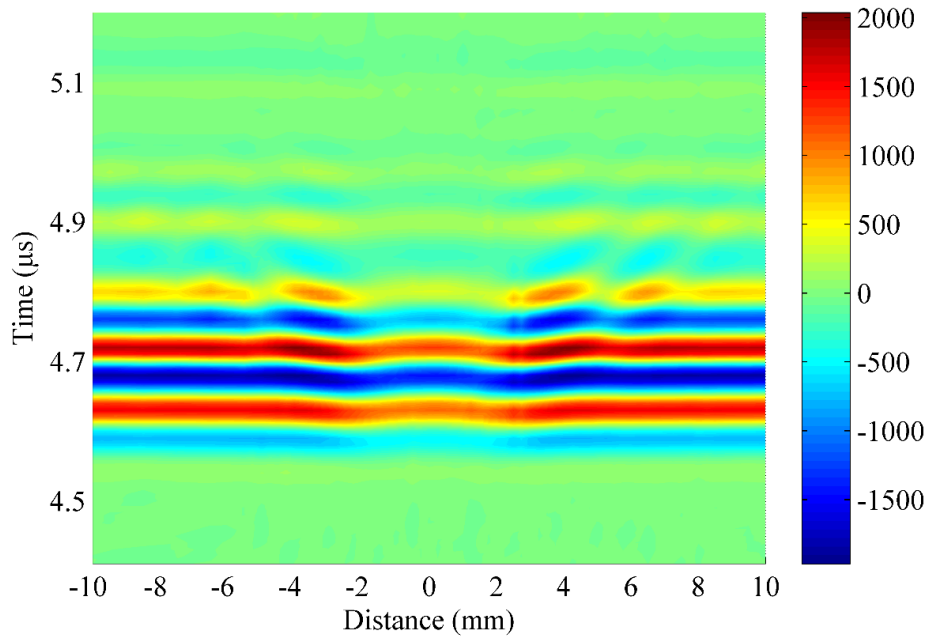
Figure 6.10. *The first reflection from the contact face of the upper plate in unloaded and loaded state. Reflection from the unloaded plate was taken as reference.*

6.2.3 B-scan data of first reflected pulses

The contact patch of the model line contact depends on the contact load, which can be predicted from the classical Hertz theory. For normal load from 2.9 to 20.7 kN, the contact width increases from 0.27 to 0.73 mm. The line contact can be driven by the stepper motor with the highest spatial resolution of 0.025 mm per step. However, the higher spatial resolution the more time consuming for the test.

In the experiment, to compromise the test time and the spatial resolution, the stepper motor moved at flexible intervals for different stages. At the beginning, the line contact was far away from the sensor location and the stepper motor moved forward to the sensor at an interval of 1mm per step. When the line contact was close to the sensor location, the stepper motor was changed to move at an interval of 0.25 mm per step. In this case, at least one reflection from the contact zone could be measured by the sensor. Once the line contact passed the sensor location, the motor interval was changed back to the interval of 1 mm per step.

For each load case, the ultrasonic reflection from the contact face when the line contact located at 70 positions were recorded and they can be shown as a B-scan map. B-scan is a two dimensional imaging system which presents the ultrasonic waveform with the location information where it is recorded. By rotating the first reflections, as shown in figure 6.9, 90 degrees in anti-clockwise direction and arranging them side by side along the distance, a B-scan map can be obtained. Figure 6.11 gives a B-scan map when the model line contact was under the contact load of 20.7 kN. In the figure, the x-axis is the distance between the moving roller and the fixed sensor ($x = 0$), the y axis is the time trace and the amplitude of the reflected pulse is represented by colors.



`Figure 6.11.` *B-scan data of the first reflection from the contact face when the line contact was at different positions under a normal load of 20.7 kN.*

In figure 6.11, the line contact passage is clear as the amplitude reduced remarkably when the distance was from -1.5 mm to 1.5 mm. Also, some ripples are observed when the line contact moved close to the sensor.

6.2.4 Spectrogram of first reflected pulses

A spectrogram displays the signal in three dimensions, by the time in the $x-$axis and by the frequency in the $y-$axis with a third dimension amplitude being presented by different colors in 2D view. The signal frequency information can be achieved by using the fast Fourier transform, and the distance is used here as it is of interest instead of the time. Figure

6.12 gives a spectrogram of the reflected pulses in figure 6.11. The amplitude of all signal was normalized with respect to the amplitude of signal when the line contact was away from the sensor. According to the definition of the reflection coefficient, the normalized amplitude in figure 6.12 is the amplitude of the reflection coefficient.

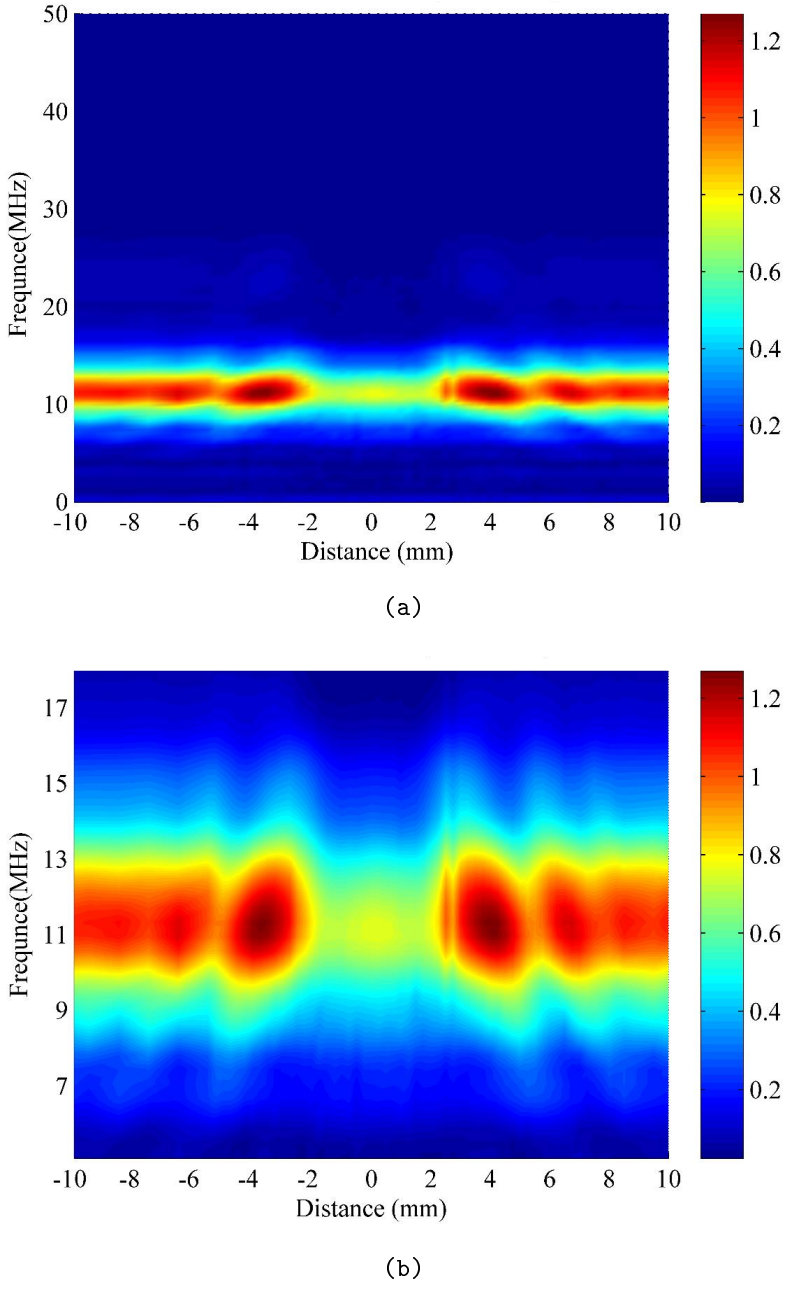


Figure 6.12. *Spectrogram of the B-scan data in figure 6.11 (a) frequency up to 50 MHz and (b) frequency from 5 to 18 MHz. The amplitude in the spectrogram is normalized and equal to the amplitude of reflection coefficient.*

The bandwidth of a transducer is commonly defined in its response curve in the frequency trace, it is the range where the main wave energy are distributed (Whittingham (1999)). The half maximum bandwidth where the ultrasonic energy drops to 50 percent of the maximum (-6 dB) is mostly used in ultrasonics and hence in this project. The central frequency of the ultrasonic transducer was measured at 11.3 MHz, although its nominal center frequency is 10 MHz, and the lower and upper frequency of the transducer bandwidth were 9.1 MHz and 13.5 MHz, respectively. This is clearly shown in the spectrogram figure, where the signal given for frequency from 0 to 50 MHz but ultrasonic energy is mainly distributed in the frequency from 9.1 to 13.5 MHz.

A magnified spectrogram for details of ultrasound in the sensor bandwidth is given in 6.12b. In the figure, the passage of the line contact is clearly observed by a reduction in the amplitude of the reflection coefficient. It was around unity before the line contact approaches, reduced to 0.67 when the line contact was in the sensor active measurement zone, and then increased back to unity when the line contact moved out of the ultrasound sight.

6.2.5 Reflection coefficient amplitude

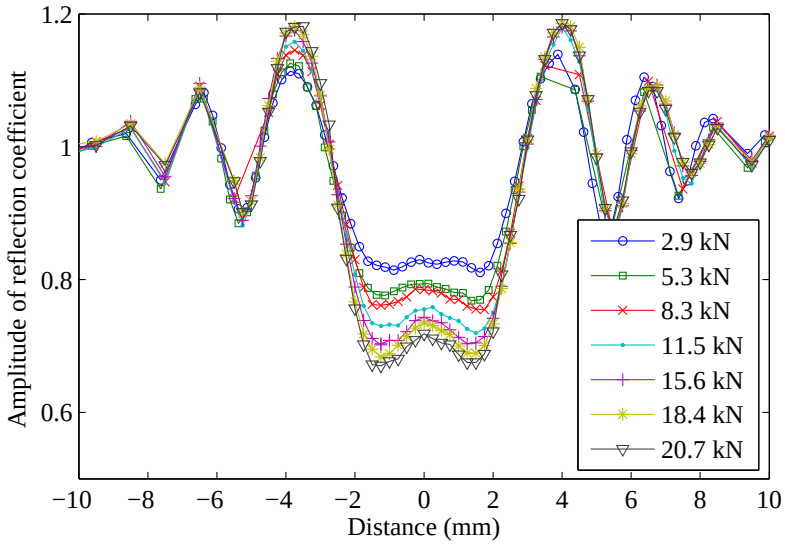


Figure 6.13. *The amplitude of the reflection coefficient at different positions when the model line contact was under seven different loads.*

The reflection coefficient amplitude at a specific frequency can be achieved by dividing the amplitude of the signal from different positions with the amplitude of the reference signal at

that frequency. Figure 6.13 shows the variation in the amplitude of the reflection coefficient at the sensor central frequency due to the model line contact pass.

It is noted that the reflection coefficient amplitude has some fringes when the roller entered into or exited the sensor active measurement zone. This is interesting and corresponding to the ripples in the B-scan data in figure 6.11. Also, it is observed that the amplitude of the reflection coefficient reduced to a minimum level when the line contact is rolled underneath the sensor, for the distance from -1.5 mm to 1.5 mm, for all load cases. The minimum reflection coefficient reduced from 0.8 to 0.67 as the contact load increased from 2.9 kN to 20.7 kN.

The reflection coefficient curves in figure 6.13 can be divided into five regions with respect to the line contact position, as shown in figure 6.14. Features of the reflection coefficient amplitude in these regions are described as follows.

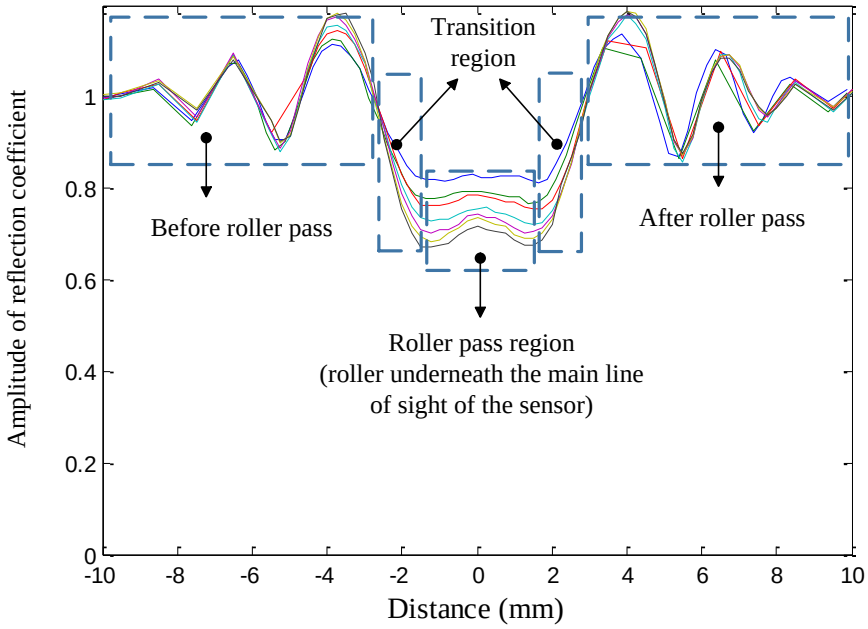


Figure 6.14. *The reflection coefficient curves can be described by five regions: 1) before roller pass region; 2) first transition region; 3) roller pass region; 4) second transition region; 5) after roller pass region.*

Firstly, oscillations in the reflection coefficient amplitude around unity were observed before the roller passed the sensor (distance from -10 mm to -3.25 mm) and after the roller passed through the sensor (distance from 3.25 mm to 10 mm). The reflection coefficient amplitude before and after the roller pass under all load cases were represented in figure 6.15. Fringes have been seen for all load cases, which implies the phenomenon is load

independent. It can be observed that the corresponding local peaks in the amplitude of reflection coefficient are quite close for different loads.

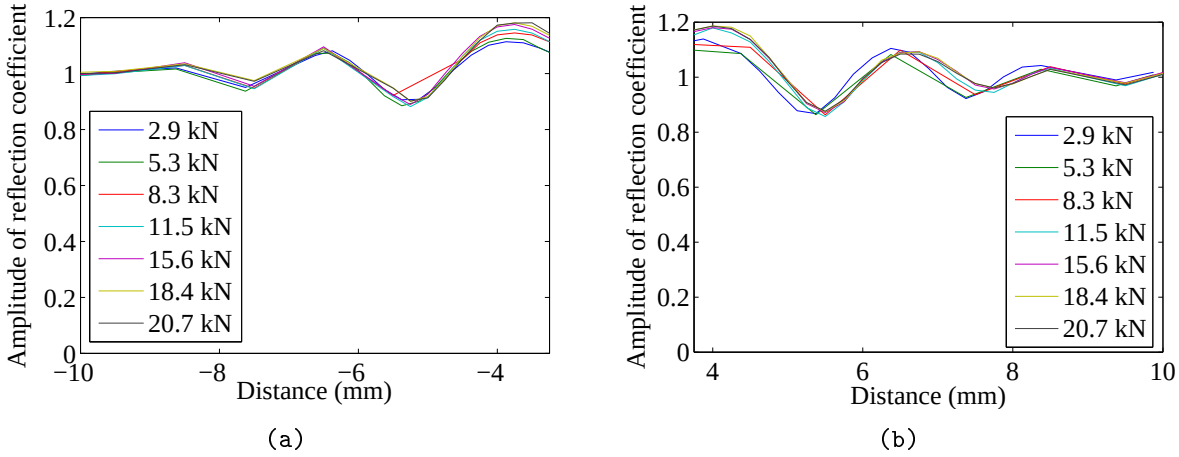


Figure 6.15. *The reflection coefficient amplitude fluctuates around unity (a) before the roller pass and (b) after the roller pass under seven different loads.*

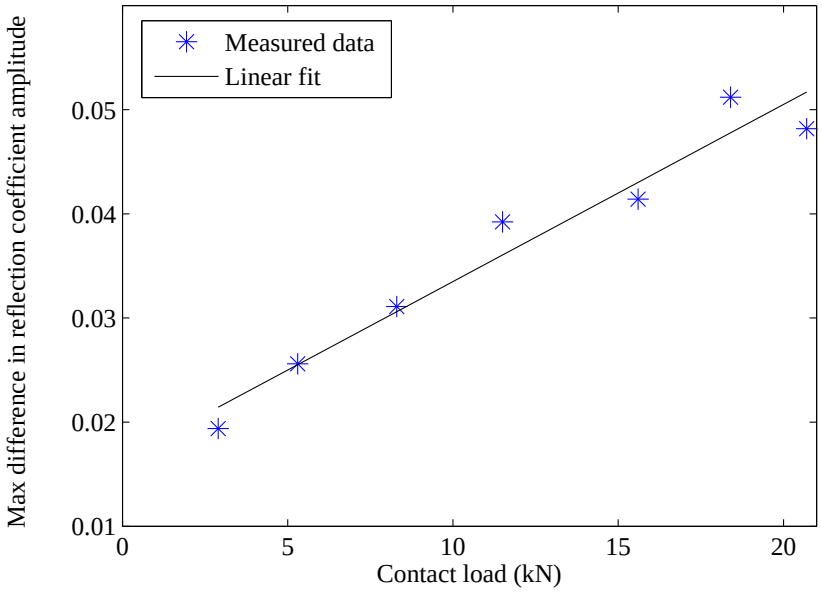
It is interesting that the reflection coefficient amplitude exceeds unity at some positions. The reference signal was taken from the steel-air interface where 99.999 percent of the incident wave was reflected. The reflection coefficient amplitude larger than one means the sensor received more energy than it emitted, which is probably caused by the wave interference.

Secondly, the reflection coefficient amplitude reduced from unity to a minimum according to the contact load in the first transition region, oscillated around the minimum value in the roller pass region and then increased back to unity in the second transition region. When the line contact moved close to the transducer effective sensing area, the reflection coefficient amplitude reduced gradually to its minimum as part of ultrasound is transmitted into the roller, which was marked as the first transition region in the figure. In the roller pass region, the amplitude of the reflection coefficient increased slightly when the line contact approached to the sensor center and then fell back to the local minimum. As the line contact exited the sensor's effective measurement zone, the reflection coefficient then gradually increased back to unity, which was marked as the second transition region in the figure.

The amplitude of the reflection coefficient in the two transition regions and the roller pass region, i.e. distance from -3.25 mm to 3.25 mm, was in a distorted W-shape. By contrast, from Hertz theory the contact patch for the model line contact is in a semi-cylindrical form. At the center of the roller pass region, the reflection coefficient tended to increase, instead of reduce as it should be as a result of the presence of the maximum contact pres-

sure at the centre. The reason why this is the case probably be the transducer average effect which will be discussed later.

Thirdly, the discrepancy between the reflection coefficient amplitude at distance of ± 1.5 mm and that at the center increased with load, resulting in a more clear W-shape profile for higher load cases. The difference between the central peak and the minimum trough for each load case was given in figure 6.16. A general linear trend was observed, clearly this difference is load dependent.



`Figure 6.16.` *Difference between the central peak and the minimum trough in roller pass region on reflection coefficient amplitude curves with increasing contact load.*

Moreover, the distance between the two minimum points in the distorted W-shaped profile, the width of the roller pass region in figure 6.14, was 3 mm for all load cases. It implies that the roller pass region measured by the ultrasonic sensor is load independent. Whilst the contact width from Hertz theory increases from 0.27 mm to 0.73 mm with load from 2.9 kN to 20.7 kN. The roller pass region in the figure is more likely the effective measure zone of the sensor rather than the actual contact zone. The ultrasonic wave from the modified piezoelectric sensor was unfocused and it spread out with the wave path in the attached component, resulting in the line of sight of the sensor at the contact face was larger than the sensor width. The real line contact width was less than 1 mm, when the contact moved in the sensor's effective measure zone, the reflection coefficient was almost kept at its minimum. This phenomenon will be discussed later in this chapter.

6.2.6 Change in ToF

The ToF change in the plate can be obtained by comparing the envelope of the loaded signal with the envelope of the reference. The Hilbert transform was used to calculate the envelope of ultrasonic signals to eliminate the phase shift effect caused by the contact stiffness. For each load, envelopes of signal for different positions were compared with the envelope of the reference and the ToF change were obtained. Figure 6.17 gives the measured variation in the ToF change with distance and load.

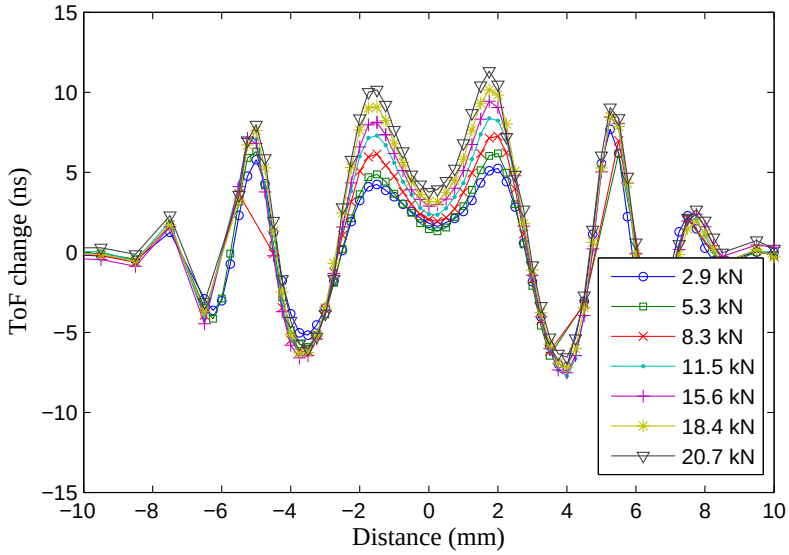


Figure 6.17. *Measured ToF change at different positions for the model line contact under seven different loads*

A few features similar to those observed on the reflection coefficient amplitude curves were observed:

- The ToF change fluctuated around zero before the roller pass and after the roller pass. This fluctuation was observed for all load cases, suggesting it is load independent.
- The ToF change in the transition region and the roller pass region, for distance from -3.25 mm to 3.25 mm, followed a M-form variation. The ToF change increased from zero to a maximum in the first transition region, reduced to a local minimum as the line contact approached the sensor center and increased back to the maximum level in the roller pass region, and then dropped back to zero in the second transition region.
- The discrepancy between the ToF change at ± 1.5 mm and the ToF change at the sensor center increased with load, resulting in a more and more clear M-shape profile

for higher load cases. The difference between the local peak and the central trough for all load cases were shown in figure 6.18. Again, a general linear trend can be observed, suggesting this difference is load dependent.

- The distance between the two maximum points in the distorted M-shaped profile, the width of roller pass region in figure 6.17, was 3 mm for all load cases, implying it is load independent.

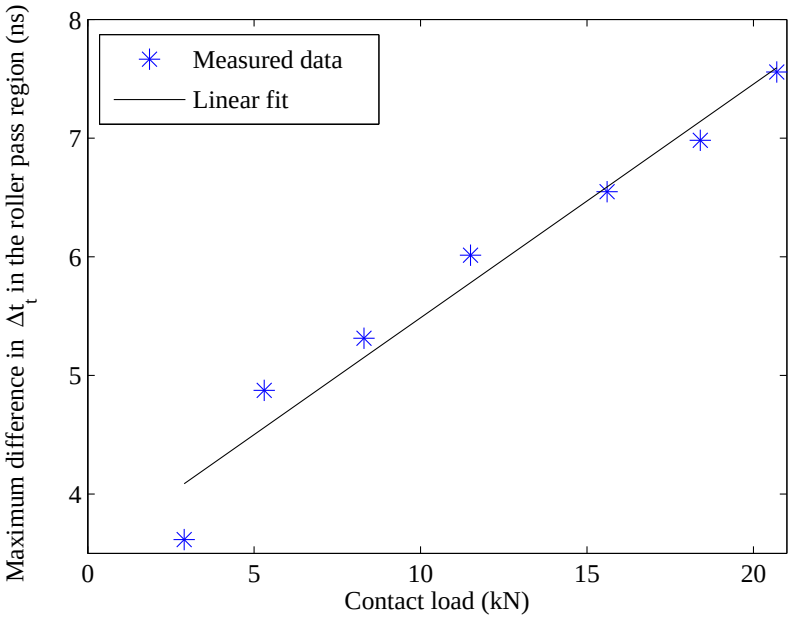


Figure 6.18. *Difference between the central peak and minimum trough in roller pass region on ToF change curves with increasing contact load.*

In the roller pass region, the reflection coefficient amplitude fell to a minimum level while the ToF change achieved a maximum. The ToF changes measured by ultrasonic reflection under seven different contact loads are given in figure 6.19. It can be seen that the measured ToF change increases with load as expected, although the experimental values are slightly larger than predictions for the low loads.

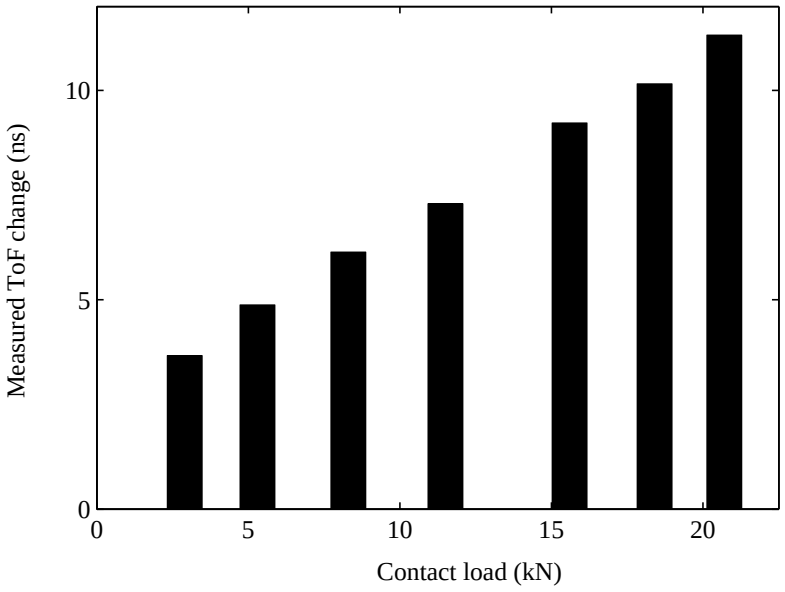


Figure 6.19. *Measured ToF change from ultrasonic reflection against contact load*

Using the relationship between the predicted change in the ToF and load, see equation (6.2) , the measured ToF change can be converted as load. Figure 6.20 compares the applied contact load and the load measured by the ultrasonic ToF method. General good agreement is seen between the measured and applied contact load. For the case of low loads, the measured load slightly over predicted the load. This is probably caused by the actual contact length was less than the effective length of the roller. The curvature of the roller resulted in the contact between the plate and the roller is not a perfect line contact, which gives higher contact stress than the predicted value from the Hertz theory. For the case of high loads, the measured load was slightly less than the applied load. This is because the measured ToF change from ultrasonic reflections is an average over the contact patch which is less than the ToF change at the contact center. The transducer average effect will be discussed in Chapter 7.

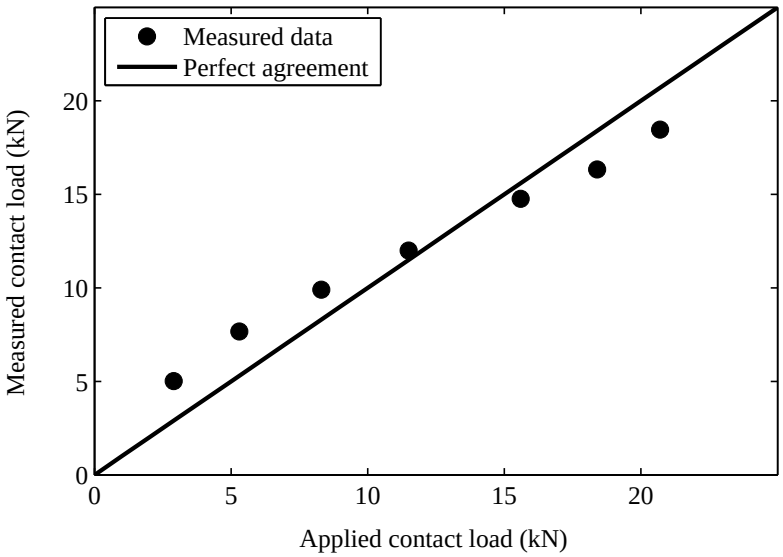


Figure 6.20. *Measured contact load from ultrasonic reflection against real contact load*

6.3 Discussion

The fringes before and after the roller pass regions and the width of the roller pass region are discussed in this section. Similar to the reflection coefficient curves, the change in the ToF according to the line contact position can be described by five regions as shown in figure 6.21.

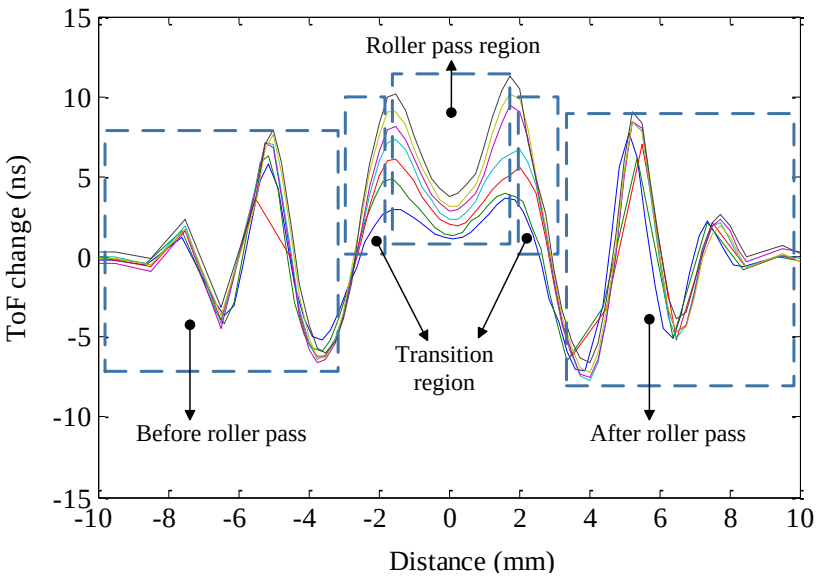


Figure 6.21. *Measured ToF change when the line contact was at different positions under seven different load.*

6.3.1 Fringes before and after roller pass

A few fringes were observed in the ultrasonic ToF before the line contact pass the sensor location, see figure 6.22a. It is noted that this phenomenon occurred for all load cases, which suggests the fringe effect does not depend on load. The magnitude of the local peaks and troughs on the ToF change are similar for different loads. It is also noted that the distance between two consecutive fringes increases when the roller approaches the sensor. To be specific, the distance between the first two troughs is 2 mm whilst the distance between the second and the third trough rises to 3 mm.

The fringe effect was also noted after the roller passed the sensor location for each load measurement, see figure 6.22b. The distance between each two consecutive fringes decreases as the roller moved away from the sensor location.

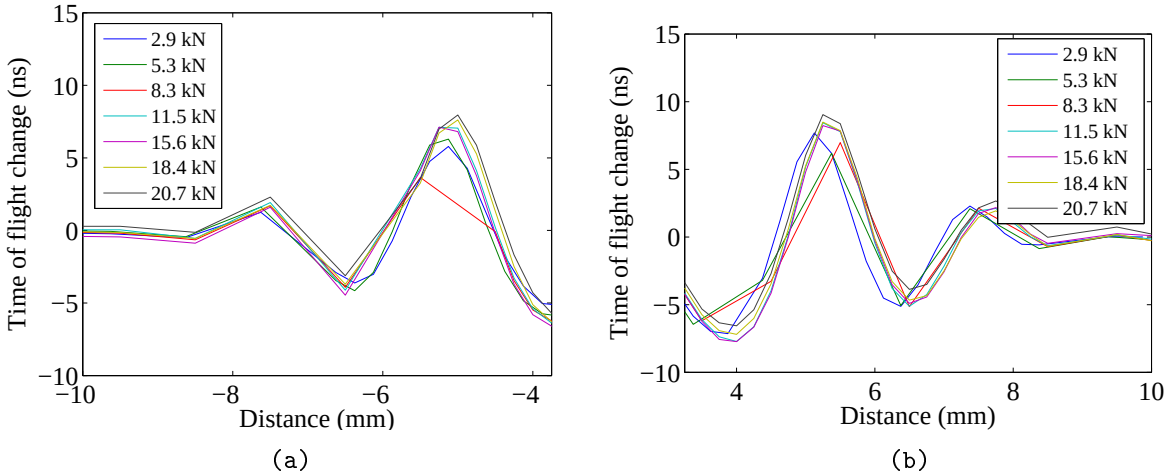


Figure 6.22. *ToF change (a) before the roller pass and (b) after the roll pass.*

6.3.1.1 Examination of subsequent reflections

To explore this fringe effect, further investigations were taken on the subsequent reflections from the contact face, the second and third reflections. The reference signal and the loaded signal are given in the time trace, shown in figure 6.23, where the waveforms consist of the first three reflections from the contact face. B-scan plots of the second and third reflection are given in figure 6.24, with the contact load of 20.7 kN.

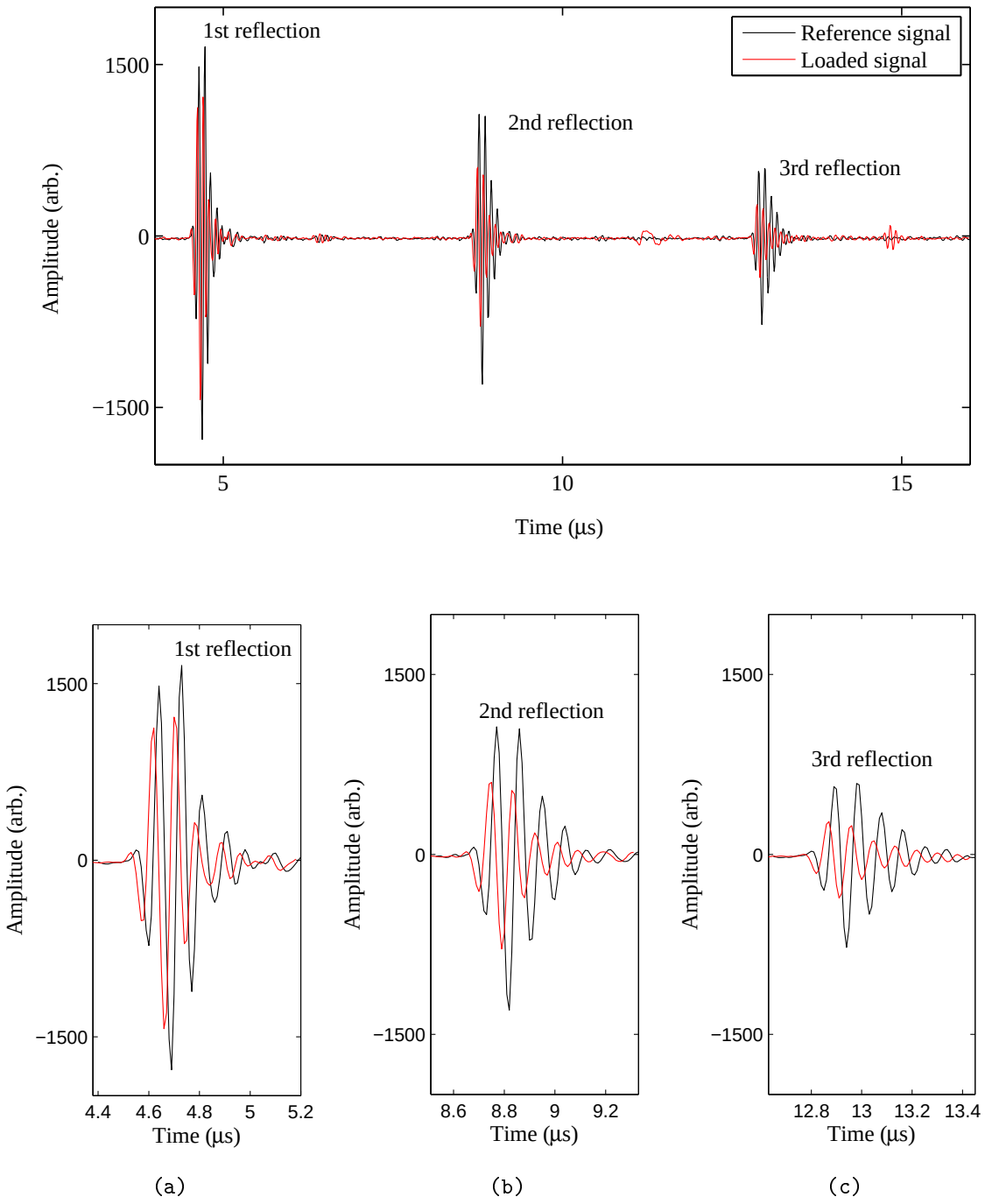


Figure 6.23. *The first three reflections from the contact face under load of 20.7 kN.*

For the second reflection, ultrasound travels four times of the upper plate thickness and the reflection happens twice. Therefore, the amplitude of the second reflected signal is reduced to the square of the reflection coefficient multiples by the amplitude of the incident

signal. The ToF difference for the second reflections in the unloaded and the loaded plate is then twice that for the first reflections. However, it is difficult to see this phenomenon in figure 6.23, since the ToF change is of the order of nanosecond and it is masked by the apparent time shift from the phase change effect.

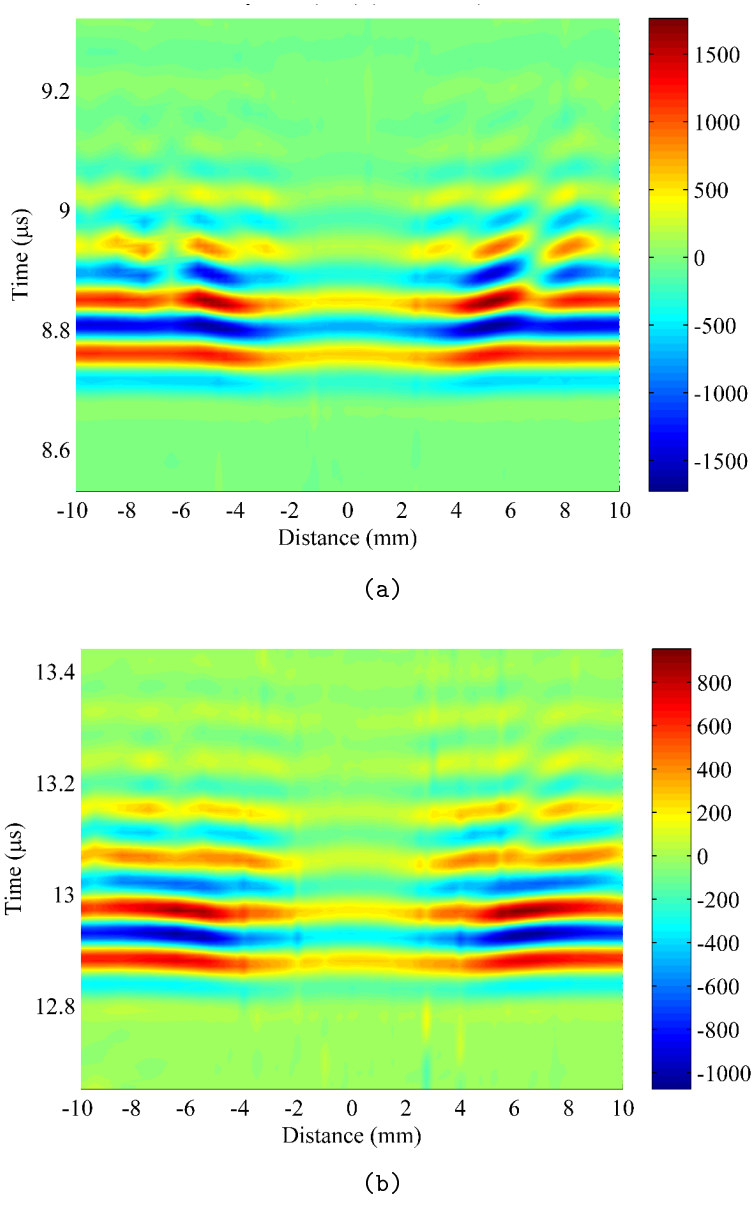
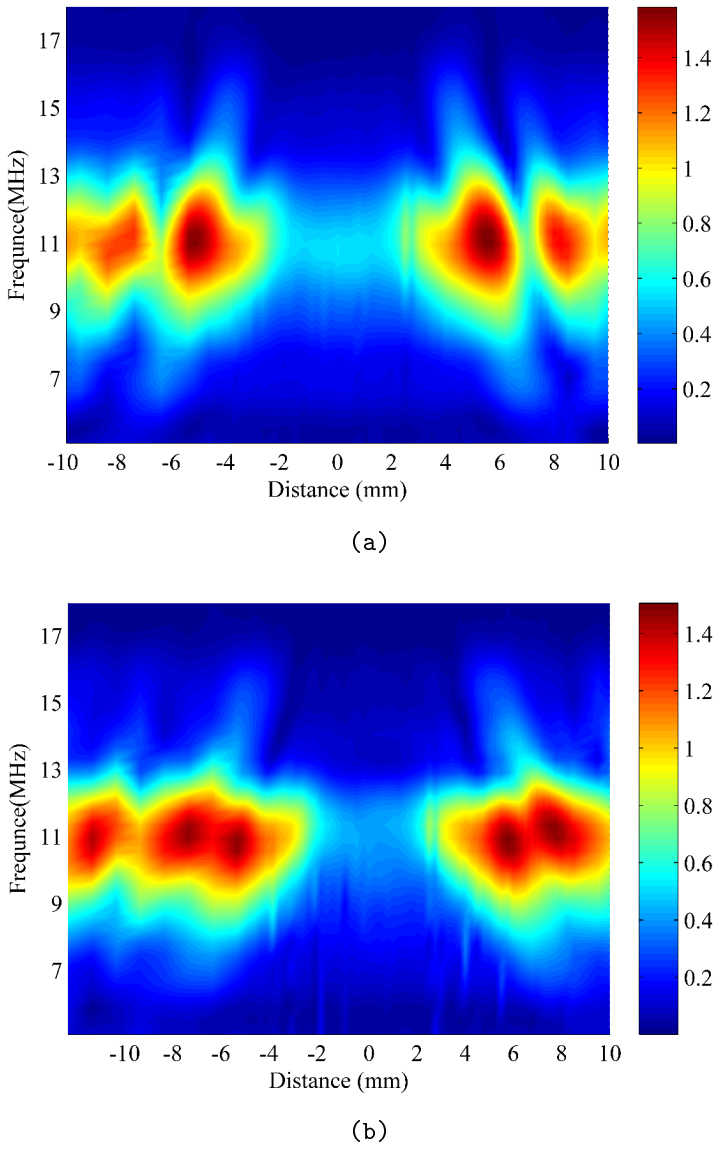


Figure 6.24. *B-scan data of (a) the second reflection and (b) the third reflection from the contact face under load of 20.7 kN.*

The spectrogram for the second and the third reflection were obtained by applying the fast Fourier transform to the B-scan data in figure 6.24. Again, the amplitude of the signal is normalized with respect to that of the reference signal. Therefore, the amplitude in the

spectrogram figure for the second and the third reflection are the square and the cube of the reflection coefficient amplitude, respectively. Figure 6.25 gives the spectrogram at frequencies from 5 MHz to 18 MHz (a) for the second reflection and (b) for the third reflection.



`Figure 6.25.` *Spectrogram map at frequencies from 5 to18 MHz for (a) the second reflection and (b) the third reflection.*

From the figure 6.12 and 6.25, it is noted that the fringes were first seen in the reflection coefficient curves for the higher order reflection in terms of the distance between the line contact and the sensor. This is highlighted in figure 6.26 which extracts the power of the reflection at the central frequency under the same load. The first fringe on the reflection coefficient curve from the third reflection occurs at - 9.5 mm, while that on the reflection

coefficient curve from the first reflection happens at -7.5mm. In the figure, it seems that the third reflection has less fringes than the first and second reflection. In addition, as the order of the reflection increases, the distance between the two peaks in the reflection coefficient curve increases. The distance between the two peaks is 7.75 mm for the first reflection, it enlarges to 10.75 mm and 11.25 mm for the second and the third reflection, respectively.

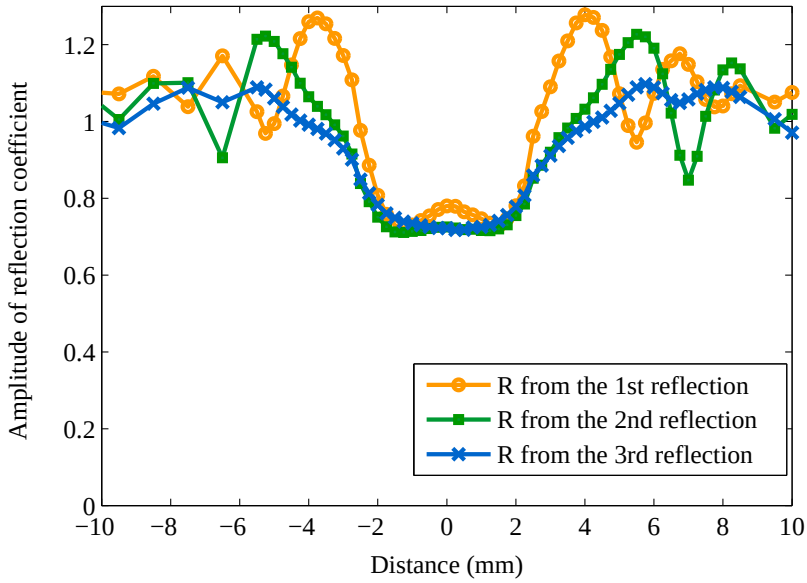


Figure 6.26. *The reflection coefficient from the first three reflections vary as the model line contact pass the sensor. The contact load is 20.7 kN.*

6.3.1.2 Mechanism of the fringe effect

The fringes were observed when the roller was close to the ultrasonic transducer. This effect may be caused by the interference of the wave as the sensor works in the near field. The length of the near field for a circular sensorN_{so} can be predicted from the diameter of the sensor D_s and the wavelength λ, as (Kräutkramer and Kräutkramer (1977))

$$N_{so} = \frac{D_s^2}{4\lambda}. \tag{6.3}$$

The length of the near field is 21.36 mm for a piezoelectric element having a diameter of 7.1 mm, according to equation (6.3) . For a rectangle sensor, the length of the near filed is determined by its length (Ginzel et al. (2014)). The near field length of an rectangular element can be calculated from equation (6.3), where L_s is the length of the rectangle and $k_a$is a correction factor which is determined by the dimension ratio (width to length) (European

Standard (2001)).

$$N_{sr} = \frac{k_a L_s^2}{4\lambda}. \tag{6.4}$$

The factor $k_a$is dominated by the length of the rectangle, and changes from 0.98 to 1.36 when the dimension ratio increases from zero to unity, Therefore, the length of the near field for the modified rectangle sensor, is slightly smaller than 21.36 mm. In the mode line contact, the thickness of the plate is 12.3 mm. Therefore, the sensor works in its near field, where the wave interference occurs.

The wave interference results in a series of maxima and minima in the sound field. Therefore, there are some side lobes in addition to the main lobe in the wave energy profile. When the contact meets with these side lobes, fringes are observed.

6.3.2 Measured roller pass region

In the experiment, the sensor was fixed on the upper plate and the line contact moved underneath the sensor. When the line contact entered into the active sensing area, i.e., the line of sight of ultrasound, the reflection coefficient amplitude reduced as part of ultrasound was transmitted into the roller. The reflection coefficient technique can be used to measure the contact patch if the line of sight of the ultrasonic sensor is comparably small to the contact dimension. This method has been applied to measure the wheel-rail contact (Marshall et al. (2006)).

However, in the ultrasonic tests, the active sensing area of the used ultrasonic sensor is too large compared to the contact area. From figure 6.21, the width of contact measured from the ultrasonic reflection is about 2.75 mm. The contact width under each load can be predicted from the Hertz theory, they are compared with the contact width from the ultrasonic measurements in table 6.2. The ultrasonic wave from the modified piezoelectric element was unfocused and it spread out along the wave path, resulting the line of sight of ultrasound was larger than the width of the modified piezoelectric element. According to the theory of diffraction, the angle of divergence for the rectangle sensor, γ_0, can be calculated from the wavelength and the sensor width, as (Kräutkramer and Kräutkramer (1977))

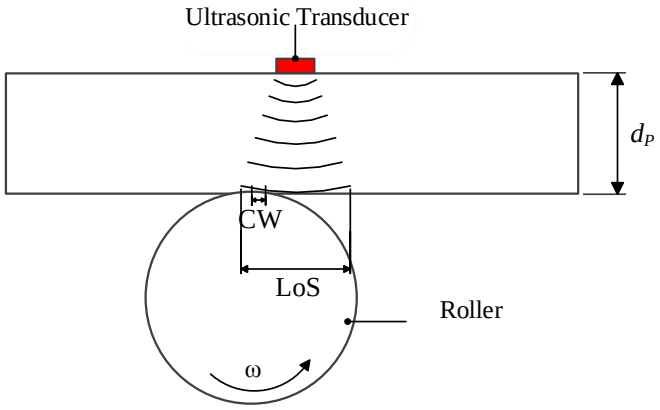
$$\gamma_0 = \arcsin \frac{\lambda}{W_s}, \tag{6.5}$$

where W_s is the width of the sensor. For the sensor used in the model line contact, the angle of divergence is 0.63 radians.

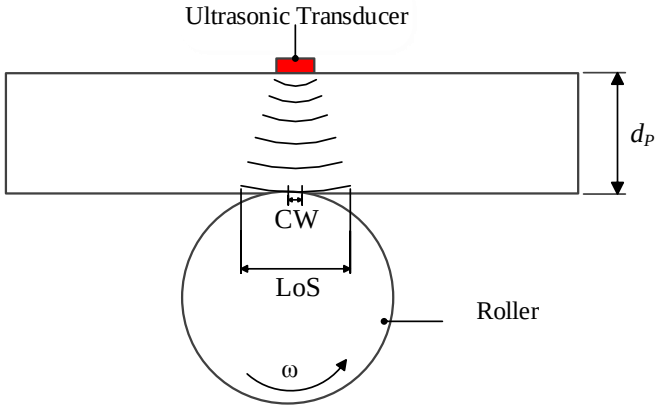
Table 6.2: The contact width from Hertz theory and ultrasound measurement

Load (kN)	Contact width from Hertz theory (mm)	Contact width from ultrasound(mm)
2.9	0.27	2.50
5.3	0.37	2.50
8.3	0.46	3.00
11.5	0.55	2.75
15.6	0.63	2.75
18.4	0.69	2.75
20.7	0.73	2.75

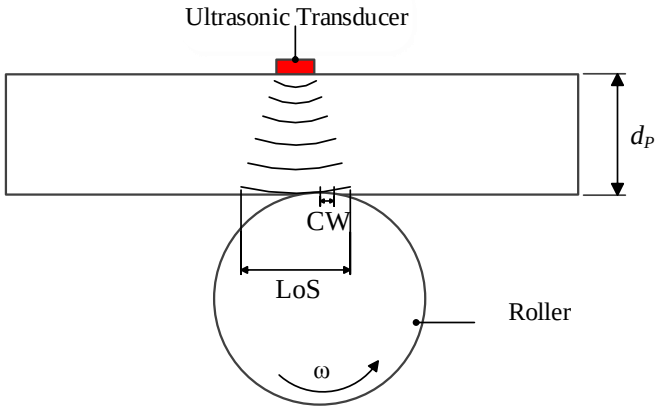
Figure 6.27 schematically shows the senor width, the line of sight of ultrasound and the contact patch. The step interval for the line contact was 0.25 mm, ensuring that the sensor could sense more than one ultrasonic reflection from the contact. This has been observed in both reflection coefficient curves in figure 6.14 and ToF change curves in figure 6.21. It also noticed that the distance between the two troughs maintained almost constant for all load in experiments, which suggests the measured roller pass region is more likely the line of sight of ultrasound instead of the real contact patch.



(a) contact at the entry of the line of sight



(b) contact in the middle of the line of sight



(c) contact at the exit of the line of sight

`Figure 6.27.` *Schematic of the sensor, the line of sight of ultrasound and the width of the contact. CW and LoS in the figure represnts the contact width and the line of sight, respectively.*

6.4 Conclusion

The ultrasonic ToF approach has been applied to the model line contact between a flat plate and a cylindrical roller. A modified narrow strip piezoelectric sensor has been instrumented on the plate. The load in the line contact has been measured via measuring the ToF change of ultrasound in the plate, since it is determined by the plate thickness and the speed of sound, both of which are load dependent. The phase change effect has been excluded by measuring the time difference between the envelope of the reference and the envelope of the loaded signal. The envelope of signal has been obtained from the reflected signal and its Hilbert transform. The measured ToF change, caused by the deflection and the acoustoelastic effect, increased from 3.67 to 11.32 ns as the load increased from 2.9 to 20.7 kN. The measured ToF change has been converted to contact load by using the relationship between load and the predicted ToF change. Good agreement has been achieved, showing that the ultrasonic ToF method is able to indicate the contact load.

Chapter 7

Load Measurements on Roller Bearing Contact

In this chapter, the method developed in Chapter 4 is used for load monitoring in the roller/raceway contact in a cylindrical roller bearing from a wind turbine gearbox. The first section is a brief introduction, followed by predictions of total ToF change under different bearing loads. The contribution of ToF change caused by the deflection, the acoustoelastic effect and the phase change effect have been studied. The next section describes the experimental apparatus including the sensor instrumentation on the roller bearing. This is followed by a section for experimental results where the measured ToF change has been used to indicate the contact load. Discussions of features observed in the ultrasonic reflections are given before the conclusion for this chapter.

7.1 Introduction

Although bearings in wind turbine gearboxes are basically designed to serve as long as twenty years, they have been reported to only survive about five years. These unexpected premature bearing failures, caused by the highly transient loading nature of wind, result in numerous wind turbine failures and make they are unreliable. It was reported that the bearing failure in the gearbox resulted in up to one-third of wind turbine failures (Christopher and Simon (2010)). The undesired downtime for exchanging these failed bearings greatly increases the cost of wind power. This makes it a less competitive source of renewable energy. The exact load information and the load history determine the bearing endurance life. Therefore, this chapter describes the experimental work of load monitoring that has been done on a cylindrical roller bearing used for the planetary gear of an epicyclical stage of a

2-megawatt wind turbine gearbox, by using a novel ultrasonic technique.

7.2 Prediction of the total ToF change

7.2.1 Load distribution in the roller bearing

When a radial load is applied to the cylindrical roller bearing, the roller at the bottom of the load line is most heavily loaded. Some of the other rollers in the bearing (located in the lower plane in figure 7.5) also share the load. In the experiment, the test bearing was operated at a low speed of 100 rpm. As the centrifugal force is not significant, the effect of the centrifugal force is ignored in the analysis of load distribution. The geometry and properties of the bearing used in the calculation has been given in table 4.2.

For the roller bearing with zero diametral clearance, the maximum load on the roller can be calculated from Harris (2001) by using the equation:

$$P_{max} = \frac{4.08W}{N}, \tag{7.1}$$

where W is the applied hydraulic load on the roller bearing, N is the roller number. The load for the roller located at ψ to the radial load line can be expressed as:

$$P_{\psi} = P_{max}\cos^{1.11}(\psi). \tag{7.2}$$

The load distribution in the roller bearing can be calculated from equation (7.1) and (7.2). As an example, the load on each roller is given in table 7.1 for the radial bearing load of 980kN.

Table 7.1: Load distribution in the cylindrical roller bearing under a radial load of 980 kN

ψ(degree)	$\cos\varphi$	P_ψ(kN)
0	1	266.56
±24	0.9135	241.09
±48	0.6691	170.7
±72	0.309	72.39
±96	-0.1045	0
±120	-0.5	0
±144	-0.809	0
±168	-0.9781	0

7.2.2 Surface deflection

Once the load on the contact is known, the dimension of the contact patch can be calculated from the Hertz theory, using equation (2.50). Figure 7.1 gives the width of the most heavily loaded contact, between the inner race and the roller, for different bearing loads (the left $y-$axis). The contact width enlarges to 1.85 mm as the bearing load increases to 1000 kN. The elastic deformation takes place in the contact area and the magnitude can be obtained by using the load-deflection relationship. According to equation (2.64), the variation of the contact deformation with bearing load can be illustrated by the right $y-$axis in figure 7.1. As a result of load increase, the contact deformation grows to 90.46 μm for the bearing load of 1000 kN.

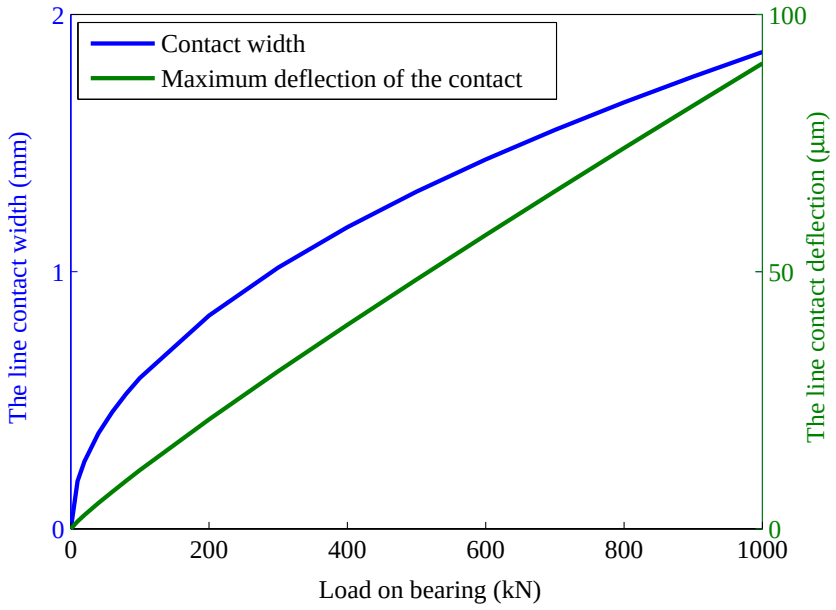


Figure 7.1. *The width and the maximum deformation for the most loaded inner race-roller contact for different bearing loads, the left and right y−axis, respectively.*

As described in Chapter 4, the ToF of ultrasound in the mating components changes with the load state, not only due to the variation in the wave path caused by the surface deflection but also the acoustoelastic effect. In addition to the ToF change caused by load, there is another component in the reflected pulse in practice, which appears as an apparent time change in simple peak to peak time determination. The following subsections predict the ToF change caused by the three effects separately.

7.2.3 ToF change caused by the deflection

When the raceway is loaded up by the rolling element, the thickness of the raceway is slightly changed as the occurrence of the elastic deformation. The wave path, twice of the raceway thickness, varies from $2d_0$ to $2d_P$. Thus the contribution to the ToF change by the geometry deformation alone in the stress state (that is $c_P = c_0$), can be given by

$$\Delta t_\delta = \frac{2\delta}{c_0}, \tag{7.3}$$

where δ is the elastic deformation in the raceway. The ultrasound speed in the unloaded race way was measured at 5900 m/s. Combining equation (2.64) and equation (7.3), the change in the ToF caused by the surface deflection alone can be calculated. The results for different bearing loads are shown in figure 4.12.

7.2.4 ToF change caused by the acoustoelastic effect

When the raceway is compressed, the material becomes more dense and the speed of sound increases, known as the acoustoelastic effect. The acoustoelastic effect of a material is quantitatively described by the acoustoelastic constant. In this study, a longitudinal wave from an ultrasonic transducer propagating through a raceway is strongly affected by the high normal stress field, as stated in Section 3.4. The acoustoelastic constant in the normal direction in a bearing steel was measured at -2.24, as described in Section 4.3.1.

The stress and/or strain in the normal direction is not constant but varies from a maximum at the contact surface reducing with depth through the component. The stress and/or strain distribution profile varies with the contact cases. Therefore, the ultrasonic speed in the normal direction is not constant but depends on the contact condition. In ultrasonic measurements, it is the speed through the component thickness that is measured. The accumulative effect of the varying stress and/or strain field must be used to determine the change in the ultrasonic wave speed. For the case of varying strain through the inner race thickness, the average strain can be expressed as:

$$\Delta\varepsilon_z = \frac{\delta}{d_0}. \tag{7.4}$$

The change in the speed of sound between loaded and unloaded cases can be calculated from deflection which is directly load dependent.

$$\Delta c_{zz} = L_{zz}\,(c_{zz})_0\,\frac{\delta}{d_0}. \tag{7.5}$$

The changes of the speed of sound have been calculated for the compressed inner race under several bearing loads, which are given in figure 7.2. As the bearing load increase to 1000 kN, the maximum contact load goes up to 272 kN, the maximum surface deflection becomes to 45.23 μm and the ultrasound speed in the compressed bearing inner race increase by up to 30.7 m/s.

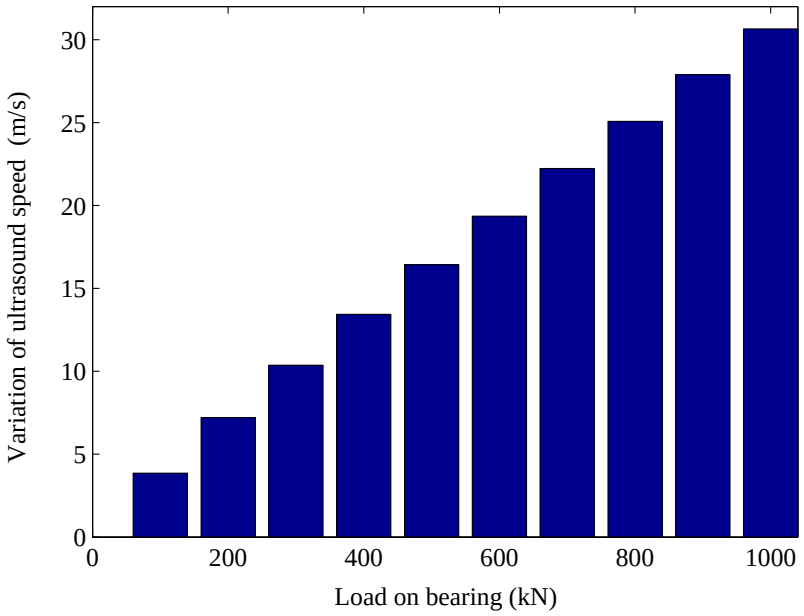


Figure 7.2. *Change in ultrasound speed under different bearing load.*

As a result of the acoustoelastic effect alone, the change in the ToF can be calculated according to equation (7.6). The corresponding ToF change has been given in figure 4.12.

$$\Delta t_c = \frac{-2L_{zz}\delta}{(c_{zz})_0}. \tag{7.6}$$

7.2.5 Apparent ToF change caused by the phase change effect

7.2.5.1 Contact stiffness

In practical ultrasonic reflection from a contact, the phase change was observed in the reflected pulse, the magnitude of which (and indeed the signal amplitude) could be predicted from given contact conditions (Reddyhoff et al. (2005)). However, the contact condition cannot be known as a priori as the bearing lubrication regime depends on applied load, bearing speed, material properties and the roughness of contact surfaces.

In rolling element bearings, for a contact separated by a thin oil layer having the thickness of h, the stiffness of the oil layer, K, is given by Dwyer-Joyce (2005), as equation (3.56). The equation has been represented here for the reader's convenience.

$$K = \frac{\rho c^2}{h}. \tag{7.7}$$

The minimum film thickness can be predicted from the well-known Dowson and Higginson equation (Dowson and Higginson (1968)) for line contacts as:

$$h_{min} = 2.65\bar{U}^{0.7}G^{0.54}\bar{Q}^{-0.13}R_x, \tag{7.8}$$

where $\bar{U}$, G and $\bar{Q}$ are referred to as the dimensionless speed, material and load parameters, respectively; they are expressed as:

$$\bar{U} = \frac{\eta_0 U\left(1-\nu^2\right)}{2ER_x} \tag{7.9}$$

$$\bar{Q} = \frac{Q\left(1-\nu^2\right)}{l_0 ER_x} \tag{7.10}$$

$$G = \frac{\lambda E}{\left(1-\nu^2\right)} \tag{7.11}$$

where $R_x$is the effective (or reduced) radius of the roller/race contact. The central oil film thickness can be predicted from the minimum film thickness, as

$$h_c = \frac{4}{3}h_{min}. \tag{7.12}$$

7.2.5.2 Phase change

The response of the reflected wave from the lubricant layer between the raceway and the rolling element has been described by the spring model. The reflection coefficient can be described by equation (3.49), which is a complex quantity having real and imaginary parts. In the complex plane, the reflection coefficient can be expressed as :

$$R = |R|cos\phi_R, \tag{7.13}$$

where $|R|$ is the amplitude of the reflection coefficient and ϕ_R is the phase angle. They are related to the wave frequency, the stiffness of the oil layer, and the acoustic property of the rolling element and the race by equation (3.51) and (3.52). The phase information of the reflection coefficient is the primary concern here as it describes the phase change between the reflected and the incident wave which appears as an apparent time change in simple time determinations.

For the roller bearing, the acoustic property of the raceway and the roller are the same,

thus equation (3.52) becomes to be :

$$\phi_R = \arctan \frac{K}{(\pi f z_1)}. \tag{7.14}$$

Therefore, the magnitude of the apparent time shift caused by the contact stiffness is directly related to the minimum thickness in the line contact by equation (7.15), where ρ and c are the oil density and sound of speed in the oil, respectively.

$$\Delta t_\phi = \frac{1}{2\pi f} \arctan \frac{\rho c^2}{\pi f z_1 h_c}. \tag{7.15}$$

7.2.6 Prediction of the total ToF change for a cylindrical roller bearing contact

The total ToF change for the line contact in the cylindrical roller bearing includes the ToF change caused by the surface deflection, the acoustoelastic effect and the phase change effect. Figure 4.12 (represented here) describes the variation of the three component contributions with increasing load. The bearing load increases from 100 to 1000 kN, giving the maximum load on the roller increases from 27.2 to 272 kN. The ToF caused by load via deflection and speed increases gradually from 6.25 to 49.68 ns, while the ToF from the phase effect maintained about 23 ns. This is due to the fact that the film thickness in the roller bearing is insensitive to load as shown by equation (7.8).

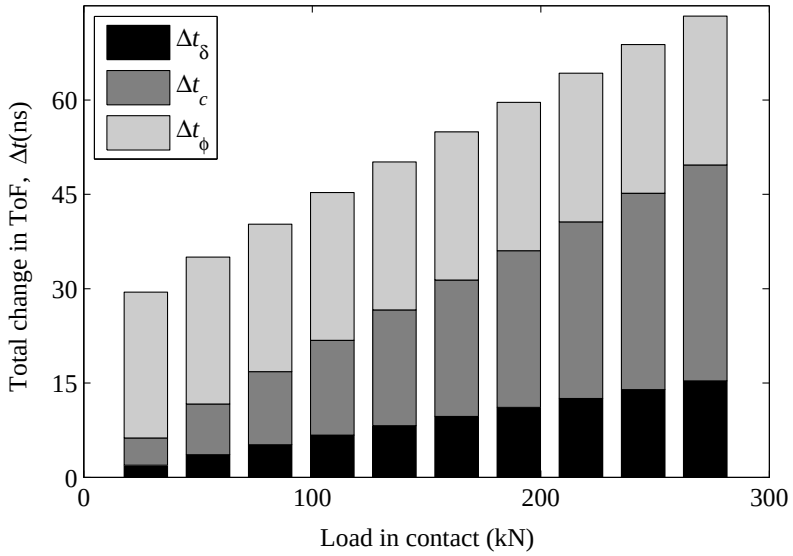


Figure 7.3. *Predicted change in ToF (ToF) for the roller bearing contact under a serials of load. For each load, the change in ToF includes contribution from surface deflection, acoustoelastic effect and phase shift.*

7.3 Experimental apparatus

A set of load measurements were carried out on the line contacts in a NU2244 cylindrical roller bearing (single row) which was used in the planetary gear of an epicyclical stage of a 2-megawatt wind turbine gearbox. The description of the test apparatus is given in this section.

7.3.1 Roller bearing test rig

Figure 7.4 shows a photograph of the experimental setup where the bearing rig was developed for tests in a pre-existing study (Wheals et al. (2011)). The rig was driven by a 7.5kW, 1450 rpm electric motor via a pulley-belt system. The test bearing was mounted within a large double row spherical roller bearing, enabling the outer raceway of the test bearing to be rotated at the desired speed. The inner raceway of the test bearing was static and the contact between the roller and the inner raceway was investigated.

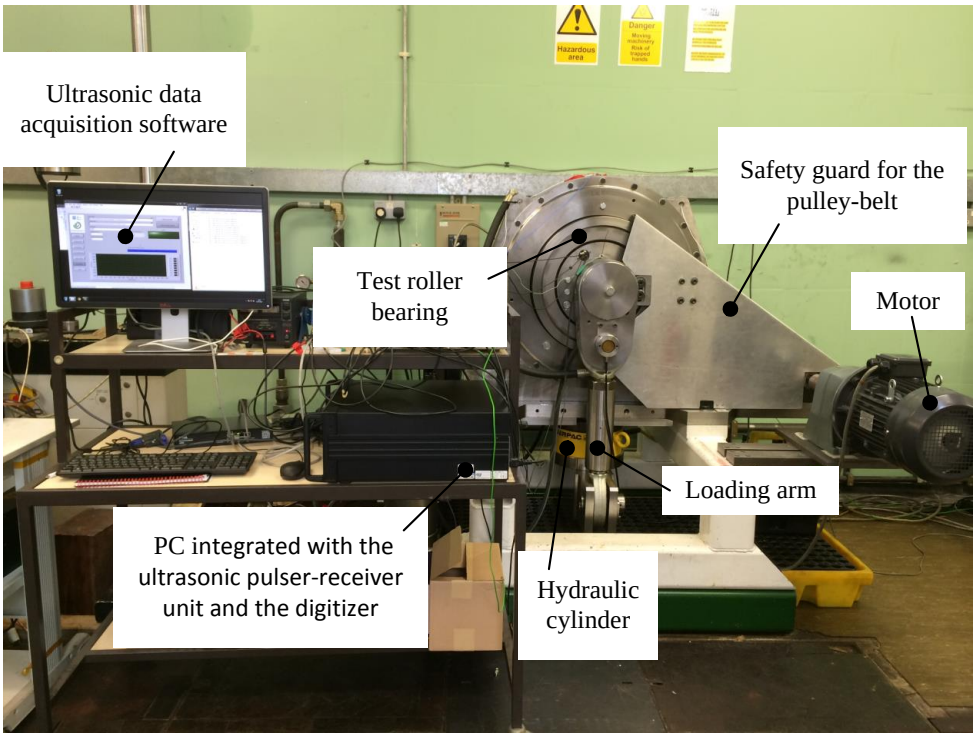


Figure 7.4. *Photograph of the bearing test rig and the ultrasonic measurement system*

The radial load was applied to the bearing through a stationary inner shaft via a large hydraulic cylinder which has a capacity of 1500 kN. The hydraulic load was applied to the roller bearing by pushing the stationary shaft downward. The hydraulic load on the bearing was indicated by two load cells. A schematic diagram is given in figure 7.5, showing the applied load and the sensor location. The dimension of the roller bearing has been given in table 4.2. The pressure in the hydraulic cylinder was controlled to maintain at the desired load on the roller bearing for each test.

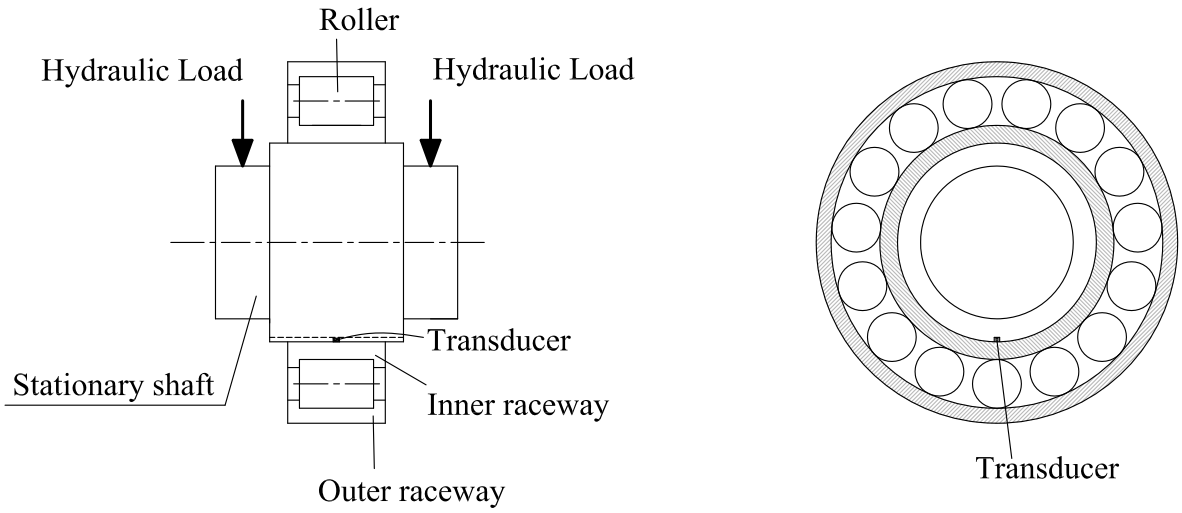


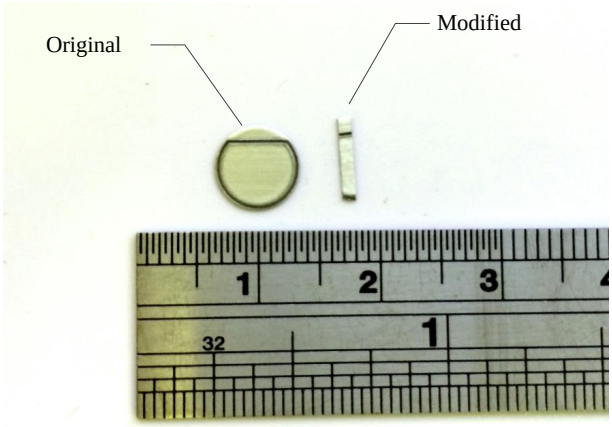
Figure 7.5. *Schematic of the cylindrical roller bearing and transducer location*

An active piezoelectric element having a nominal central frequency of 10 MHz was used to generate longitudinal waves and the sensor instrumentation is described in the following subsection.

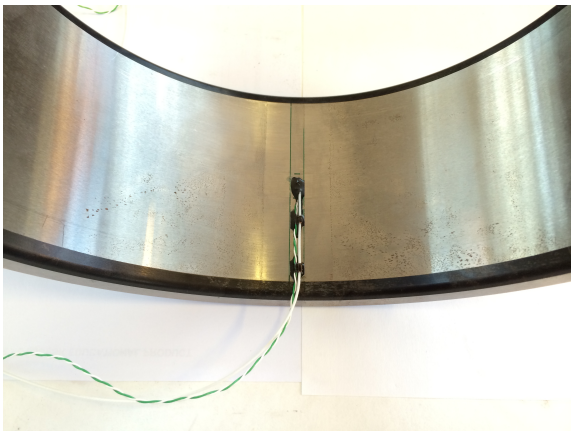
7.3.2 Instrumentation of the ultrasonic sensor

The piezoelectric element was modified from its original disc shape into a narrow strip having width of 1 mm and length of 7 mm. It was directly bonded on the bore surface of the inner raceway by using a high temperature adhesive. A layer of the epoxy was placed over the element to protect it and partially suppress the sensor 'ringing'. Figure 7.6a gives a photograph of the original (as bought) and the modified element, figure 7.6b gives a photograph of the instrumentation of the sensor on the bore surface of the inner raceway. A thermocouple was closely placed to the sensor to monitor the race temperature. The inner race was then mounted on a stationary shaft on which a slot with width of 5 mm was made for the sensor. The circumferential location of sensor was chosen as shown in figure 7.5 so that the sensor measured the most highly loaded contact in the bearing.

In this study, the principal interest is to measure the most heavily loaded rolling contact as it gives most measurable changes in ultrasonic reflections, so only one sensor is used without loss of generality. This approach could be applied to multiple contacts; several modified PZT elements could be arranged on the inner raceway so that the load transmitted at each contact can be monitored.



(a)



(b)

Figure 7.6. *(a) Photograph of the ultrasonic sensor before and after modification, and (b) the modified sensor instrumented on the bore surface of the inner raceway.*

The modified piezoelectric element was connected to the pulser-receiver unit which aims to generate short duration voltage pulses and receive the reflected signals. A short duration signal of 100 ns in the top-hat from was employed to excite the sensor and an ultrasonic pulse was generated. The wave propagated from the bore surface to the contact surface of the inner raceway, partly reflected back at the contact face and then received by the same sensor. The reflected signal was digitized at 100M samples per second by the DAQ system and streamed directly to the computer for post data analysis.

7.4 Experimental results

7.4.1 Changes in ultrasonic signal

Just as the model line contact in Section 6.1.4, the sensor picked up several reflections back from contact interface, some of them or all can be recorded as required. Here, the first reflection was isolated from the whole waveform and recorded. When a roller approached the sensor location, the raceway under the sensor was loaded up, then the amplitude and the arrival time of the ultrasonic signal were modified. For each load case, ultrasonic reflections from the contact face when it was loaded and unloaded have been stacked together and stored for data processing.

With the attempt to get more reflected pulses as each roller passage, high pulse repetition rate was used. In the experiment, the pulse repetition rate of the pulser-receiver unit was set to 10k samples per second. In this case, there will be over 27 reflected pulses from the contact can be measured (Note the sampling rate in figure 7.7b is 1 kHz to clearly show transitions caused by the roller passage). The ultrasonic data was acquired for a desired period during which a number of rollers passed underneath the sensor location.

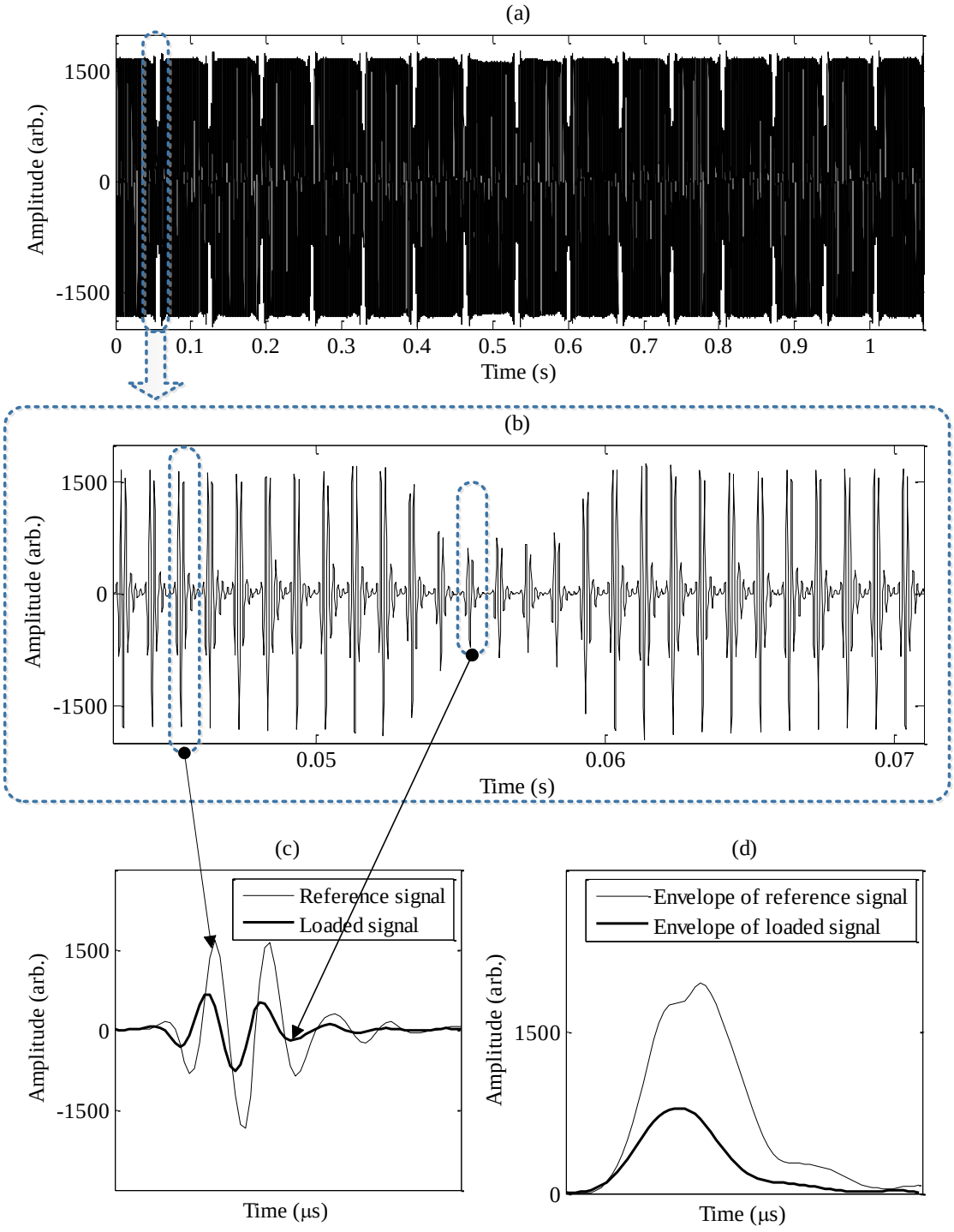


Figure 7.7. *An example of ultrasonic data recorded under the radial bearing load of 980 kN. (a) Streamed data of the first reflection for one complete revolution of the bearing; (b) a section of data when a roller approaches the sensor position; (c) Extracted single pulses for contact and out of contact; (d) envelopes of the signal for contact and out of contact.*

Figure 7.7 gives an example of data obtained from the cylindrical roller bearing. Figure 7.7a shows a sequence of the first reflections from the contact face for a complete revolution with the radial load of 980 kN on the bearing. As described in Chapter 5, a signal window was used to isolate the required reflection from the whole waveform and the isolated part was stored in a First-in-First-out buffer. Then they were sequentially stacked together to form an array. With the information of the pulse rate and the window length, the ultrasonic data can be plotted in time domain. In the figure, the data consists of the first reflections from the contact surface.

There are 15 significant drops in the amplitude, corresponding to each of the 15 rollers passing underneath the sensor. A series of data showing the contact before and after passing the sensor position has been extracted and is shown in figure 7.7b. The effect of the roller passage on the amplitude of the reflection signal is obvious. When the roller was away from the fixed sensor, the ultrasound was reflected from the raceway-lubricant interface and the amplitude was large. The amplitude then varied slightly as the roller moved close to the sensor. With the roller approached the sensor active sensing area, the amplitude fell quickly to a large extent due to a proportion of the energy was transmitted into the roller. When the roller exited the line of sight of the sensor, the raceway under the element was then released of load; the ultrasound was reflected from the raceway-lubricant interface again and the amplitude went back to the normal level as it was before the roller pass.

Ideally, the reference signal should be taken from the unloaded raceway-air interface. However, in real bearing applications, it is more convenient to use the reflection from the unloaded raceway-lubricant interface. Because the rolling contacts are tiny compared with the bearing geometry, the majority of the reflections are from the unloaded raceway-lubricant interface. In the experiment, the modal reflection was obtained from all reflections, representing the most frequently reflected signal from the unloaded raceway-lubricant interface which was then used as reference.

Besides the change in the reflected wave amplitude, the ToF in the inner race also changed. A reference signal was compared with a reflected signal from the contact in figure 7.7c. The change in the ToF is clear; the reflected signal (the thick line) shifts to the left, which consists of the apparent time shift resulted from the phase change effect and the time change from load (the deflection and the speed change effect). The contact dependent phase shift was removed from the envelopes which were obtained by using the Hilbert transform as described in chapter 4. The envelope of the reference signal and that of reflected signal from the contact is given in figure 7.7d.

7.4.2 Results

7.4.2.1 B-scan data

To show the patten of the ultrasonic signal, figure 7.8 gives the B-scan data of a sequence of ultrasonic signal which has been presented in figure 7.7b. The first reflection in the whole waveform in time history, located from 6.8 to 8.3 μs (see figure 5.4b, was isolated and rotated 90 degrees in anti-clockwise direction. Then a serials of the first reflections were arranged side by side as a roller passed the sensor location. In the figure, a significant decrease in the amplitude of the first reflection is observed as the roller approaches the fixed sensor. Besides, some ripples in the signal are observed when the roller is close to the sensor location.

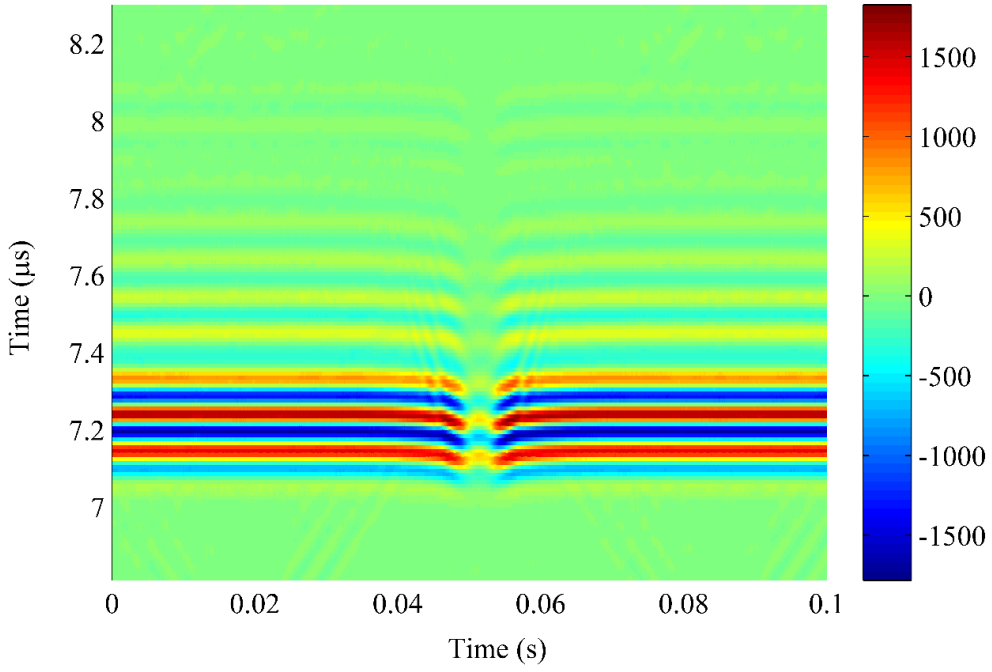


Figure 7.8. *B-scan data for a sequence of the first reflections which shows the passage of a roller, representing data in figure 7.7 (b).*

7.4.2.2 Spectrogram and reflection coefficient

The spectrogram of the first reflections can be obtained by applying the fast Fourier transform (FFT) to the ultrasonic data. The amplitude of the signal was normalized with that of the reference signal, which means the amplitude of the spectrogram is the amplitude of the reflection coefficient.

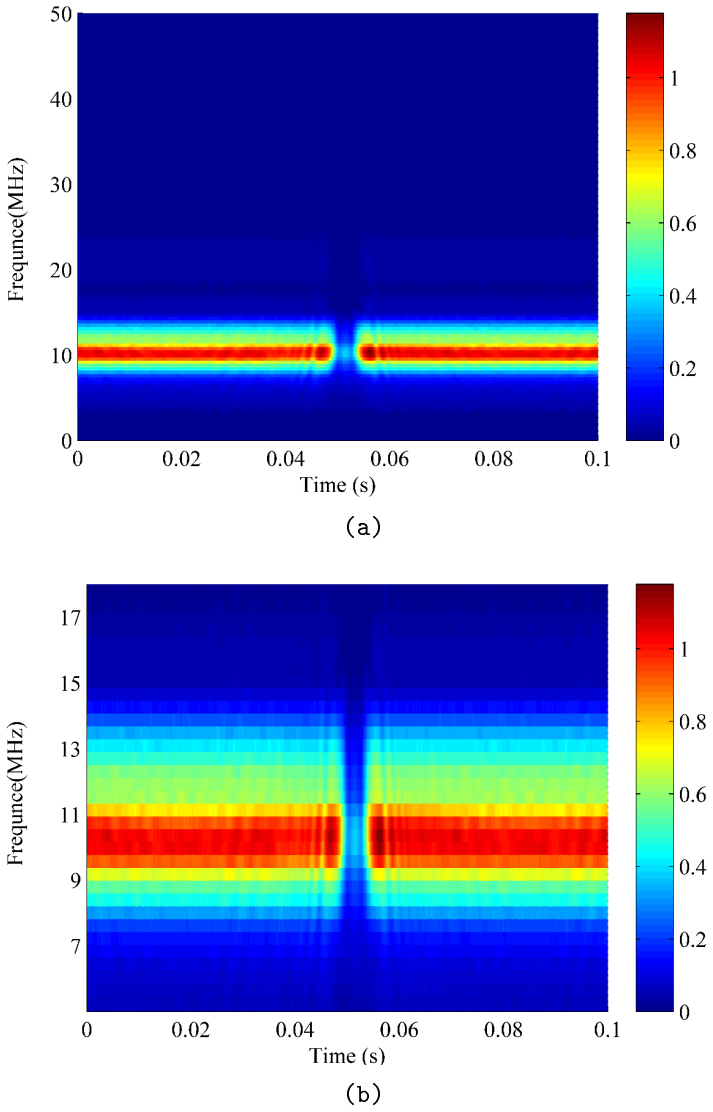


Figure 7.9. *The spectrogram for the first reflections as a roller pass underneath the sensor location under bearing load of 980 kN; (a) frequency from 0 to 50 MHz and (b) frequency from 5 to18 MHz.*

Figure 7.9 gives the spectrogram for the B-scan data in figure 7.8. From figure 7.9a, it can be observed that the wave energy is distributed in a relative narrow frequency band. To be more clear, the spectrogram is given at frequency range from 5 to 18 MHz, shown as 7.9b. It is observed that the central frequency of the sensor bonded on the raceway is 10.2 MHz and the wave energy is mainly distributed in the frequency range from 8.6 to 12.3 MHz. Moreover, the roller passage is evidenced by a large drop in the amplitude of reflection coefficient at the sensor bandwidth frequencies (8.6 - 12.3 MHz).

The reflection coefficient at the central frequency was calculated, shown in figure 7.10,

where the bearing load varied from 380 to 980 kN while the speed of the bearing was set at 100 rpm. The change in the amplitude of the reflection coefficient as each rolling contact passes the sensor is clear in the figure. The amplitude of the reflection coefficient for the inner race-roller contact reduces from 0.5 to 0.3 as the bearing load increases from 380 to 980 kN.

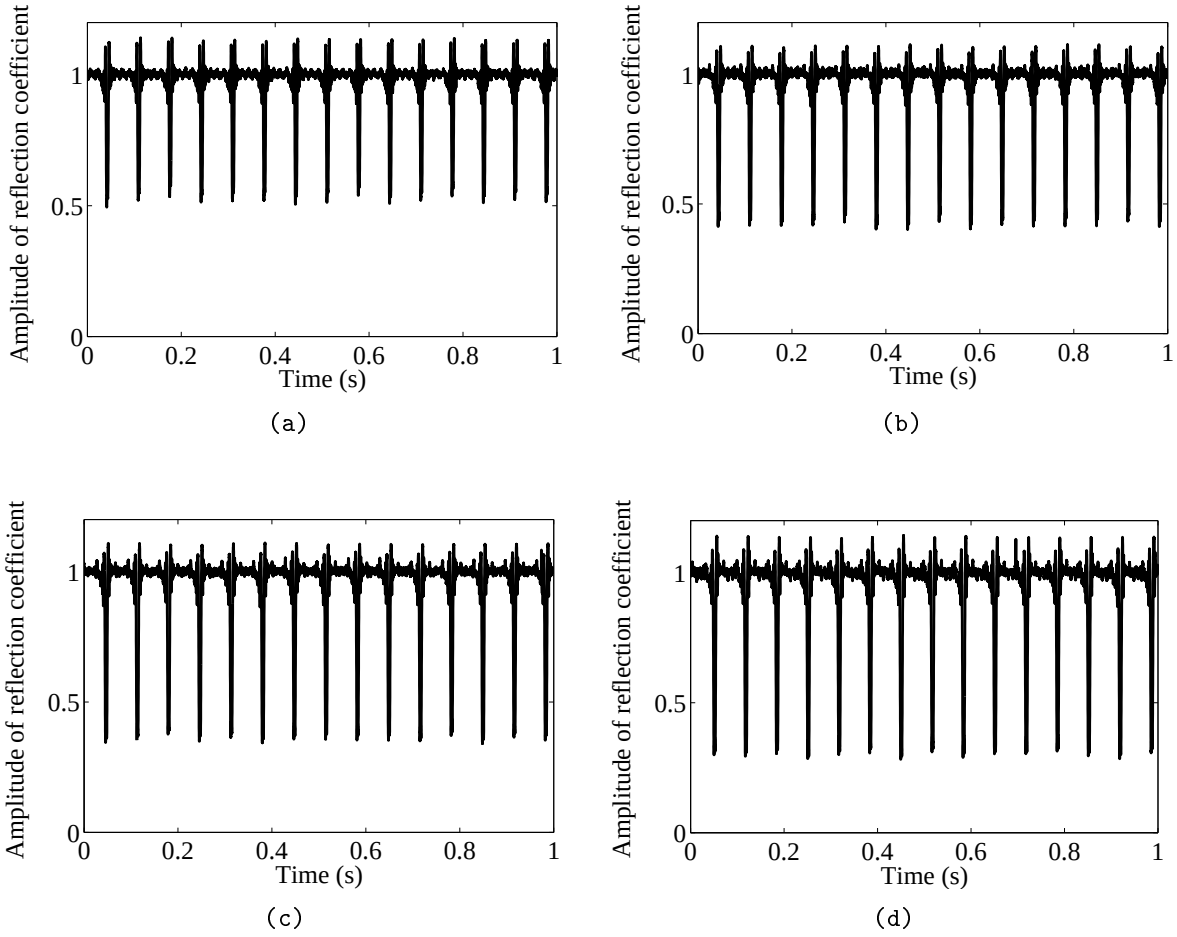
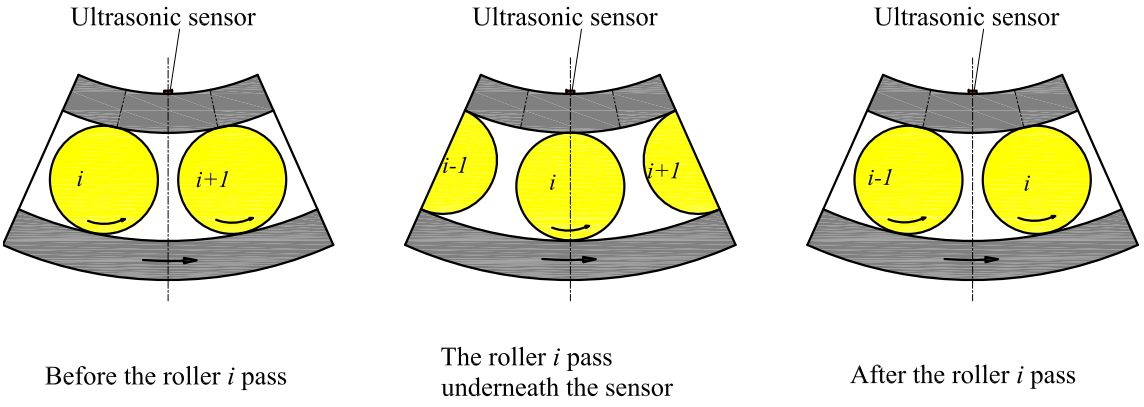
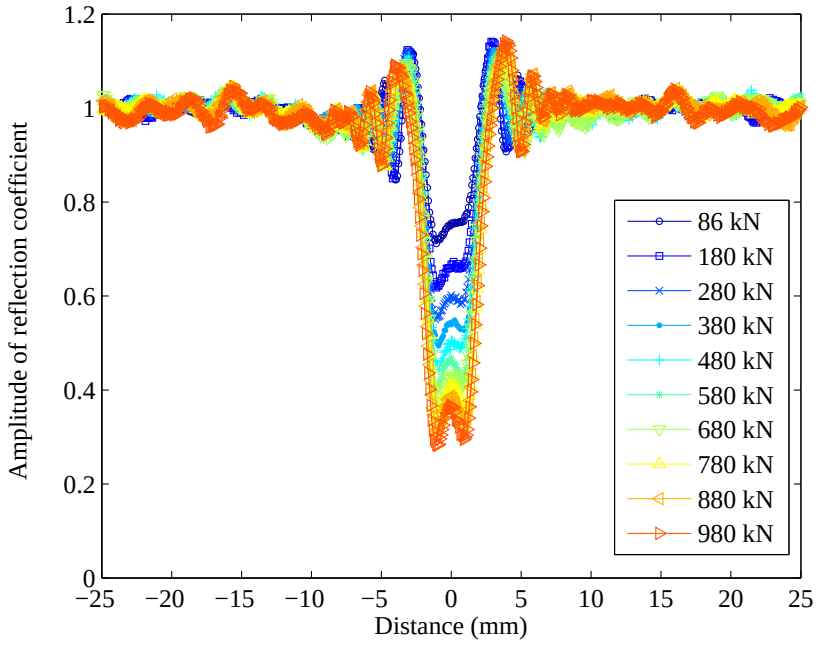


Figure 7.10. *The amplitude of the reflection coefficient changes as the roller pass underneath the sensor location during one second. The bearing was subjected to four different radial loads: (a) 380 kN, (b) 580 kN, (c) 780 kN and (d) 980 kN.*



(a)



(b)

Figure 7.11. *(a) Schematic of rolling rollers and the fixed sensor for three typical stages; (b) the change in the reflection coefficient amplitude due to the roller pass under 10 bearing loads.*

The variation of the reflection coefficient amplitude with the relative distance between the moving roller and the fixed sensor is given in figure 7.11, where the bearing load increases from 86 to 980 kN. Figure 7.11a schematically shows three stages for the roller *i*

passing through the sensor: before the roller pass, the roller pass and after the roller pass. Given the bearing speed and the geometry, the cage speed can be obtained which in turn can be used to calculate the distance between the roller and the sensor. The roller pitch of the test bearing, the circular distance between two rollers, is 54.24 mm, so figure 7.11b gives the reflection coefficient amplitude for positions from -25 to 25 mm. The reflection coefficient curves for all loads are aligned at the center of the sensor. The reflection coefficient amplitude is close to 1 when the roller was away from the sensor, although some oscillations can be observed again when the roller was close to the sensor. It falls greatly when the roller was in the sensor active sensing area, especially for high bearing loads.

7.4.2.3 ToF change

The change in the ultrasonic ToF in the raceway when it is unloaded and loaded can be obtained by comparing the corresponding envelopes. This ToF change is caused by the deflection and the acoustoelastic effect only, the effect of phase shift has been eliminated. Figure 7.12 gives the variation in the ToF change with rolling rollers for one second under four different bearing loads.

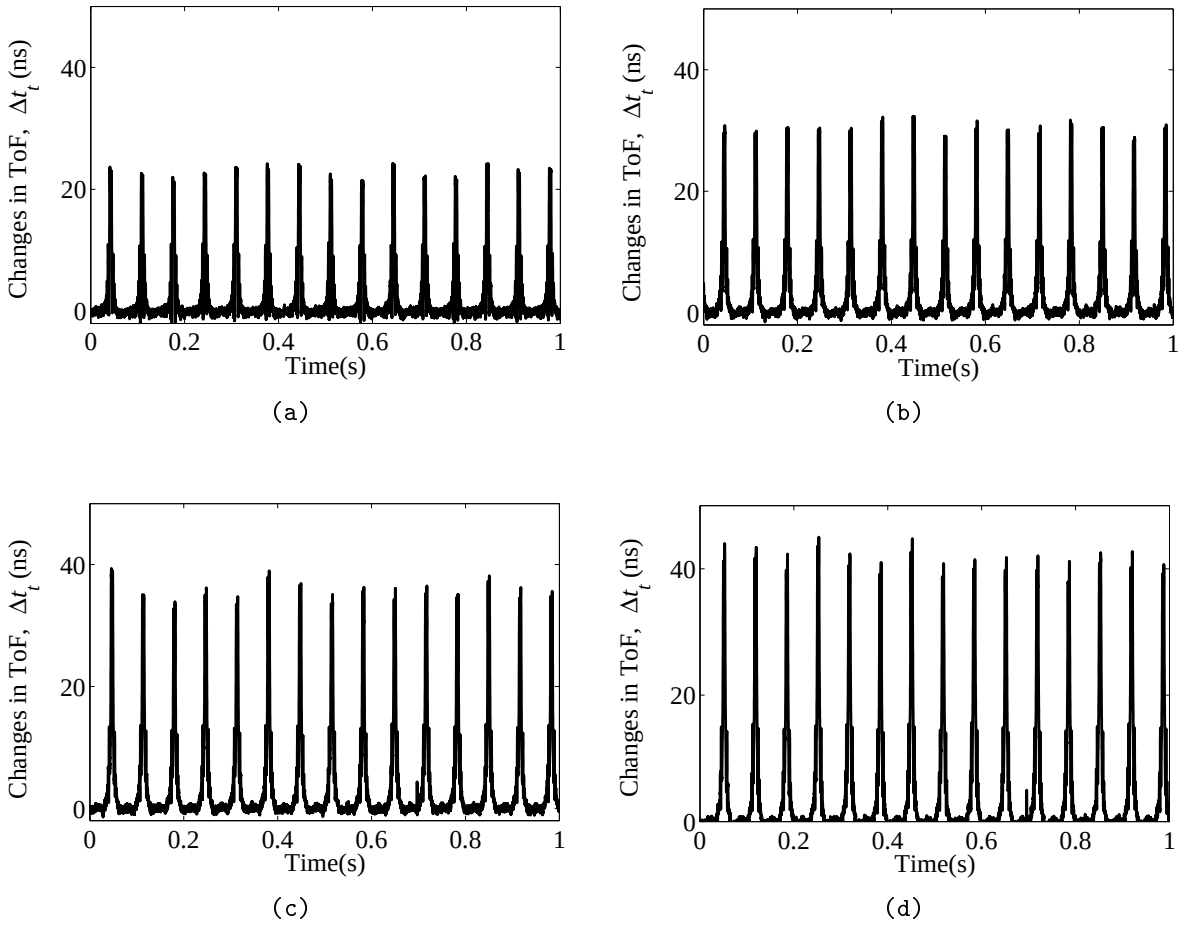


Figure 7.12. *ToF change as rollers pass underneath the sensor location for one second. The bearing was subjected to four different radial loads: (a) 380 kN, (b) 580 kN, (c) 780 kN and (d) 980 kN.*

The ToF in the bearing race reduces with the increase of bearing load, thus the ToF change (i.e. the difference in ToF) increases with load when comparing with that in unloaded raceway. The measured ToF change increases from 23.1 to 43.9 ns when the bearing load increased from 380 to 980 kN. It can be observed that the ToF change under a particular load is not exactly the same for each of the 15 rollers, there appears to be a slight periodic variation. Currently, it is not clear why this should be the case as the roller are all nominally the same geometry and should be subject to the same load as they move into contact. Possibly the tolerance in the roller effective length and some slight misalignment of the cage with respect to the raceway may be the cause.

Figure 7.13 illustrates the variation in ToF change for a roller/raceway contact passage under ten bearing loads. It is noted that the ToF change grows as the contact rolls forward

to the sensor and it is in a distorted M-shape when the contact is close to the sensor. To be more specific, it reaches a maximum when the contact enters into the active sensing area of the sensor, followed by a drop to a local minimum when the contact is right underneath the sensor, and then it grows back to a second maximum when the contact exits the active sensing area of the sensor.

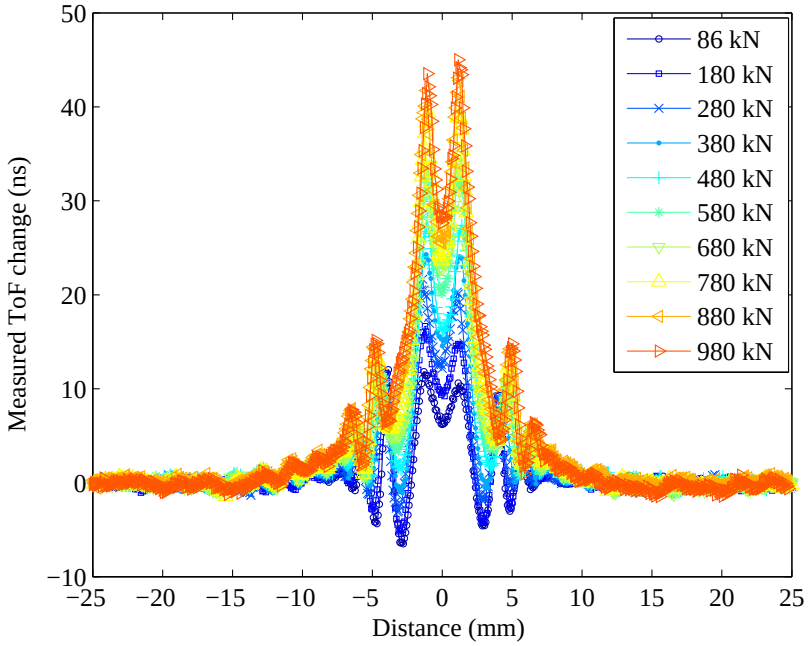


Figure 7.13. *Variation of the ToF change when a roller passes underneath the sensor, for ten bearing loads.*

The reason why the ToF change is formed in a distorted M-shape around the sensor is not clear at this stage. This is may be caused by the non-parallel geometric property between the raceway and the roller. Since the reflection coefficient achieves its minimum at the entry of line of sight, the peak of the ToF change at this position exhibits most sensitivity to the load on the contact, it was then extracted and converted to load in the following section.

7.4.2.4 Comparison between measured and applied contact load

By substituting equation (2.64) into equation (7.3) and equation (7.6), the ToF change caused by the deflection and the acoustoelastic effect can be calculated, respectively. Therefore, the ToF change caused by load, i.e., the sum of the ToF change from the deflection and

the acoustoelastic effect, Δt_t, can be related to the contact load by equation (7.16),

$$\Delta t_t = 3.84 \times 10^{-8} \frac{(1 - L_{zz}) P_{max}^{0.9}}{(c_{zz})_0 \, (l_0 \times 1000)^{0.8}}, \tag{7.16}$$

where the unit of load on the contact is Newton and the unit of length is meter.

By rearranging equation (7.16), the relationship between the contact load and the ToF change can be expressed as:

$$P_{max} = \left[\frac{(c_{zz})_0 \, (l_0 \times 1000)^{0.8}}{3.84 \times 10^{-8} (1 - L_{zz})} \right]^{\frac{10}{9}}. \tag{7.17}$$

Using equation (7.17), the measured change in the ToF can be used to indicate load in the contact. Figure 7.14 compares the contact load measured from ultrasonic reflections and the load applied on the contact. Good agreement is observed, although as with the data for the model line contact in Chapter 6 (see figure 6.20), the load is slightly over-predicted at the low end. It is possibly caused by some misalignment and the roller curvature which have disproportional effects at low loads. For low load cases, the contact areas were small, then the stresses in the contact had a large dependency on the effective length of the line contact which could be varied as a result of some sort of bearing misalignment and the roller curvature. As a consequence, the ultrasonic data has a greater susceptibility to errors for the low load end.

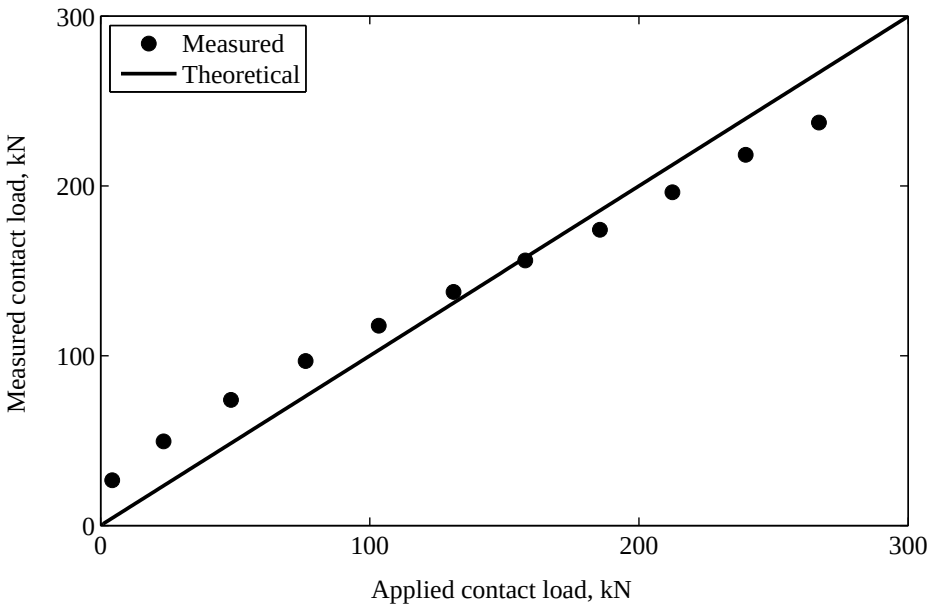


Figure 7.14. *Measured contact load from ultrasonic reflections compared with applied contact load*

As with data in figure 6.20 in Chapter 6, the load from ultrasonic measurements in figure 7.14 is slightly under-predicted for high applied loads. This is caused by the average effect of the transducer. From the Hertz theory, it is known that the surface deformation and pressure are in the elliptic shape. The deformation at the contact centre achieves the maximum and it decreases to zero outside the contact. The distributed deformation causes a distributed variation in the ultrasound ToF. If the incident wave from the sensor is now assumed to behave as a series of discrete pulses, the ToF of each individual pulse depends on the local wave path (or the local deformation) at that particular part of the contact face. However, the transducer senses the cumulative result from the contact when all discrete pulses returned back to the transducer. That means the transducer 'observed' the equivalent deformation effect over the contact zone on the ToF. From the deformation distribution profile (elliptic shape), the equivalent deformation over the contact is less than the maximum contact deformation at the centre. As the contact patch increase with load, this effect is clearer at higher applied load. Therefore, the measured result is less than the predicted value for high load cases.

7.5 Discussion

A few interesting features have been observed in both the reflection coefficient amplitude curves and the ToF change curves for rollers passage. The curves are represented by fives regions again and they are before roller pass region, first transition region, roller pass region, second transition region and after roller pass region. Figure 7.15a plots reflection coefficient amplitude and figure 7.15b depicts the ToF change in the different regions for a roller pass under bearing load of 980 kN. The features in different phases are discussed in the following subsections.

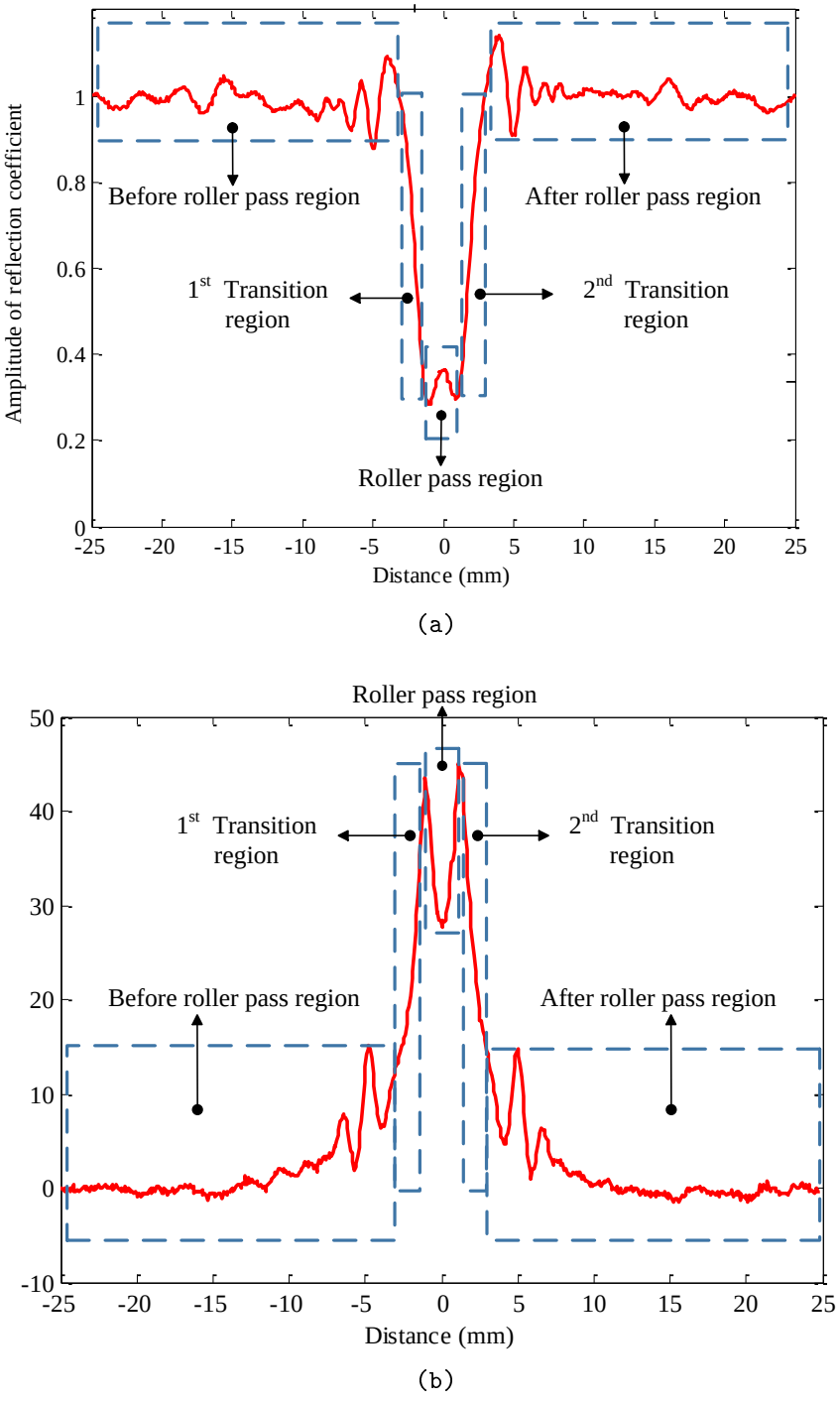


Figure 7.15. *(a) The amplitude of reflection coefficient and (b) the ToF change in different regions for a roller passage when the bearing was under load of 980 kN.*

7.5.1 Fringes before and after the roller pass

When the roller is away from the sensor, before or after rolling through the sensor location, the reflection coefficient amplitude is unity and the ToF change is zero as the ultrasound is reflected from the unloaded raceway. However, a few fringes have been seen when the roller approached close to the sensor, as with that observed in the model line contact test in Chapter 6. Fringes are presented in all load cases (see figure 7.11 and 7.13), which indicates the fringe effect does not depend on load.

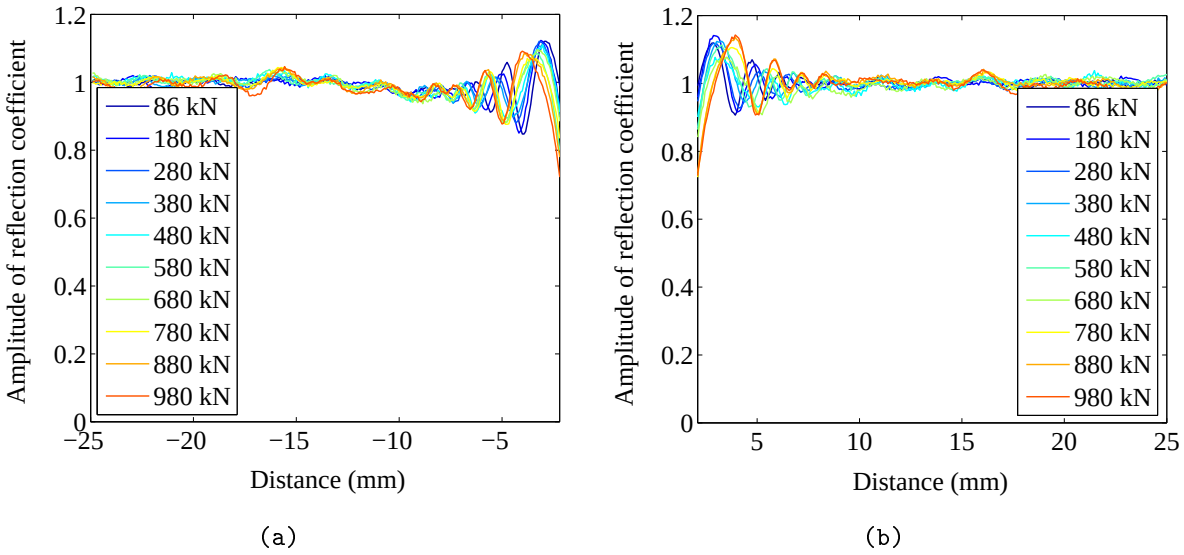


Figure 7.16. *The amplitude of reflection coefficient (a) before the roller pass the sensor and (b) after the roller pass.*

The amplitude of the reflection coefficient (a) before the roller pass and (b) after the roller pass are given in figure 7.16. It is noted that the reflection coefficient amplitude exceeds over 1 at some positions which means the amplitude of the reflected wave are larger than that of the reference. This is probably caused by the wave interference at the contact interface.

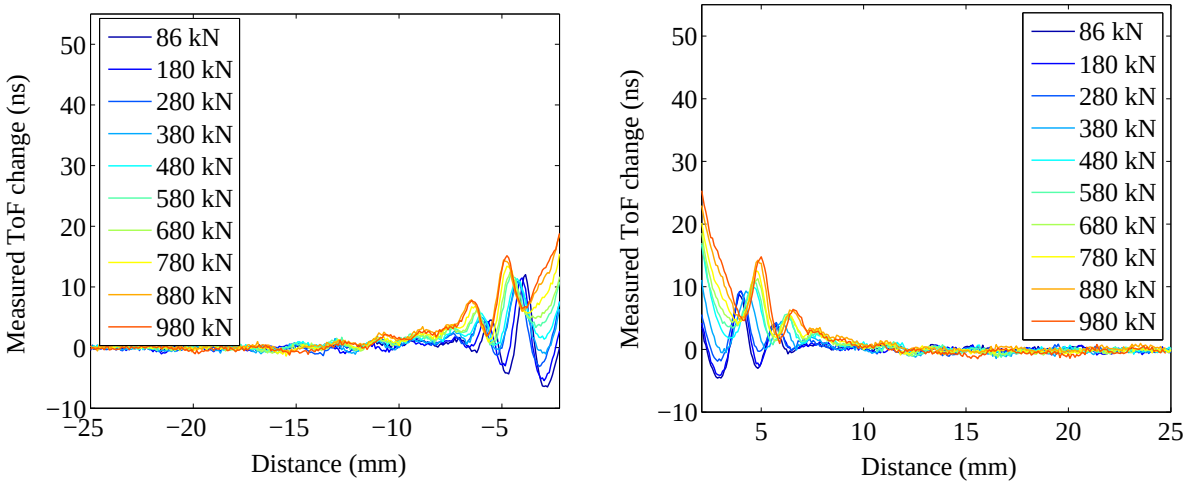


Figure 7.17. *The amplitude of reflection coefficient (a) before the roller pass the sensor and (b) after the roller pass.*

The change in the ToF (a) before the roller pass the sensor and (b) after the roller pass the sensor position is given in figure 7.17. Although the local peaks in the amplitude of reflection coefficient are almost the same for all load cases, the local minimums in ToF changes are different. In fact, the fringes in the ToF change curves increased gradually with increasing bearing load.

The width of the fringe depends on the distance between the roller and the sensor rather than applied load. It is small when the roller is far away from the sensor and increases as the roller moves forward to the senor. This can be observed in the measured ToF change curves and the reflection coefficient amplitude curves, however, it is more obvious in figure 7.16. The non-parallel contact surfaces and the interference effect between non-focused waves may be the cause.

On the other hand, load affects the location where the first fringe occurs before roller pass (or the last fringe disappears after the roller pass). Indeed, the fringe effect appears earlier for higher load. For example, the first fringe peak in the measured ToF change curve for load of 980 kN is observed around -6.5 mm while the corresponding fringe peak occurs around -5.5 mm for load of 86 kN, as shown in figure 7.17a.

7.5.2 Measured roller pass region

7.5.2.1 Width of roller pass region

The trough in the reflection coefficient amplitude curve indicts the roller pass region. However, the roller pass region measured from the first ultrasonic reflection, is larger than the Hertz contact width. From Hertz theory, the contact width increases with load, which is shown in figure 7.1. The width of the contact grows from 0.23 to 1.84 mm as bearing load rises from 86 to 980 kN. However, the width of roller pass region in figure 7.11, is close to 2 mm for all load cases. The independence of this measured roller pass region on load suggests that there are some matters related to the geometry.

Again, this can be explained by the spread phenomenon of ultrasound beam along the wave path. In the experiment, the ultrasonic wave from the sensor was not focused and it spread as traveling in the raceway, which makes the line of sight at the contact interface is larger than the width of the modified piezoelectric element, as shown by figure 6.28. Therefore, the roller pass region in figure 7.15 is the sensor effective measuring width rather than the real contact width.

7.5.2.2 Differences in roller pass region

The amplitude of reflection coefficient in the roller pass region is not constant but fluctuates around a minimum value for each load. It drops to the minimum when the line contact entries ultrasound line of sight, which is determined by the contact condition, rises slightly to a local peak at the center, and then falls back to its minimum value when the contact exits the line of sight. The difference between the minimum and the local peak for each load in figure 7.11b is plotted in figure 7.18. From this figure, it is clear that this difference depends on load.

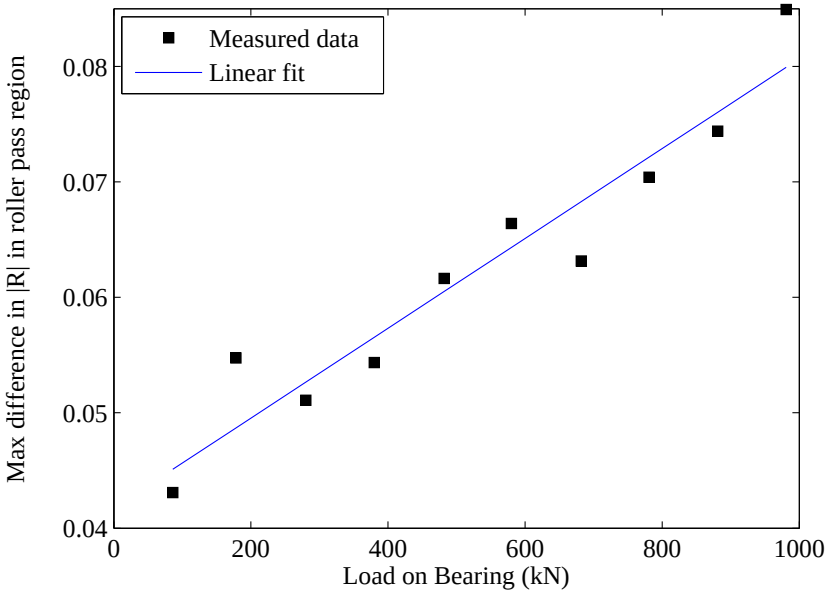


Figure 7.18. *Difference between the minimum and the local maximum on the reflection coefficient amplitude curves in the roller pass region.*

Corresponding, the change in ToF when the roller passes through the sensor is not constant. The ToF change achieves the maximum value at the entry of the ultrasound line of sight. It reduces to some extent at the center and then increase back to it maximum at the exit of the line of sight. The difference between the maximum and the local minimum values under all loads in figure 7.13 are given with load, see figure 7.19.

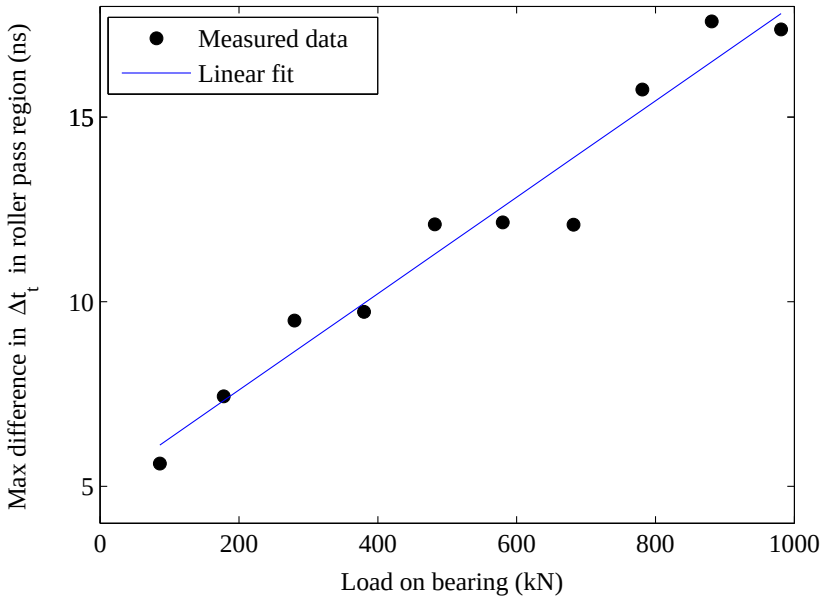


Figure 7.19. *Difference between the maximum and the local minimum on the ToF change curves in the roller pass region.*

For both the amplitude of reflection coefficient and the change in ToF in the roller pass region, the difference increases with load. This is possibly caused by the interference effect among the wave from different parts of the sensor if the sensor is treated as a collection of small point wave sources. As the bearing load increases, the wave path in the raceway reduces and the wave interference has an increasing effect.

As the sensor works in the near field, where complex interference occurs, the features described in Section 7.5.1 and 7.5.2 need further study. Future work can be done in modeling the sound field in the bearing race and measuring from experiments.

7.5.3 Subsequent reflections

Further investigations were taken on the subsequent reflections from the contact face with the aim of exploring features observed from the first reflection.

Figure 7.20a gives a B-scan data for the second reflection under bearing load of 980 kN. The amplitude of the second reflection reduces greater than that of the first reflection when the contact approaches the sensor as the ultrasound travels four times of the thickness of the raceway and reflection happens twice. By using the FFT, the B-scan data was transformed into frequency domain and the spectrogram for the second reflection was obtained. Again, the second reflection from the unloaded raceway was taken as reference. Figure 7.20b gives

the spectrogram for the second reflection in which the amplitude is normalized by the reference amplitude. According to the definition of reflection coefficient, the amplitude in the figure is then the square of the amplitude of the reflection coefficient. The spectrogram is only given at frequencies from 5 to 18 MHz.

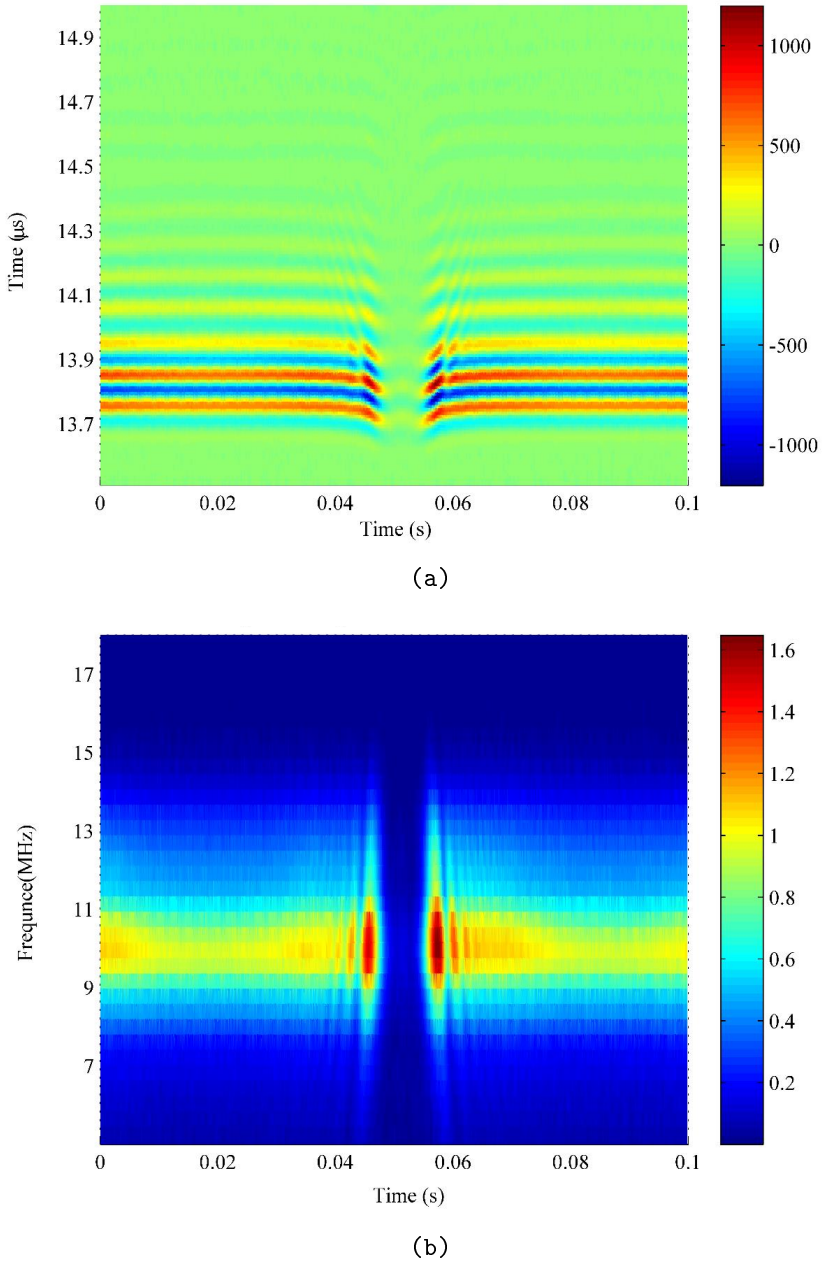


Figure 7.20. *(a) B-scan data for the second reflection and (b) the corresponding spectrogram under bearing load of 980 kN. The amplitude of spectrogram was normalized, so it is the square of amplitude of reflection coefficient.*

Figure 7.21a shows B-scan data for the third reflected pulses under bearing load of 980 kN. The same procedure was applied to the B-scan data and the corresponding spectrogram was obtained in figure 7.21b. Then the amplitude in spectrogram is the cube of the reflection coefficient amplitude.

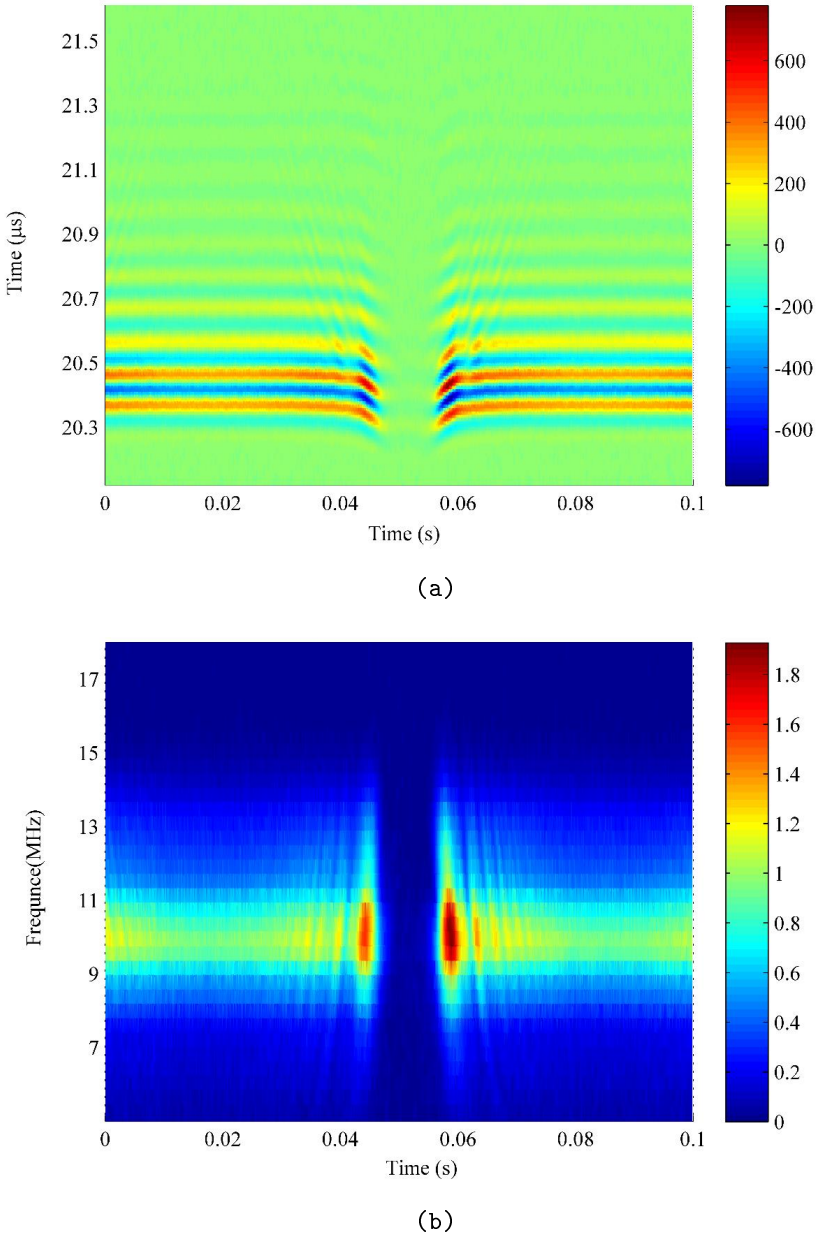


Figure 7.21. *(a) B-scan data for the third reflection and (b) the corresponding spectrogram under bearing load of 980 kN. The amplitude of spectrogram was normalized, so it is the cube of amplitude of reflection coefficient.*

By comparing the B-scan data for the first three reflection in figure 7.8, figure 7.20a

and figure 7.21 a, it is observed that there are more ripples when the roller approaches the sensor as the order of the reflection increases. Correspondingly, there are more fringes in the spectrogram of in the second and the third reflection than that of the first reflection, as can be seen in figure 7.9b, figure 7.20b and figure 7.21b.

The amplitude of reflection coefficient at the central frequency for the first three reflections were obtained when a roller passes through the fixed sensor under bearing load of 980 kN. They are shown in figure 7.22. The fringe effect appears first in the third reflection, followed by the second reflection, then in the first reflection when a roller passes through the fixed sensor. Except the fringes, the reflection coefficient amplitude for the first three reflections follow the same trend; they are close to 1 as the roller is far away from the sensor, drop to 0.3 as the roller approaches the sensor and then increase back to 1 as the roller rolling away. This is useful in showing that the reflection coefficient amplitude method works on subsequent reflections for indicating the contact.

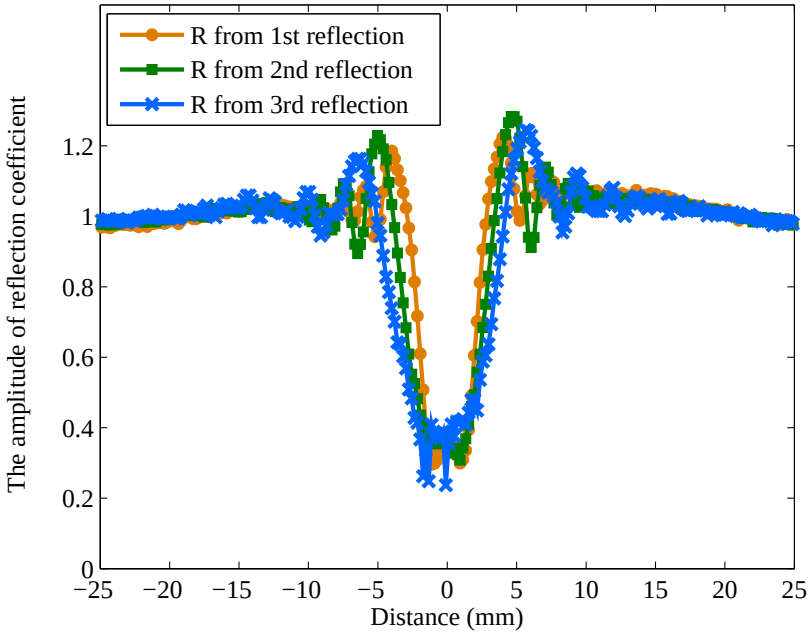


Figure 7.22. *The reflection coefficient amplitude, from the first three reflections, as a roller pass underneath the sensor.*

7.6 Conclusion

Direct load measurements were carried out on the roller/raceway contacts in a cylindrical roller bearing using the ToF method of ultrasonic reflections. A piezoelectric sensor was

bonded onto the bore surface of the inner raceway to monitor the rolling contacts formed between the inner race and the roller. The first ultrasonic reflections from the contact face under a series of load were recorded and analyzed. For each load, the ultrasonic ToF in the race reduced when the contact approached the sensor since the race was loaded by the coming roller, which resulted in the surface deformation and the increase in the wave speed. The Hilbert transform was used in the signal processing to calculate the signal envelope, aiming at eliminating the apparent time caused by phase change in the reflected signal which depends on the contact stiffness. Comparing the envelope of signal from the unloaded race with the envelope of signal from the loaded race, the ToF change in envelopes can be obtained, which is load dependent and only caused by the surface deflection and the acoustoelastic effect. By using the elastic model developed in Chapter 4, the measured change in the ToF was used to predict load. It has been found that load from the ultrasonic ToF method agreed well with the applied load. Although some interesting features have been observed, which can be better understand by modeling the sound propagation in the bearing race in future, this technique shows promise for application to large bearings when both the contact dimension and the raceway deflection are large and hence lead to readily measurable ToF change.

Chapter 8

Load measurements on Ball Bearing Contact

The ultrasonic ToF method developed in Chapter 4 has been applied to a roller bearing application for load monitoring as described in Chapter 7, and a ball bearing application which is given in this chapter. After a brief introduction, predictions of the total ToF change in the elliptical contact under different bearing loads are given, where the contribution of the deflection, the acoustoelastic effect and the phase change effect have been studied. It continues to describe the experimental apparatus, including the sensor instrumentation on the ball bearing. Followed by the section of experimental results where the measured ToF change has been used to indicate the contact load. A discussion of the experimental work is given before a conclusion for this chapter.

8.1 Introduction

Ball bearings are widely used in industrial applications and their performance significantly affect the efficiency of the involved machines. Their failures can cause catastrophic events if they are not detected in time and can result in costly downtime and rather high maintenance cost. The load and the load history information are important for engineers to evaluate the bearing residual service life and arrange proper maintenance or replacement in advance. Although the bearing load can be inferred from load cells or strain gauges on the shaft or bearing housing, this is not always simply and uniquely related to the real load transmitted by balls directly to the raceway. In this chapter, the load transmitted in the rolling contacts in a ball bearing can be directly measured by using the ultrasonic ToF technique. A number of load measurements have been done on elliptical contacts in a deep groove ball bearing

using ultrasonic reflections.

8.2 Prediction of the total ToF change

8.2.1 Load distribution in the ball bearing

When the ball bearing is subjected to a radial load as described by figure 2.3, the ball at the bottom of the load line accommodate a maximum load and balls in the lower plane share the bearing load. The number of balls carrying load depends on the bearing geometry. For a ball bearing 6410, the geometry and property parameters of the bearing is given in table 8.1, with which the load distribution in the ball bearing can be calculated using equations in Section 2.2. From the calculation, the bearing load is shared by three balls out of seven: one on the load line and another two positioned an angle of 51.43 degree to the loading line. As the bearing load goes to 40 kN, the maximum load on the ball increases to 26.03 kN and the load on the other two balls increase to 11.09 kN. This is described by figure 8.1.

Table 8.1: Geometry and properties of the deep groove ball bearing

Symbol	Description	Value
OD	Outer race diameter, m	0.1146
ID	Inner race diameter, m	0.06533
D	Ball diameter, m	0.0246
r_o	Groove radius, m	0.01279
N	Number of balls	7
E	Elastic modulus of, GPa	207
ν	the Poisson's ratio	0.3
d_0	Thickness of the outer race, m	0.0077

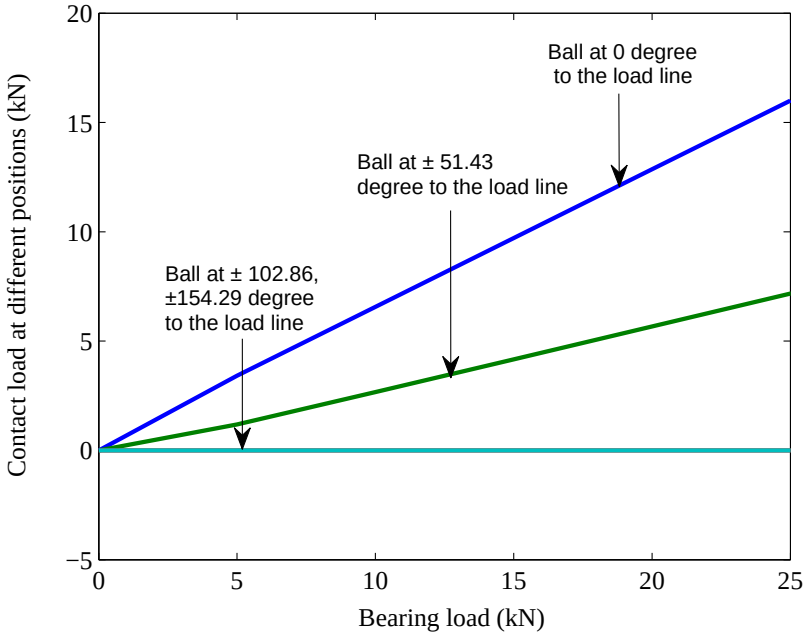


Figure 8.1. *Variation of contact load for balls at different positions with bearing load.*

8.2.2 Surface deflection

The load distribution for each contact in the ball bearing can then be used to determine the contact size, the stress and the surface deflection in the contact. The size of the contact ellipse can be calculated using equations (2.17) and (2.18). For the most loaded outer race-ball contact, the major contact length (2a) increases to 9.2 mm and the minor contact length (2b) enlarges to 1.3 mm as the bearing load grows to 25 kN. The pressure in the contact area is in the form of ellipsoid. At the contact centre, the maximum pressure increases to 2.49 GPa and the corresponding maximum contact deflection is 48.3μm, which can be obtained from equation (2.24) and (2.46), respectively. The major and minor contact length under different bearing loads are given in figure 8.2, represented by the left vertical axis, while the maximum contact deflection is described by the right vertical axis.

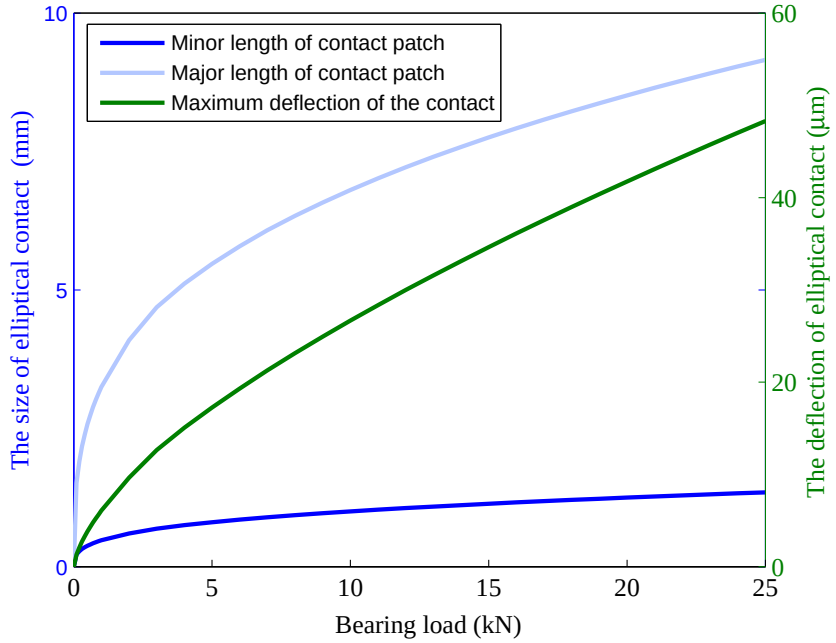


Figure 8.2. *The effect of load on the elliptical contact size and the maximum contact deflection, for the most loaded outer race-ball contact.*

8.2.3 Predicted ToF change

As stated in Chapter 4, the total ToF change is consisted of three components caused by the deflection, the acoustoelastic effect and the phase shift in practical ultrasonic measurements. The contribution of these three components, for the most heavily loaded outer race-ball contact, is studied in this section.

As the occurrence of the elastic deformation in the bearing race, the ultrasonic wave path reduces and hence the ultrasonic ToF reduces. The change in the ToF caused by the surface deflection alone can be calculated by using equation (4.8). Besides the reduction in the wave path, the ultrasound speed increases with load, resulting in further drop in the ultrasonic ToF. The magnitude of the ToF change caused by the acoustoelastic effect can be predicted from equation (4.17). Note that the deflection in these equations is the surface deflection of the outer ring which is half of the contact deflection in figure 8.2.

In addition, the phase shift also causes an apparent change in the ToF which is significant for the ball bearing case. The magnitude is determined by the stiffness of the thin oil layer which is related to the oil thickness. The minimum film thickness for a ball bearing contact

has been developed by Hamrock and Dowson (Hamrock and Dowson (1977)):

$$h_c = 2.69\bar{U}^{0.67}G^{0.53}\left(1-0.61e^{-0.73k}\right)\bar{Q}^{-0.067}R'. \tag{8.1}$$

The central film thickness can then be used to calculate the phase shift by using equation (3.61). As a consequence, the apparent ToF change can be obtained from equation (4.20).

The ToF change due to the surface deflection, the increase in the speed of sound and the phase change effect, under different load conditions with the constant speed at 120 rpm, are given in figure 8.3. First, it can be observed that the total ToF change is small, about 49.6 ns when the bearing load increases to 25 kN. Second, it is clear that all three components are significant and cannot be neglected. The apparent ToF caused by the phase change is comparable large with respect to that caused by the deflection and the speed change, especially for low load cases. Third, it is noticed that the first two components increase with bearing load while the last component remains the same for all loads. To be specific, the ToF change contributed by the surface deflection increases to 8.2 ns, and the ToF change from the acoustoelastic effect increases from 18.3 ns. By contrast, the apparent ToF change from the phase change rarely change with load. This is because the film thickness does not change much with load as shown by equation (8.1); the exponent of the load parameter in the equation is only -0.067.

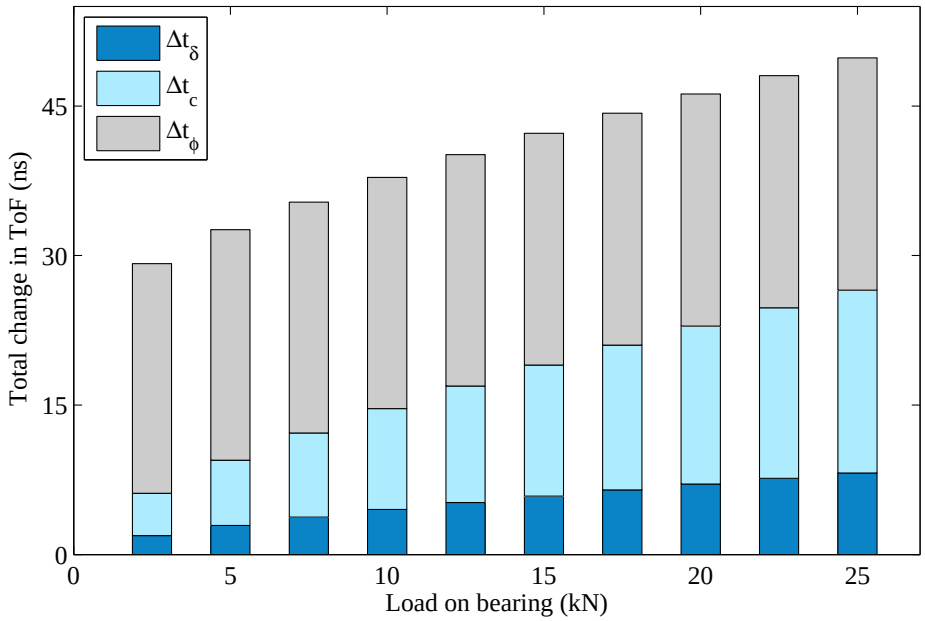


Figure 8.3. *Variation in total ToF change in the ball bearing outer race with bearing load, caused by the deflection, the speed change and the phase change.*

8.3 Experimental apparatus

This section gives a description of the experimental setup, followed by the instrumentation of sensor on the ball bearing. The ultrasonic measurement system described in Chapter 5 has been used to communicate with the sensor and record the data.

8.3.1 Apparatus

Figure 8.4 gives a photograph of the experimental apparatus. The test bearing is a deep groove ball bearing (SKF 6410) and mounted on a rotating shaft which is driven by an induction motor through a pulley-belt system. As such a case, the inner race of the ball bearing is rotational while the outer race is stationary. The motor has a speed controller and the maximum speed of the ball bearing can be up to 1500 rpm. The radial load was applied to the ball bearing through a loading lever arm by a hydraulic cylinder having a load capacity of 100 kN. The ball bearing was lubricated by oil (Shell T68) at a constant flow rate of 0.8 ml/s controlled by a peristaltic oil pump.

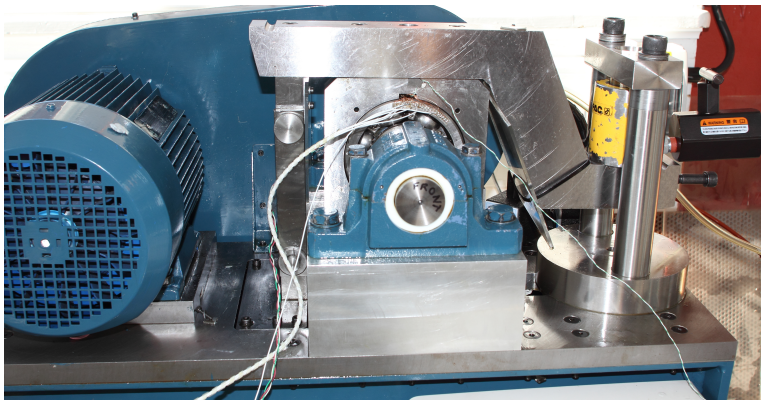


Figure 8.4. *A photograph of the test apparatus for a deep groove ball bearing*

Hydraulic load was not directly applied on the test bearing but via the lever arm which pressed the ball bearing downwards at its center. The radial load on the ball bearing was coupled by a radial load which has same magnitude but acted in the opposite direction (acted upon the loading arm) according to the Newton's third law. The radial load can be determined by analyzing forces upon the loading lever arm.

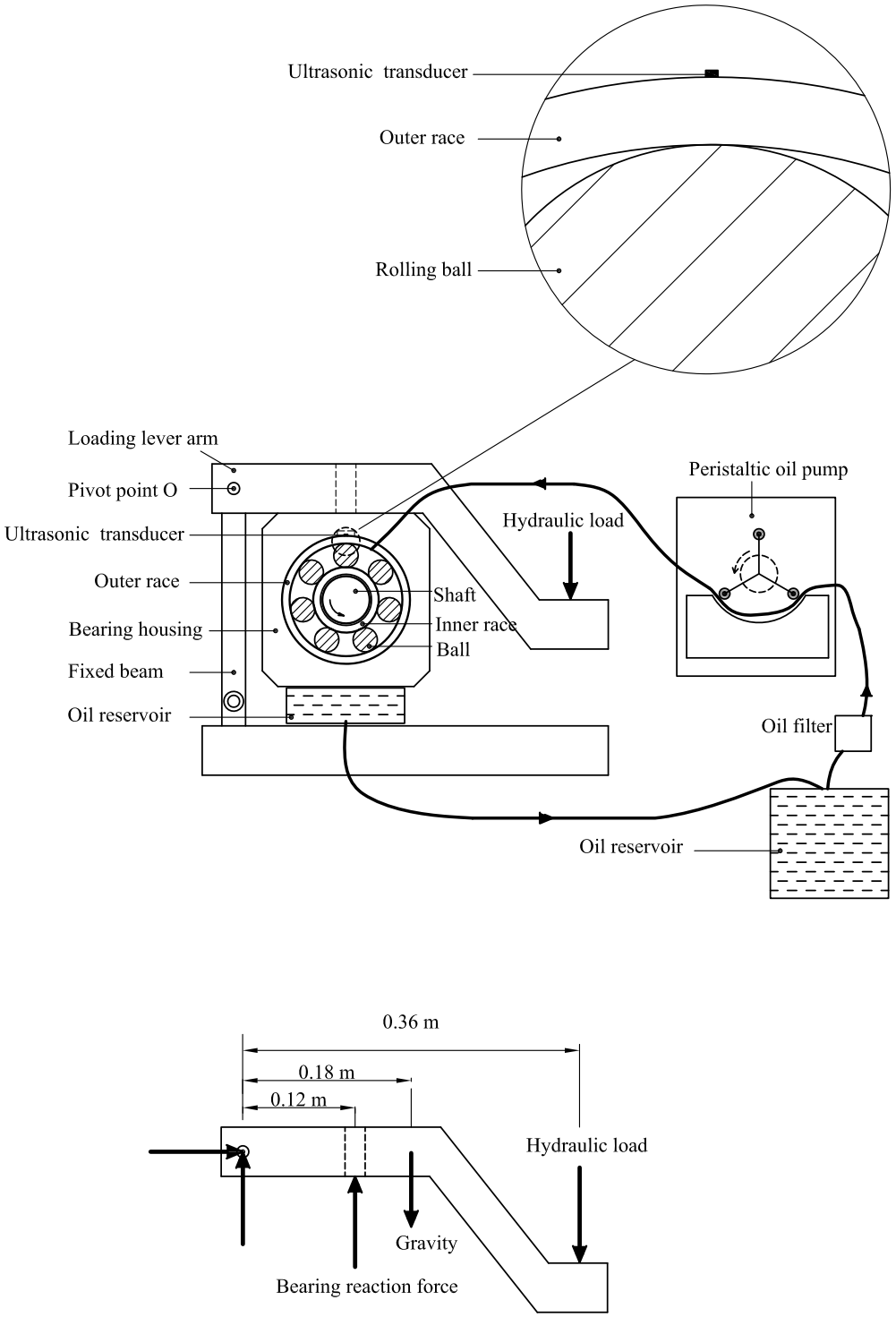


Figure 8.5. *Schematic diagram of the ball bearing test apparatus and force diagram analysis for the loading arm.*

A force diagram analysis is shown in figure 8.5. The distance between the ball bearing centre and the point O was measured at 0.12 m. The loading level arm was so designed that its geometrical center was 0.18 m away from the point O. The distance between the loading position and the point O was 0.36 m. According to the point O, the moment is equilibrium, which gives the expression of normal bearing load, W , in terms of hydraulic load on the loading arm, W_h, as:

$$W = 3W_h + 1.5mg, \tag{8.2}$$

where mg is the gravity of the lever arm, measured at 157 Newtons.

8.3.2 Instrumentation

Ultrasonic measurements can be carried out for all contacts in the bearing if each contact is equipped with an ultrasonic transducer, and thus the load profile in the bearing can be experimentally built up. However, the principal interest of this investigation was given to monitoring load in the most loaded contact.

A modified piezoelectric element, having a nominal centre frequency of 10 MHz, was bonded directly on the external surface of the outer ring. The sensor location was chosen to be on the radial loading line of the ball bearing as shown by figure 8.5 so that the sensor measured the most highly loaded contact. A thermal couple located close to the sensor was used to monitor the outer race temperature. A small slot was made on the bearing housing for the sensor. The modified piezoelectric sensor was a narrow strip having width of 2 mm and length of 3 mm. The sensor width direction was aligned with the ball rolling direction, as shown in a magnification view of the sensor and outer race, the upper diagram of figure 8.5.

The sensor was connected with the ultrasonic pulser-receiver unit to send out and receive ultrasound. The element was excited when it received a short duration electrical signal from the pulser. Ultrasonic pulse was then generated and propagated through the raceway. It was reflected back at the contact face and picked up by the same element. The reflected wave was converted to an electrical signal by the sensor and then sent to the receiver. The signal was digitized by the DAQ system and prepared for the data acquisition process. In the experiment, a PCI cards based PC system, housing the customized pulser and receiver cards as well as the digitizer card, was used to study the reflections from the ball bearing contact.

8.4 Experimental results

8.4.1 Ultrasonic reflection

With the developed sensor and ultrasonic measurement system, the ultrasonic response from the outer race was recorded continuously during the bearing rotation. The outer race was loaded up whenever a ball approached the sensor location and it was released of load when the ball rolled away. The ball passage altered both the amplitude and the arrival time of the reflected signal, which is similar to that happens in the roller bearing case which has been shown in figure 7.7.

In the ball bearing case, an ultrasonic wave diffracts to a large extent due to the presence of the groove curvature. Therefore, the amplitude of the reflected signal is much smaller than that of the incident wave. The amplitude decreases further when a ball come into contact as part of energy is transmitted into the ball. The situation deteriorates with increasing load as larger load makes the outer raceway contact with the ball at a lager area and more ultrasonic energy is transmitted. Also, the curvature on the raceway results in some undesired pulses which makes the identification of normal reflected pulse more difficult. As such a consequence, load measurements in this study were carried out for bearing load up to 22.7 kN, exceeding which the first reflected signal becomes difficult to identify in the whole waveform, as shown by figure 8.6.

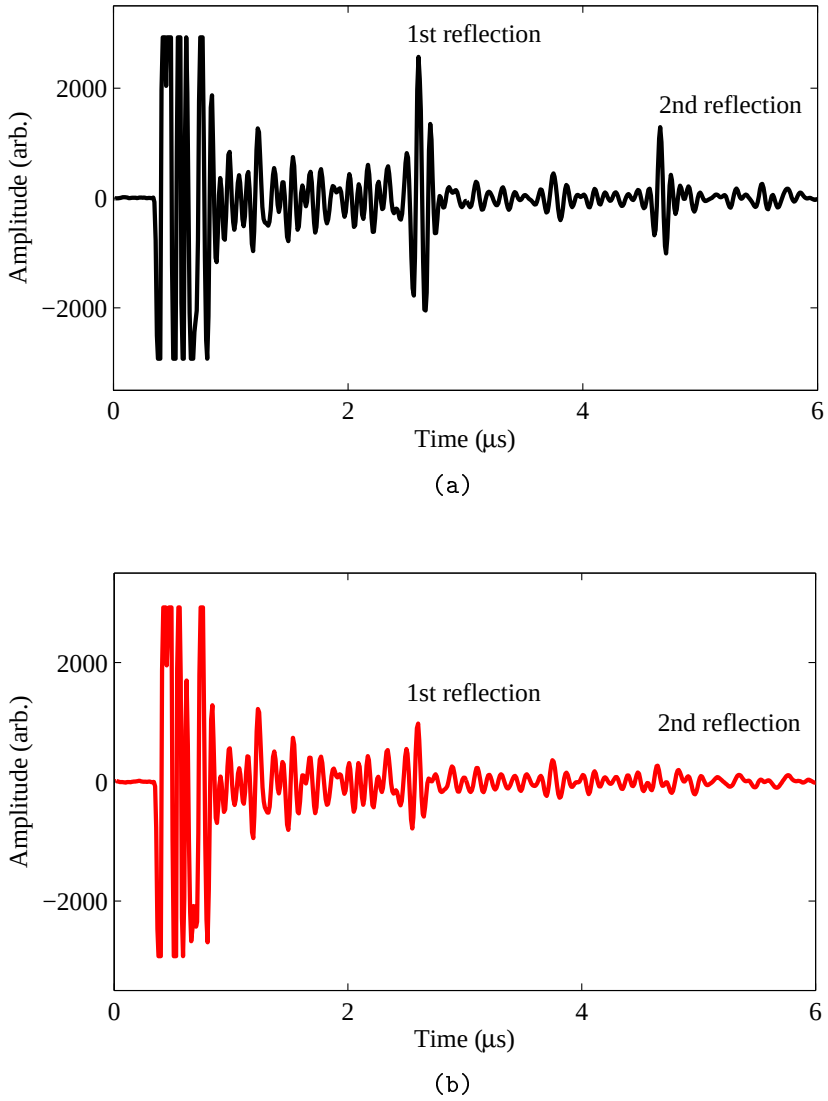


Figure 8.6. *Typical waveform from the ball bearing outer race (a) when it is unloaded and (b) when it is under bearing load of 22.7 kN.*

Figure 8.7 gives a series of the first reflections from the contact interface under different load conditions with shaft speed of 120 rpm. The ultrasonic signal was sampled at the rate of 4000 samples per second. The effect of load on the amplitude of reflected signal is obvious, it reduces with increasing load; while the load effect on the ToF is much smaller, the loaded signal slightly shift to the left in the time trace. The 'kink' in the left of the first reflection is possibly caused by the curvature of the ball bearing outer race. As the presence of the curvature, the wave path for points on the race are different and this may cause wave interference, which is then 'observed' by the sensor. To better understand it, further work is required, like modeling the wave propagation in the bearing race. However,

the principal interest of this work is to measure the ToF change caused by load, and the highlighted portion of the signal has been used to minimize the interference effect.

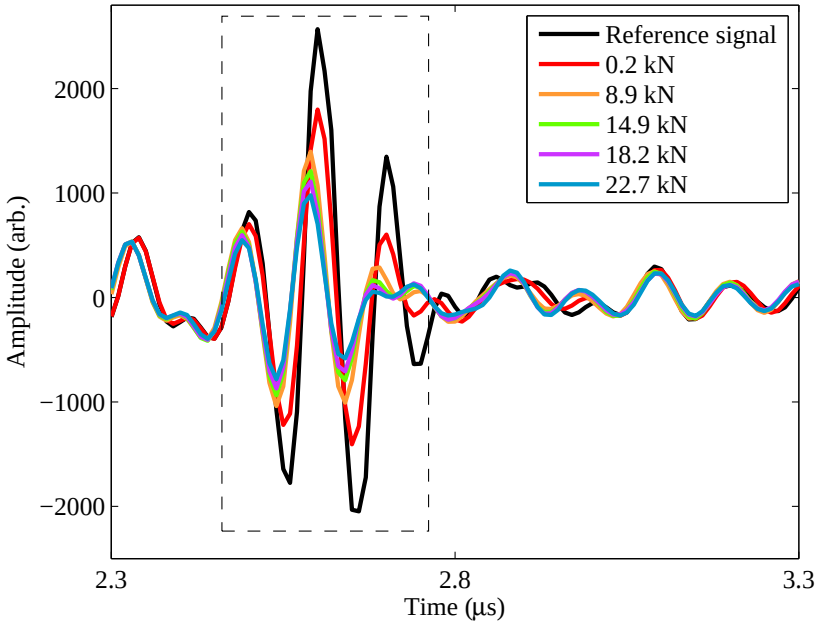


Figure 8.7. *A series of the first reflection from the outer raceway-ball contact under four different bearing loads.*

The length of the near field for the sensor on the ball bearing is 3.81 mm, calculated from equation (6.4), and the sensor worked in the far field. Therefore, the fringe effect, which has been observed in line contacts in Chapter 6 and 7, has not been seen in the ball bearing measurements.

To eliminate the apparent time from the phase change effect, in the signal processing it is the time shift between the envelops that has been measured rather than the time shift between the loaded and the reference signal. The techniques used to measure this time shift has been documented in Section 5.3.1. This approach was applied to the reflected signal for all load conditions, then the ToF change, caused by the deflection and the speed change in outer raceway, was obtained for each load case.

8.4.2 ToF change for different loads

A series of hydraulic loads were applied to the ball bearing and the ToF change under each load condition was measured. The change in ToF under four different loads are given in figure 8.8 as an example. The load on the bearing increased from 8.9 to 22.7 kN, giving a

corresponding maximum contact load from 5.9 to 15 kN. Compared to the unloaded raceway (when the ball was far away from the sensor), the ultrasonic ToF reduced by 25 ns when it was loaded by the ball under a bearing load of 8.9 kN. The difference in the ToF in the outer ring increased to 36 ns as the bearing load went up to 22.7 kN.

Again, it is obvious that the ToF change is not the same for each ball passage even these balls are nominally the same and should carry the same load. The light fluctuation in the ToF change may be attributed to some misalignment of the cage with respect to the raceway. The variation in ToF change between seven balls reduces with increasing load.

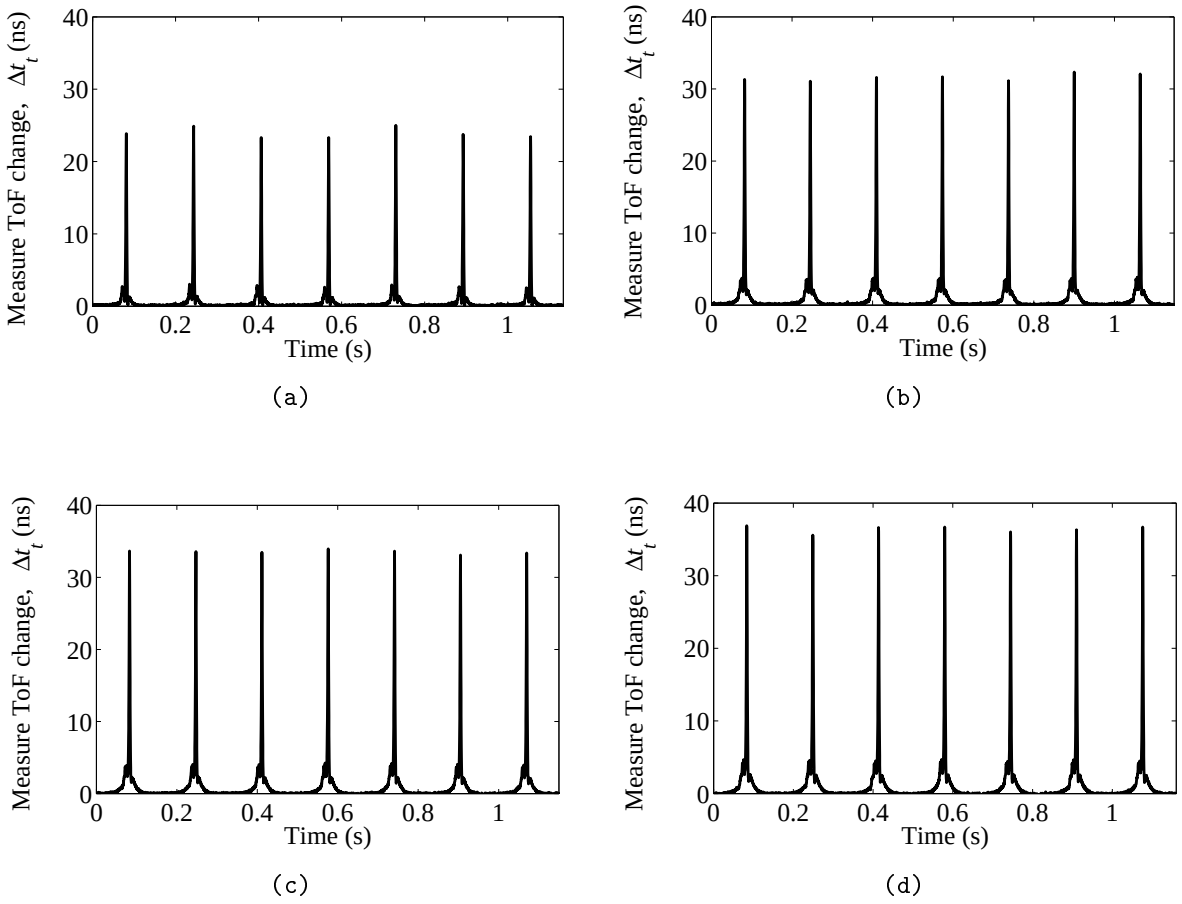


Figure 8.8. *Measured ToF change for one complete revolution under four different bearing loads, caused by deformation and ultrasound speed change. The bearing load are (a) 8.9 kN, (b) 14.9 kN, (c) 18.2kN and (d) 22.7kN.*

For each particular load, ultrasonic measurements have been done several times. The average ToF change between balls has been obtained for each load case. Figure 8.9 gives the measured ToF change under 12 loads. It can be seen that the measured ToF change

non-linearly increases with the bearing load; it has a relative large value when the bearing has no hydraulic load.

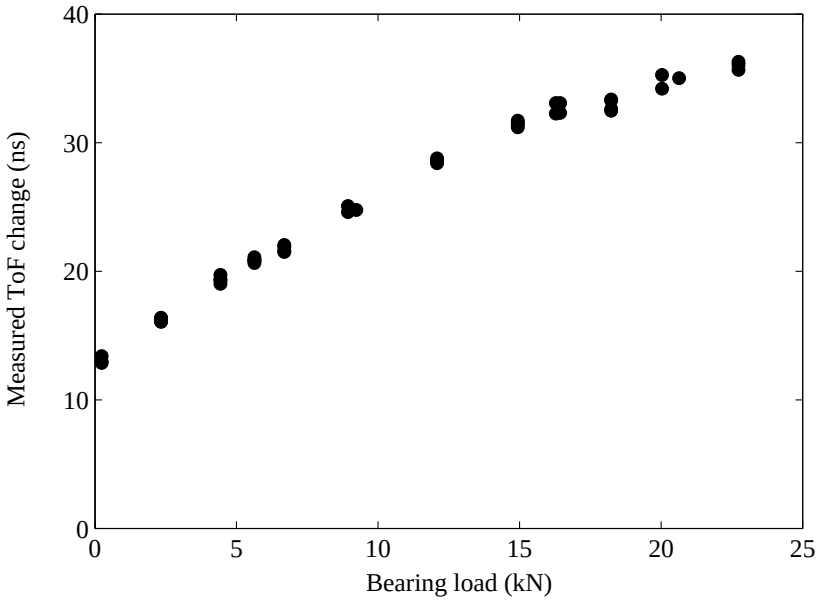


Figure 8.9. *Measured ToF change under different bearing loads.*

8.4.3 ToF change for the initial contact in unloaded bearing

This section explores the interesting feature that the ultrasonic ToF 'observed' relative large change when the bearing has no applied hydraulic load. Figure 8.10 shows the first reflections from the contact face of the outer ring when the bearing had no hydraulic load. In the figure, the reference signal (the dark line) was taken when the ball was away from the sensor, the contact signal (the red line) was taken when the ball was right under the sensor position. The effect of the ball-race contact is clear in the figure: the amplitude of the contact signal reduces greatly.

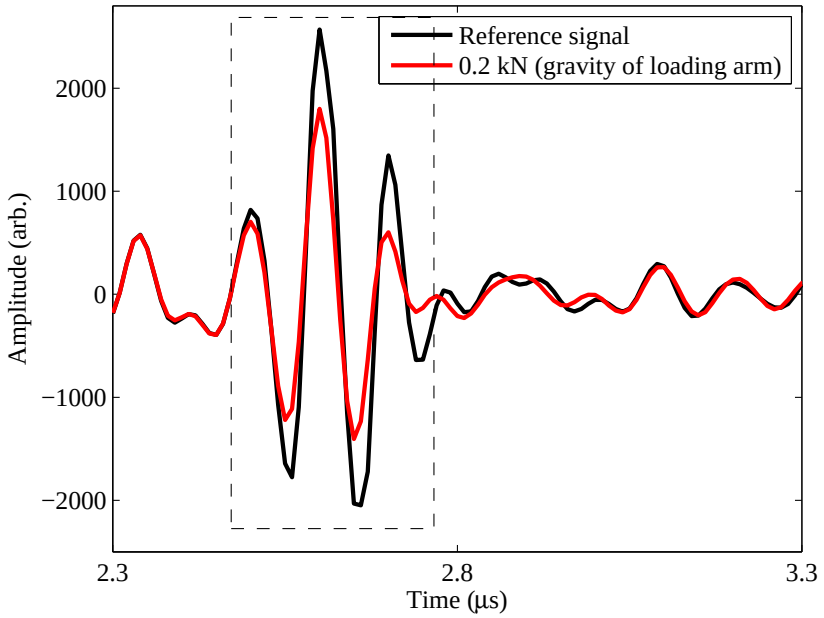


Figure 8.10. *The effect of the initial ball-race contact on the first reflection when the bearing has no hydraulic load.*

In figure 8.10, the time shift between signals is not clear. However, it can be obtained for each initial contact passage by using techniques in the signal processing as described in Section 5.3.1. Figure 8.11 shows the variation in the ToF as rolling balls pass the sensor location when the bearing had no applied hydraulic load. In the figure, the ToF change is given for a complete revolution and seven peaks are clear which represent for seven ball passages. It can be observed that the ToF reduces by about 12.3 ns for a ball passage even no hydraulic load was applied on the bearing.

When no hydraulic load was applied, the bearing was under the gravity effect of the loading arm alone; according to equation (8.2), this load was only 235 N. Under such a small load, the theoretical ToF change, caused by the deflection and the acoustoelastic effect, was calculated at only 1.5 ns. There is a difference of 10.8 ns between the theoretical value and the ToF change observed in figure 8.11. This is interesting because the apparent time shift from the phase change effect has been excluded and there should be no such a difference. This difference has been observed in all measurements and it is not load dependent. The reason why this is the case is currently unknown; one possible reason is the wave interference caused by the conformal contact between the ball and the grooved outer race.

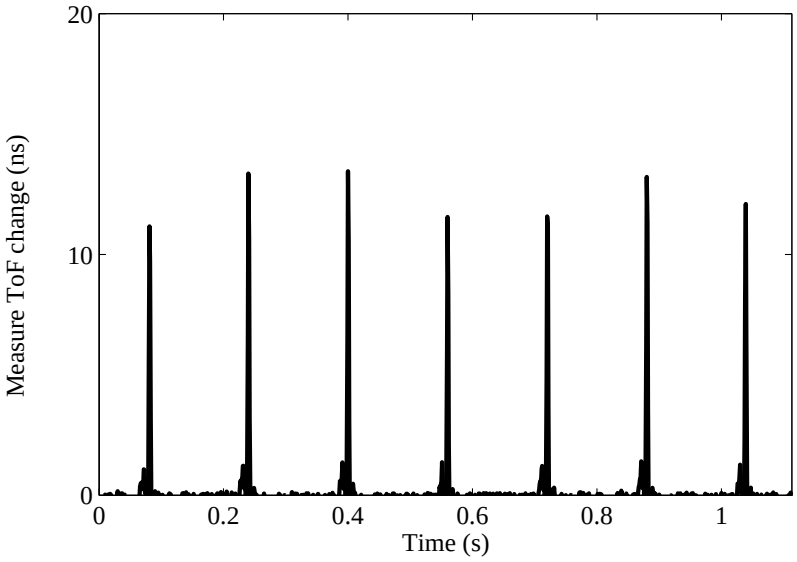


Figure 8.11. *Measured ToF change when the bearing was not applied hydraulic load*

As shown by figure 8.11, some undisclosed factors (like geometry) other than load causes the ToF change for each initial ball passage and they take place in all tests. In order to exclude these effects, all measured ToF change for ball passage are subtracted by a corresponding ToF change of 10.8 ns. Figure 8.12 gives the corrected ToF change with increasing bearing load. In the figure, the average ToF change for all ball passages was used for each load case. The theoretical ToF change for each load condition can be predicted by using equations (8.2), (2.11), (2.46) and (4.25). The results are given in figure 8.12.

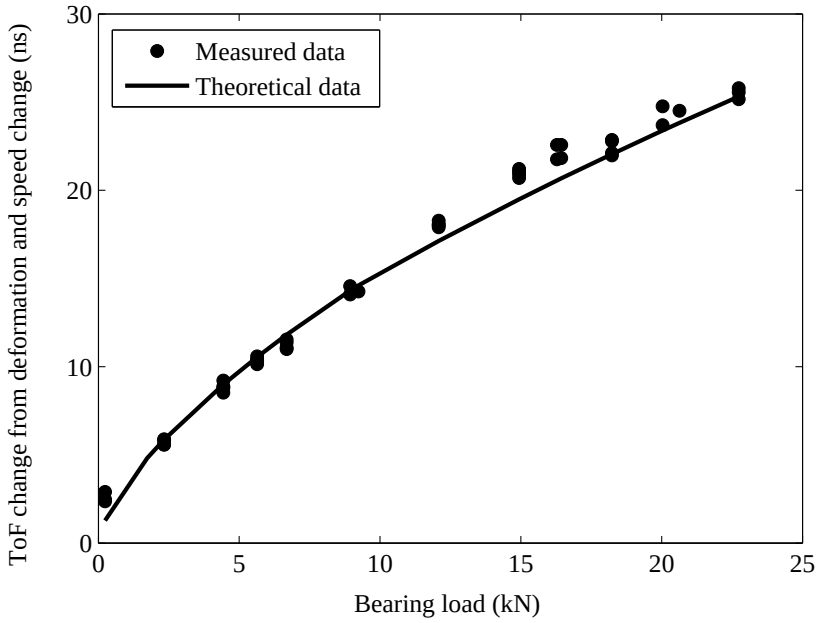


Figure 8.12. *Measured ToF change under different bearing loads compared with predictions.*

8.4.4 Comparison between the measured and the applied contact load

Once the change in the ToF is measured from ultrasonic reflections, it can be interpreted as the surface deflection by using equation (4.25). As the deflection dependents on the contact load (see equation (2.46)), the measured ToF change can be used to indicate the contact load. Note that the deflection in the bearing race δ is half of the contact deformation δ_m in equation (2.46). The relationship between the contact load and the ToF change can be described by equation (8.3), obtained by substituting equation (2.46) into (4.25).

$$\Delta t_t = \frac{(1-L_{zz})}{(c_{zz})_0} F \left[\frac{9}{2\varepsilon R'} \left(\frac{P}{\pi k E'} \right)^2 \right]^{\frac{1}{3}}, \tag{8.3}$$

This relationship can be described by figure 8.13, which is similar to figure 8.3 while the bearing load is replaced by the corresponding contact load.

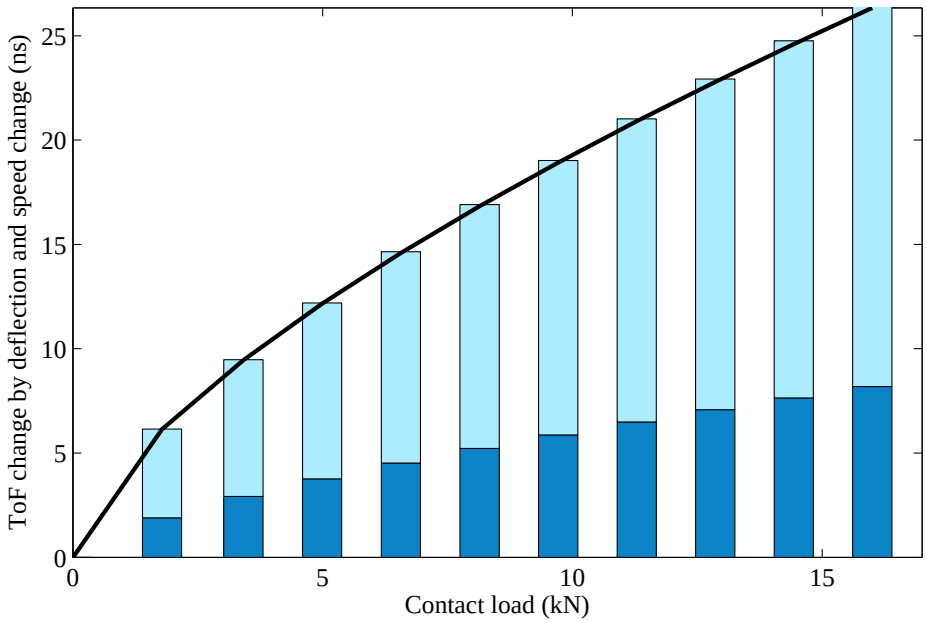


Figure 8.13. *Predicted ToF change by the deflection and the acoustoelastic effect under different contact loads.*

From figure 8.13, it can be observed that the ToF non-linearly increase with the contact load. By rearranging equation (8.3), the measured ToF change can be used to indicate the contact load, by using equation (8.4).

$$P_{max} = \pi k E^{'} \left[\frac{(c_{zz})_0 \Delta t_t}{F(1-L_{zz})} \right]^{\frac{3}{2}} \left(\frac{2\varepsilon R^{'}}{9} \right)^{\frac{1}{2}}, \tag{8.4}$$

The loads from the ultrasonic ToF changes were then compared with applied contact load in figure 8.14. Clearly, the measured contact loads corresponded well with the contact loads, the maximum difference is about 2 kN when the contact load was 10.5 kN. Some slight differences between measured and theoretical contact load are observed for contact loads between 8 and 12 kN but good agreement dominated in the whole load range.

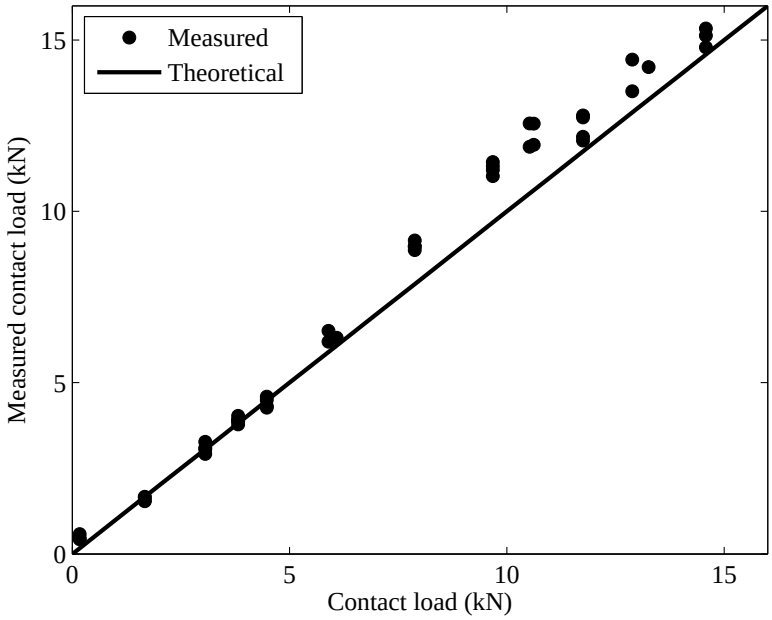


Figure 8.14. *Comparison of contact load from ultrasonic measurement and Hertz theory*

By comparing figure 8.14 with figure 7.14, a conclusion that the accuracy for load measurement on the ball bearing is better than the accuracy for the roller bearing, can be obtained. This is because the ball bearing is less likely to have edge loading or misalignment compared with the roller bearing.

8.5 Conclusion

The ultrasonic ToF technique has been applied to a ball bearing to monitor the contact load, by using a modified piezoelectric element with central frequency of 10 MHz. The transducer was directly coupled on the outer race to measure the most heavily loaded outer race-ball contact. The ToF change of ultrasound in the outer ring caused by load were measured for a number of load conditions, by comparing envelopes of the unloaded and the loaded reflected pulses. For the ball bearing, the phase change effect is significant, resulting in an apparent change in time of 23 ns which corresponds to an effect of a contact load of 13 kN. The Hilbert transform has been used to calculate signal envelopes and the phase change effect has eliminated.

In the ball bearing load measurements, it was found that the groove curvature had large effect on the ultrasonic signal, making the identification of reflected pulses more difficult

and introducing wave interference. The ToF changed by 10.8 ns more than it should be as a ball approached the sensor. It has been found that this difference is load independent. Since the principal interest of this work is the ToF change caused by load, measured ToF change for each ball passage was corrected. The corrected ToF change, caused by the surface deflection and the acoustoelastic effect alone, was then converted to contact load by using the elastic model. The loads from the measured ultrasonic ToF change were then compared to the applied load. Good agreement has been achieved, showing that this approach is able to monitor the load that transmitted from one bearing race to another by balls.

Chapter 9

Conclusions

This thesis has presented the work to assess the ability of using ToF of ultrasonic reflections to measure contact loads, which are transmitted by rolling elements in a bearing assembly. Conclusions that can be drawn from this study are detailed in this Chapter, including the development of the analytical model for the ultrasonic ToF and applications of this approach to measure load in both line contacts and elliptical contacts in bearings. Load measurements have been done on a model line contact, line contacts in a cylindrical roller bearing for wind turbine gearbox, and elliptical contacts in a deep groove ball bearing.

9.1 Significance of this work

As stated in Chapter 1, there are many problems associated with bearing failures, which significantly increase the maintenance costs, especially those resulted in unscheduled downtime. Wind turbines have demonstrated the importance of bearings as their effectiveness is determined by the main shaft and gearbox bearings performance. However, bearing failures have been reported to cause wind turbine failures. Although these bearings have a designed service life of twenty years, they frequently only survive for as less as five years. These premature failures are typically associated with highly transient loading, as a nature of wind.

Load is significant to bearing performance. Monitoring the accurate bearing load and the load history is an effective way to predict the bearing remaining service life. However, there is little work on measurements of load that are transmitted from one bearing race to another by rollers or balls. The work presented here is the only method the author is aware of to directly measure the transient load that is transmitted from one bearing ring to another by rolling elements. A novel technique has been developed on the basis of the ultrasonic ToF. Ultrasound has been used due to its non-invasive and effectiveness characteristic. This tech-

nique has been successfully applied on both a model line contact and real bearing contacts. The main achievements of this work are detailed as follows.

9.2 Modeling ToF change with contact load

After introducing the ultrasound propagation in solid components and the response at the boundaries in Chapter 3, the ultrasonic ToF change has been modeled in Chapter 4.

In tribological contacts, such as rolling element bearing contacts, the mating components are normally subjected to high load, while they contact with each other over a quite small area. As a consequence, the contact stress is rather high, in the order of Gigapascals. For example, the compression stress in the inner race of the cylindrical roller bearing was as high as 2.37 GPa when the bearing was under a radial load of 1000 kN, and the stress in the outer race of the deep groove ball bearing went up to 2.48 GPa when the bearing was under a radial load of 25 kN. As a result of the induced high stress, the contact surface deforms and the speed of sound in the compressed elements increases. The ToF of ultrasound is governed by the distance and the speed, therefore, it reduces with load. In practical ultrasonic measurements, it has been found that an apparent time shift appeared which was a result of the high contact stiffness. In this work, the load effect on the ultrasonic ToF, via the surface deformation and the increase in the speed of sound, as well as the phase change effect on the ToF, have been quantified.

9.2.1 ToF change caused by the deflection

Contacting surfaces deform when they are subjected to load. The magnitude of deformation is determined by the applied load, the geometry and property of the contact surfaces, which has been obtained from the Hertz theory using equations in Chapter 2. When the thickness of the component reduced from d_0 to d_P due to the surface elastic deflection of δ, then the ToF decreased from ToF_0 or ToF_P. In the model, the ToF change caused by the surface deflection, Δt_δ, has been elvaluated when considering the deflection alone (provided the speed was the same) and it has become to $\Delta t_\delta = 2\delta/(c_{zz})_0$. The model has shown that the ToF change was quite small, in the order of nanoseconds, since the elastic deformation in the bearing race was in the order of microns and the ultrasound speed in steel was a few thousands m/s. For example, the speed of sound in steel was 5900 m/s while the surface deflection was only 45.23 μm when the load on the roller bearing went up to 1000 kN. Hence, the ToF change caused by this deflection was only 15.3 ns.

9.2.2 ToF change caused by the acoustoelastic effect

The ultrasound speed in a bearing race is not constant, it changes with the stress state. In a compressed bearing race, the speed of sound increased from $(c_{zz})_0$ to $(c_{zz})_P$, then the ToF decreased from ToF_0 or ToF_P. The effect of speed change on ultrasonic ToF in a bearing raceway, Δt_c, has evaluated in the model by considering the acoustoelastic effect alone, i.e, the race thickness has been assumed to be constant at this point. The ToF change caused by the speed change become to $\Delta t_c = -2L_{zz}\delta/(c_{zz})_0$.

The effect of the speed change on the ToF has been predicted for different load states with known initial wave speed and bearing acoustoelastic constant. The acoustoelastic constant has been measured at - 2.24 for a typical bearing steel from experiments. The model has shown that the ToF change was small, also in the order of nanoseconds, since $L_{zz} = -2.24$ and δ was in the order of the order of microns while the speed was a few thousands m/s. For instance, the ToF only changed by 34.3 ns when the roller bearing was under load of 1000 kN. The speed of sound in this case increased by 30 m/s, about 5 thousandths of the initial speed.

9.2.3 ToF change caused by the phase change effect

In ultrasonic measurements, it has been found that the phase of the reflected wave from the contact differed from that of the incident wave and this phase change had effect on the determination of ToF. The phase change in reflected waves altered the arrival time of the reflected pulse and it appeared as an apparent time shift. This effect has been evaluated in the model.

The phase change in the reflected wave depends on the stiffness of the contact interface. For a lubricated contact, the stiffness of the lubricant layer can be determined from the film thickness which has been well evaluated in the fluid lubrication theory. The film thickness in a rolling contact is quite small, in the order of sub-microns, which gives a considerably stiff lubricant layer and thus a large phase change. The phase change approaches to $\frac{\pi}{2}$ for an infinite stiff interface and appears as an apparent ToF; it is a quarter of the signal period in the time trace, for instance, 25 ns for a 10 MHz ultrasonic pulse.

The phase change effect on the ToF change has been illustrated by figure 4.12. It can be seen that the apparent time shift is significant and cannot be neglected, especially for the low load cases. For example, when the roller bearing was under 100 kN radial load, the ToF change caused by the phase change was about 23 ns, while the ToF change caused by the deflection was 1.9 ns and that caused by the speed change was 4.3 ns. The ToF change from the phase effect was similar when the bearing load increased to 1000 kN, while the ToF

caused by the deflection and the speed change increased to 15.3 and 34.3 ns, respectively. The phase change contribution has shown a comparatively weak dependency on load.

9.3 Development of the ultrasonic measurement system for load monitoring

An ultrasonic measurement system for load monitoring has been developed in Chapter 5 and it has been used for all load tests in this work. It includes the ultrasonic apparatus, the data acquisition software and the signal processing routine. Bare piezoelectric elements have been used to transmit and receive ultrasound. They have been modified to narrow strips with respect to the considerable small contact patches in rolling bearings to provide high spatial resolution. A post data analysis software has been developed to measure the change in ToF. The Hilbert transform has been used to calculate the envelope of signals. It has found that the phase change had no effect on the envelope of the reflected pulses although it altered the arrival time of reflected signals. The time difference between the envelopes has been obtained by using the cross correlation technique. This time difference is the change in ToF caused by load alone, via the surface deflection and the acoustoelastic effect, since the phase change effect has been eliminated.

9.4 Application to a model line contact

The ToF approach has been applied to a model line contact between a steel plate and a cylindrical roller, which has been documented in Chapter 6. A modified narrow strip piezoelectric sensor has been instrumented on the external surface of the plate and load exerted upon the line contact has been measured using ToF change of ultrasonic waves in the plate. The ToF change in the plate has been measured for several load conditions. The phase change effect has been excluded by measuring the time difference between the envelope of the reference and the envelope of the loaded signal. The envelope of signal has been obtained from the reflected signal and its Hilbert transform. The measured ToF change, caused by the deflection and the acoustoelastic effect, increased from 3.67 to 11.32 ns as the load increased from 2.9 to 20.7 kN. The measured ToF change has been converted to contact load by using the relationship between load and the predicted ToF change. Good agreement has been achieved, showing that the ultrasonic ToF method is able to indicate the load in contact.

9.5 Application to a cylindrical roller bearing

The ultrasonic ToF approach has been applied to a line contact in a real roller bearing application which was used for a 2-megawatt wind turbine gearbox, as described in Chapter 7. A modified narrow strip piezoelectric sensor has been instrumented on the stationary inner race, to measure the ToF in the race when it was loaded by rolling rollers. The measurements have been done for a series of bearing loads, from 86 to 980 kN, at constant bearing speed. These loads resulted in contact loads from 4.4 to 266.8 kN, and contact pressure from 0.3 to 2.34 GPa. The corresponding maximum deflection in the inner ring was from 1 to 44.45 μm, giving the ToF change Δt_t (the deflection and the speed change effect) from 1.2 to 48.9 ns.

For each particular load, the ToF in the race has been continuously monitored and compared with that in the unloaded raceway. It has been found that the ToF change for each of the 15 rollers was slightly different, which should not be the case as the rollers had nominally the same geometry and should carry the same load when they located at the same load position. Therefore, an average value has been taken over the 15 rollers for each specific load.

The measured data has been interpreted as contact load by using the relationship between the theoretical ToF change and the applied contact load. The ultrasonic technique has slightly over-predicted at the low loads, possibly caused by some sort of bearing misalignment and the roller curvature, which have disproportion effects at low load end. The ultrasonic technique has slightly under-predicted the load for high load cases, due to the sensor average effect. But in general, the load measured from the ultrasonic reflections using the ToF approach corresponds well with the applied contact load. This technique shows promise for application to large bearings when both the contact dimension and the raceway deflection are large and hence lead to readily measurable ToF change.

9.6 Application to a ball bearing

The ultrasonic ToF approach has been applied to elliptical contacts in a deep groove ball bearing. A modified narrow strip sensor has been instrumented on the stationary outer race, to study the load effect on ultrasonic ToF in the outer ring. Experiments have been carried out for a number of radial loads at a constant bearing speed.

It has been found that the presence of the curvature of the raceway significantly reduced the reflected wave amplitude, challenging the identification of the intended pulse when the load on the bearing went up to 22.7 kN. An apparent time change has been observed as

a ball passage when the bearing has no hydraulic load. It has presented in all ultrasonic measurements, indicating it is not load dependent, possibly caused by the wave interference resulted from the conformal contact between the curved outer race and the ball. Therefore, it has been subtracted from all ultrasonic results. The slight variation in the ToF change for each ball passage has been observed for each load condition, and an average ToF change has been used instead.

After the correction, the measured ToF change has been used to indicate the load in contact, by using the relationship between the applied contact load and the theoretical ToF change. Contact load from the ultrasonic ToF approach shows excellent agreement with contact load predicted from the Hertz theory. The over prediction at low loads in the model line contact and the roller bearing application in Chapter 6 and Chapter 7, respectively, was not seen in the ball bearing application.

9.7 Future work

In this work, it has been found that the ultrasonic ToF approach has slightly over-predicted the load in at low loads, for both the model line contact and line contacts in the cylindrical roller bearing. This is possibly caused by the difference in the contact size, between the actual and the nominal contact. Further study can be done in the area using appropriate methods to measure the actual contact patch, such as using the pressure sensitive film or ultrasonic scanning with focusing transducer (small spot sizes).

Another area where future work can be done is to develop ultrasonic sensors with small dimensions so that the ultrasonic technique can be extended to small tribological contacts. Ultrasonic sensors used in this study had a width of 1 mm and a limited spatial resolution as the width of contact patch in typical elastohydrodynamic contacts is of the same order. In this study, the contact area predicted from the Hertz theory is less than the sensor effective sensing area, it is unable to measure the real contact width. If it is possible to have a sufficiently small ratio between the spot size of ultrasound beam and the contact patch, the stress profile of a contact can be built up. The stress distribution in the whole bearing can be profiled by instrumenting an ultrasonic sensor for each contact.

Further work can be done in modeling the sound propagation in components, especially those involving complex geometries. It can be quite useful to understand the interesting phenomenons observed in ultrasonic measurements if the exact wave front of the reflected wave is known.

Future work can be done in applying the ultrasonic ToF approach to industrial bearings. The ultrasonic sensors can be instrumented on the bearing element to monitor the transient

load in bearing contacts. A portable ultrasonic measurement system can be developed for engineers in industry so that they can easy collect the ultrasonic data, get the exact load transmitted between raceways by rolling elements.

Moreover, the technique can be applied to other non-steady loading elements, such as main shaft and gearbox bearings in wind turbines. The life and durability of such components depends on whether the applied load causes overload in the form of plastic deformation, impact damage or buckling. In other cases, components eventually fail by contact fatigue since the applied load is below a yield point but it highly fluctuates. This kind of tribological machine elements can be benefit from monitoring the transient load in contacts.

Further, the transient load can be used to determine the load history of the involved component, which can be helpful to study the accumulation of stress cycles. The accumulation of stress cycles can be used to determine the expected residual fatigue life of the measured subject. Main shaft and gearbox bearing failures in wind turbines always associated with excessive transient stress. Provided with the stress information from the ultrasonic ToF, engineers can predict bearings residual service life and arrange proper actions, or the extreme case to schedule proper maintenance work.

In addition, this technique developed in this work can be extended to other tribological contacts, such as gear teeth contacts, cam/tappet contacts, piston ring/liner contacts, metal roll contacts, wheel/rail contacts. Last but not least, future work can be done in the area of developing non-contact transducers and using shear waves to study the stress.

References

A. Adhikari, T. Sigurdsson, M. A. Topiwala, and J. A. Gordon. Cross-correlation of instantaneous amplitudes of field potential oscillations: A straightforward method to estimate the directionality and lag between brain areas. *Journal of Neuroscience Methods. A.*, 191: 191–200, 1998.

S. Al-Dossary, R. I. R. Hamzah, and D. Mba. Observation of changes in acoustic emission waveform for varying seeded defect sizes in a rolling element bearing. *Applied Acoustics*, 70(1):58–81, 2009.

A. M. Al-Ghamd and D. Mba. A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size. *Mechanical Systems and Signal Processing*, 20(7):1537 – 1571, 2006.

I. S. Al-Tubi and H. Long. Prediction of wind turbine gear micropitting under variable load and speed conditions using iso/tr 15144-1: 2010. *Proc. Inst. Mech. Engrs. Part C: J. Mech. Eng. Sci.*, 227:1898–1914, 2013.

R. J. Alfredson and J. Mathew. Time domain methods for monitoring the condition of rolling element bearings. *Mechanical Engineering Transactions, Institution of Engineers (Australia)*, ME 10(2):102–107, 1985.

Y. C. Angel and J. D. Achenbach. Reflection and transmission of elastic waves by a periodica array of cracks. *Journal of Applied Mechanics*, 52:33–40, 1985.

D. G. Astridge and M. D. Longfield. Capacitance measurements and oil film thickness in a large-radius disc and ring machine. *Proceedings of the Institution of Mechanical Engineers, Conference Proceedings*, 182(14):89–96, 1967.

H. L. Balderston. The detection of incipient failuer in bearings. *Material Evaluation*, 27: 121–128, 1969.

C. Barus. Isothermals, isopiestics, and isometrics relative to viscosity. *American Journal of Science*, 45:87–96, 1893.

N. M. Belajev. Bulletin of institution of ways and communications. *St Petersbug*, page 66, 1917.

J. Blitz. *Fundamentals of Ultrasonics*. BUTTERWORTHS, 1963.

R. N. Bracewell. *The Fourier Transform and its Application*. McGraw-Hill, 1978.

D. E. Bray and W. Tang. Subsurface stress evalution in steel plates and bars using the lcr ultrasonic wave. *Nuclear Engineering and Design*, 207:231–240, 2001.

L. M. Brekhovskikh. *Elastic waves in layered media*. New York ; London : Academic Press, 1960.

D. Brewe and B. Hamrock. Simplified solution for elliptical-contact deformation between two elastic solids. *ASME Trans. J. Lub. Tech*, 101(2):231–239, 1977.

P. W. Bridgman. Viscosities to 30,000 kg/cm^3. *Proc. Amer. Acad. Sci.*, 77:117–128, 1949.

J. J. Broderick, R. F. Burchill, and H. L. Clark. Design and fabrication of prototype system for early warning of impending bearing failure. *MTI Report MTI-71 TR-1 (prepared for NASA)*, 1972.

R. F. Burchill. Resonant structure techniques for bearing fault analysis. *Proceedings of 18th Meeting of Mechanical Failures Prevention Group,National Bureau of Standards NBSIR*, 73-252:21–29, 1973.

P. F. J. Burgess. Antifriction bearing fault detection using envelope detection. *Trans. Instn. Prof. Engrs., New Zealand*, 15(2):77–82, 1988.

A. Cameron and R. Gohar. Theoretical and experimental studies of the oil film in lubricated point contact. *Proceedings of the Royal Society*, 291:520–536, 1966.

K. W. Chambers. An online ferromagnetic wear debris for machinery condition and fault detection. *Wear*, 128:325–337, 1988.

J. S. Chen and K. W. Chen. Bearing load analysis and control of a motorized high speed spindle. *International Journal of Machine Tools and Manufacture*, 45:1487–1493, 2005.

A. Choudhury and N. Tandon. A theoretical model to predict vibration response of rolling bearings to distributed defects under radial load. *Transaction ASME , Journal of Vibration Acoustics*, 120(1):214–220, 1998.

A. Choudhury and N. Tandon. Application of acoustic emission technique for the detection of defects in rolling element bearing. *Tribology International*, 33:39–45, 2000.

S. G. Christopher and J. W. Simon. Physics of failure approach to wind turbine condition based maintenance. *Wind Energy*, 13:395–405, 2010.

D. Cooper and A. J. Moore. Application of the ultrathin elastohydrodynamic oil film thickness technique to the study of automotive engine oils. *Wear*, 175(1-2):93–105, 1994.

M. S. Darlow and R. H. Badgley. Applications for early detection of rolling element bearing failures using the high-frequency resonance technique. *American Society of Mechanical Engineers Paper 75-DET-46*, 1975.

M. S. Darlow, R. H. Badgley, and G. W. Hogg. Applications of high frequency resonance techniques for bearing diagnostics in helicopter gearboxes. *U.S. Army Air Mobility Research and Development Laboratory, USAAMRDL-TR-74-77*, 1974.

P. J. Dempsey, G. Kreider, and T. Fichter. Investigation of tapered roller bearing damage detection using oil debris analysis. In *Aerospace conference,2006 IEEE*, 2006.

D. Dowson and G. Higginson. Elastohydrodynamics. *Proceedings of the Institute of Mechanical Engineers*, 182:151–167, 1968.

D. Dowson and G. R. Higginson. *Elastohydrodynamic lubrication: The fundamentals of roller and gear lubrication*. Pergamon, Oxford, 1966.

B. W. Drinkwater, R. S. Dwyer-Joyce, and P. Cawley. A study of the interaction between ultrasound and a partially contacting solid-solid interface. *Proc. R. Soc. A*, 452:2613–2628, 1996.

R. S. Dwyer-Joyce. The application of ultrasonic ndt techniques in tribology. *Proceedings of the Institute of Mechanical Engineers,Part J: Journal of Engineering Tribology*, 219(5): 347–366, 2005.

R. S. Dwyer-Joyce, B. W. Drinkwater, and C. J. Donohoe. The measurement of lubricant-film thickness using ultrasound. *Royal Society London : Proceeding A*, 459:957–976, 2003.

R. S. Dwyer-Joyce, P. Harper, and B. W. Drinkwater. A method for the measurement of hydrodynamic oil films using ultrasonic reflection. *Tribology Letters*, 17:337–348, 2004.

R. S. Dwyer-Joyce, T. Reddyhoff, and B. W. Drinkwater. Operating limits for acoustic measurement of rolling bearing oil film thickness. *Tribology Transactions*, 47(3):366–375, 2006.

R. S. Dwyer-Joyce, C. Yao, J. Zhang, R. Lewis, and B. W. Drinkwater. Feasibility study for real time measurement of wheel-rail contact using an ultrasonic array. *ASME, Journal of Tribology*, 131(4):041401, 2009.

D. Dyer and R. M. Stewart. Detection of rolling element bearing damage by statistical vibration analysis. *Transaction ASME , Journal of Mechanical Design*, 100(2):229–235, 1978.

D. M. Egle and D.E. Bray. Measurement of acoustoelastic and third-order elastic constants for rail steel. *J Acoust Soc Am*, 60:741–744, 1976.

S. I. El-Sisi and G. S. A. Shawki. Measurement of oil-film thickness between disks by electrical conductivity. *Transactions of the ASME*, 82(1):12–16, 1960.

M. Elforjani and D. Mba. Monitoring the onset and propagation of natural degradation process in a slow speed rolling element bearing with acoustic emission. *Tranaction ASME, Journal of Vibration and Acoustics*, 130(4):041013, 2006.

M. Elforjani and D. Mba. Accelerated natural fault diagnosis in slow speed bearings with acoustic emission. *Engineering Fracture Mechanics*, 77(1):112 – 127, 2010.

Engineering Sciences Data Unit. *General Guid to the Choice of Journal Bearing type*. Institution of Mechanical Engineers, item 65007 edition, 1965.

P. Eschmann, L. Hasbargen, and K. Weigand. *Ball and roller bearings: their theory, design, and application*. London: K. G. Heyden, 1958.

European Standard. Non-destructive testing - characterization and verification of ultrasonic examination equipment-part 2: Probes. EN12668-2, 2001.

M. Flanagan, J. R. Jordan, and H. W. Whittington. An inductive method for estimating the composition and size of metal particle. *Measurement Science Technology*, 1:381–384, 1990.

D. Gasni. *Ultrasonic reflection for measurement of oil film thickness and contact between dissimilar materials*. University of sheffield, phd thesis edition, 2012.

E. Ginzel, A. G. Ekhlas, M. Matheson, P. Cyr, and B. Brown. Near field length compensation options. *The e-journal of Nondestructive Testing*, 19, 2014.

S. B. Glavatskih, ö Uusitalo, and D. J. Spohn. Simultaneous monitoring of oil film thickness and temperature in fluid film bearings. *Tribology International*, 34(12):853–857, 2001.

R. Gohar. Oil film thickness and rolling friction in elastohydrodynamic point contact. *Journal of Tribology*, 93(3):371–379, 1971.

O. G. Gustafsson and T. Tallian. Detection of damage in assembled rolling element bearings. *ASLE Transactions*, 5:197–209, 1962.

G. M. Hamilton. Explicit equations for the stresses beneath a sliding spherical contact. *Proc. Instn. Mech. Engrs.*, 197C:53–61, 1983.

G. M. Hamilton and L. E. Goodman. The stress field created by a circular sliding contact. *Tansaction ASME, Journal of Applied Mechanics*, 33:371–376, 1966.

B. J. Hamrock and W. J. Anderson. Rolling-element bearings. *NASA Reference Publication*, 1105, 1983.

B. J. Hamrock and D. Dowson. Isothermal elastohydrodynamic lubrication of point contacts. *Transactions of the ASME*, 99:264–276, 1977.

B. J. Hamrock and D. Dowson. *Ball bearing lubrication-the elastohydrodynamics of elliptical contacts*. Wiley-Interscience, New York, 1981.

B. J. Hamrock, B. O. Jacobson, and S. I. Bergstrom. Measurement of the density of base fluids at pressures to 2.2 gpa. *ASLE Trans.*, 30(2):196–202, 1987.

B. J. Hamrock, S. R. Schmid, and B. O. Jacobson. *Fundamentals of Fluid Film Lubrication*. Marcel Dekker, Inc., 2nd edition, 2004.

T. A. Harris. *Rolling Bearing Analysis*. John Willey-Sons, 4th edition, 2001.

D. R. Harting. Demodulated resonance analysis-a powerful incipient failure detection technique. *ISA Transactions*, 17:35–40, 1978.

H. Hertz. The contact of elastic solids. *Reine Angew. Math.*, 92:156–171, 1881.

B. Hosten. Bulk heterogeneous plane-wave propagation through viscoelastic plates and stratified media with large values of frequency-domain. *Ultrasonics*, 29:445–450, 1991.

B. Hosten, M. Deschamps, and B. R. Tittmann. Inhomogeneous wave generation and propagation in lossy anisotropic solids. application to the characterization of viscoelastic composite materials. *Acoustical Society of America*, 82(5):1763–1770, 1987.

N. E. Huang, Z. Shen, and et al. R. S. Long. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A.*, 454:903–995, 1998.

D. S. Hughes and J. L. Kelly. Second order elastic deformation of solids. *Physical Review*, 92(5):1145–1149, 1953.

T. M. Hunt. *Handbook of wear debris analysis and particle detection in liquids*. Elsevier Applied Science, 1993.

IAS motor reliability working group. Report of large motor reliability survey of industrial and commercial installations: part i. *IEEE Transactions on Industry Applications*, IA-21(4):853–864, 1985.

T. Igarashi and H. Hamada. Studies on the vibration and sound of defective rolling bearings (first report: vibration of ball bearings with one defect). *Bulletins of the JSME*, 25(204): 994–1001, 1982.

International Organization for Standards. *Rolling Bearings- Dynamic load ratings and rating life*. International standard ISO 281, 2010.

B. O. Jacobson and P. A. Vinet. A model for the influence of pressure on the bulk modulus and the influence of temperature on the solidification pressure for liquid lubricants. *ASME J. Tribol.*, 109:709–714, 1987.

K. L. Johnson. *Contact Mechanics*. Cambridge University Press, 1st paperback edition, 1987.

G. J. Johnston, R. Wayte, and H. A. Spikes. The measurement and study of very thin lubricant films in concentrated contacts. *Tribology Transaction*, 34:187–194, 1991.

W. R. Jones, R. L. Johnson, W. O. Winer, and D. M. Sanborn. Pressure-viscosity measurements for several lubricants to 5.5×10^8 newtons per square meter and 149 °C(300°F. *ASLE Trans.*, 18(40):249–262, 1975.

T. Kamei and A. Serizawa. Measurement of 2-dimensional local instantaneous liquid film thickness around simulated nuclear fuel rod by ultrasonic transmission technique. *Nuclear Engineering and Design*, 184(2-3):349–362, 1998.

S. Kasolang, D. I. Ahmed, R. S. Dwyer-Joyce, and B. F. Yousif. Performance analysis of journal bearings using ultrasonic reflection. *Tribology International*, 64:78 – 84, 2013.

P. Y. Kim. A review of rolling element bearing health monitoring (ii): preliminary test results on current technologies. *Proceedings of Machinery Vibration Monitoring and Analysis Meeting, Vibration Institute, New Orleans, LA, 26-28 June*, pages 127–137, 1984.

V. K. Kinra, P. T. Jaminet, C. Zhu, and V. R. Iyer. Simultaneous measurement of the acoustical properties of a thin-layered medium: the inverse problem. *J. Acoust. Soc. Am.*, 95: 3059–3074, 1994.

M. T. Kirk. The hydrodynamic lubrication of perspex. *Nature*, 194:965–966, 1962.

C. H. Knapp and G. C. Carter. The generalized correlation method for estimation of time delay. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 24:320–327, 1976.

J. Kräutkramer and H. Kräutkramer. *Ultrasonic testing of materials*. Springer-Verlag, New York, 1977.

O. P. Lelikov. *Design of machine elements. In: K.-H. Grote and E. K. Antonsson eds. Springer Handbook of Mechanical Engineering*, volume 10. Springer Berlin Heidelberg, 2009.

G. Lundberg and H. Sjövall. *Stress and Deformation in Elastic Contacts*. Institute of Theory of Elasticity and Strength of Materials, Chalmers Inst. Tech., Gothenburg, pub.4 edition, 1958.

G. Y. Luo, D. Osypiw, and M. Irle. On-line vibration analysis with fast continuous wavelet algorithm for condition monitoring of bearing. *Journal of Vibration and Control*, 9:931–947, 2003.

M. B. Marshall, R. Lewis, R. S. Dwyer-Joyce, U. Olofsson, and S.Björklund. Experimental characterization of wheel-rail contact patch evolution. *Transactions of the ASME: Journal of Tribology*, 128(3):493–504, 2006.

K. F. Martin and P. Thorpe. Normalised spectra in monitoring of rolling bearing elements. *Wear*, 159:153–160, 1992.

J. Mathew and R. J. Alfredson. The condition monitoring of rolling element bearings using vibration analysis. *Transaction ASME, Journal of Vibration, Acoustics, Stress, and Reliability in Design*, 106:447–453, 1984.

J. R. Mathews. *Acoustic emission*. Gordon and Breach Science Publishers Inc., New York, 1983.

E. McEwen. Stresses in elastic cylinders in contact along a generatrix. *Philosophical Magazine*, 40:454–459, 1949.

P. D. McFadden and J. D. Smith. Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of Sound and Vibration*, 96(1):69 – 82, 1984a.

P. D. McFadden and J. D. Smith. Vibration monitoring of rolling element bearings by the high-frequency resonance technique-a review. *Tribology International*, 17(1):3–10, 1984b.

L. D. Meyer, F. F. Ahlgren, and B. Weichbrodt. An analytic model for ball bearing vibrations to predict vibration response to distributed defects. *Transaction ASME , Journal of Mechanical Design*, 102:205–210, 1980.

J. L. Miller and D. Kitaljevich. In-line oil debris monitor for aircraft engine condition assessment. *2000 IEEE Aerosp Conference Proceeding*, 6:49–56, 2000.

K. Mori, N. Kasashima, T. Yoshioka, and Y. Ueno. Prediction of spalling on a ball bearing by applying the discrete wavelet transform to vibration signals. *Wear*, 195:162–168, 1996.

F. D. Murnaghan. The mompressibility of media under extreme pessures. *Proceedings of the National Academy of Sciences of the United States of America*, 30(9):244–247, 1944.

M. Nässelqvist, R. Gustavsson, and J. O. Aidanpää. Bearing load measurement in a hydropower unit using strain gauges installed inside pivot pin. *Experimental Mechanics*, 52:361–369, 2012.

C. C. Osuagwu and D. W. Thomas. Effect of inter - modulation and quasi-periodic instability in the diagnosis of rolling element incipient defect. *Trans. ASME, J. Mech. Design*, 104: 296–302, 1982.

A. Palmgren. *Ball and Roller Bearing Engineering*. Burbank, Philadelphia, 3rd edition, 1959.

Y. H. Pao, R. R. Gajewski, and A. N. Ceranoglu. Acoustic emission and transient waves in an elastic plate. *Journal of Acoustical Society of America*, 65(96):96–102, 1979.

E. P. Papadakis. *Ultrasonic instruments and devices*. Academic Press, 1999.

T. Pialucha and P. Cawley. The detection of thin embedded layers using normal incidence ultrasound. *Ultrasonics*, 32(6):431 – 440, 1994.

T. Pialucha, C. H. Guyott, and P. Cawley. Amplitude spectrum method for the measurement of phase velocity. *Ultrasonics*, 27:270–279, 1989.

T. P. Pialucha. *Reflection coefficient from interface layers in ndt of adhesive joints*. Imperial Colledge, phd thesis edition, 1992.

B. Pottier, L. Rasolofondraibe, P. Marconnet, and D. Nuzillard. Static stress measurement based on capacitive probes integrated inside rolling rings. *IEEE SENSORS JOURNAL*, 11(11):2919–2925, 2011.

S. Prabhakar, A. R. Mohanty, and A. S. Sekha. Application of discrete wavelet transform for detection of ball bearing race faults. *Tribology International*, 35(12):793–800, 2002.

R. Prabhu. Rolling bearing diagnostics. *Proceedings of the Indo-US Symposium on Emerging Trends in Vibration and Noise Engineering, New Delhi*, pages 311–320, 1996.

H. Prasad. The effect of cage and roller slip on the measured defect frequency response of rolling element bearings. *ASLE Transactions*, 30(3):360–367, 1987.

H. Qiu, J. L., J. L., and G. Yu. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of Sound and Vibration*, 289(4-5):1066–1090, 2006.

L. Rasolofondraibe, B. Pottier, P. Marconnet, and X. Chiementin. Capacitive sensors device for measuring loads on bearings. *IEEE SENSORS JOURNAL*, 12(6):2186–2191, 2012.

T. Reddyhoff. *Ultrasonic measurement techniques for lubricant films*. University of sheffield, phd thesis edition, 2007.

T. Reddyhoff, S. Kasolang, R. S. Dwyer-Joyce, and B. W. Drinkwater. The phase shift of an ultrasonic pulse at an oil layer and determination of film thickness. *Proceedings of the Institute of Mechanical Engineers,Part J: Journal of Engineering Tribology*, 219(5): 387–399, 2005.

L. Reti. Leonard on bearings and gears. *Seientific American*, 224(2):101–110, 1971.

C. J. A. Roelands. *Correlation aspects of the viscosity-temperature-pressure relationship of lubricanting oils*. V. R. B. Druk, Groingen, Netherlands, 1966.

L. M. Rogers. The application of vibration analysis and acoustic emission source location to on-line condition monitoring of anti-friction bearings. *Tribology International*, 12(2): 51–59, 1979.

B. J. Roylance. Ferrography-then and now. *Tribology International*, 38(10):857 – 862, 2005.

B. J. Roylance and T. M. Hunt. *The wear debris analysis handbook*. Coxmoor Publishing Company, Oxford, UK, 1999.

R. Rubini and U. Meneghetti. Application of the envelope and wavelet transform analyses for the diagnosis of incipient faults in ball bearings. *Mechanical Systems and Signal Processing*, 15:287–302, 2001.

A. Sackfield and D. A. Hills. Some useful results in the classical hertz contact problem. *Journal of Strain Analysis*, 18:101–108, 1983.

M. M. A. Safa, J. C. Anderson, and J. A. Leather. Transducers for pressure, temperature and oil film thickness measurement in bearings. *Sensors and Actuators*, 3:119–128, 1982-1983.

M. Schoenberg. Elastic wave behaviour across linear slip interfaces. *J Acoust Soc Am*, 68(5):1516–1521, 1980.

W. W. Seifert and V. C. Westcott. A method for the study of wear particles in lubricating oil. *Wear*, 21(1):27 – 42, 1972.

J. Shiroishi, Y. Li, S. Liang, T. Kurfess, and S. Danyluk. Bearing condition diagnostics via vibration and acoustic emission measurements. *Mechanical System and Signal Processing*, 11(5):693–705, 1997.

SKF. Ball bearings. http://www.skf.com/group/products/bearings-units-housings/ball-bearings, acessed on 28-02-2015.

R. T. Smith. Stress-induced anisotropy in solids- the acousto-elastic effect. *Ultrasonics*, 1(3):135–147, 1963.

R. Stribeck. Characteristics of plain and roller bearings. *Zeit. Ver. Deut. Ing.*, 46:1341–1348, 1902.

Y. T. Su and S. J. Lin. On initial fault detection of a tapered roller bearing: frequency domain analysis. *Journal of Sound and Vibration*, 155:75–84, 1992.

C. S. Sunnersjö. Varying compliance vibrations of rolling bearings. *Journal of Sound and Vibration*, 58(3):363–373, 1978.

T. E. Tallian and O. G. Gustafsson. Progress in rolling bearing vibration research and control. *ASLE Transaction*, 8(3):195–207, 1965.

C. C. Tan. Application of acoustic emission to the detection of bearing failures. *The Institution of Engineers Australia,Tribology conference, Brisbane*, pages 110–114, 1990.

E. Tanala, G. Bourse, M. Fremiot, and J. F. De Belleval. Determination of near surface residual stress on welded joints using ultrasonic methods. *NDT I& E International*, 28(2): 83–88, 1995.

N. Tandon and A. Choudhury. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology International*, 32(8):469 – 480, 1999.

N. Tandon and B. C. Nakra. Defect detection of rolling element bearings by acoustic emission method. *Journal of Acousti Emission*, 9(1):25–28, 1990.

N. Tandon and B. C. Nakra. Vibration and acoustic monitoring techniques for the detection of defects in rolling element bearings-a review. *Shock and Vibration Digest*, 24(3):3–11, 1992.

N. Tandon and B. C. Nakra. Detection of defects in rolling element bearings by vibration monitoring. *J. Instn. Engrs. (India) - Mech. Eng. Div.*, 73:271–282, 1993.

W. Tang and D. E. Bray. Stress and yielding studies using critically refracted longitudinal waves. *NDE Engineering Codes and Standards and Material Characterization, Proceedings 1996 ASME Pressure Vessels Piping Conference*, Montreal, PQ, 1996.

H. G. Tattersall. The ultrasonic pulse-echo technique as applied to adhesion testing. *Journal of Physics D:Applied Physics*, 6(7):819–832, 1973.

J. I. Taylor. Identification of bearing defects by spectral analysis. *Trans. ASME, J. Mech. Design*, 102:199–204, 1980.

H. R. Thomas and V. A. Hoersch. Stress due to the pressure of one elastic solid upon another. *University of Illinois, Engineering Experimental Station*, Bulletin No. 212:66–99, 1930.

Timken. Roller bearings. http://www.timken.com/en-us/products/bearings/product-list/roller/Pages/default.aspx, acessed on 28-02-2015.

Vishay. M-bond 610: Strain gage adhesive. 11013 ed., 2010.

W. Wang and W. Zhang. A model to predict the residual life of aircraft engines based upon oil analysis data. *Naval Research Logistics*, 52(3):276 – 284, 2005.

J. C. Wheals, P. Guern, and R. S. Dwyer-Joyce et al. Ricardo multilife bearing programme for increased reliability of offshore wind turbines. *European Wind Energy Conference*, 2011.

T. A. Whittingham. Broadband transducers. *European Radiology*, 9, 1999.

S. Witting, A. Glahn, and J. Himmelsbach. Influence of high rotational speeds on heat transfer and oil film thickness in aero-engine bearing chambers. *Transaction of the ASME*, 116(2):395–401, 1994.

H. Yang, J. Mathew, and L. Ma. Fault diagnosis of rolling element bearings using basis pursuit. *Mechanical Systems and Signal Processing*, 19:341–356, 2005.

T. Yoshioka and T. Fujiwara. New acoustic emission source locating system for the study of rolling contact fatigue. *Wear*, 81(1):183–186, 1982.

J. Zhang, B. W. Drinkwater, and R. S. Dwyer-Joyce. Acoustic measurement of lubricant-film thickness distribution in ball bearings. *Acoustical Society of America*, 119:863–871, 2006.

Appendix A

Appendix: Associated publications

Chen, W., Mill, R., Dwyer-Joyce, R. " Direct load monitoring in rolling bearing contacts using ultrasonic time of flight", Accepted by Proceedings of Royal Society A: Mathematical, Physical & Engineering Sciences.

Chen, W., Dwyer-Joyce, R. " Active piezoelectric-sensor based condition monitoring system for rolling element bearings ", 5th World Tribology Congress, WTC 2013. 2: 170-1793.

Chen, W., Howard, T., Dwyer-Joyce, R. " Load profiling of a rolling element bearing contact with ultrasound ", accepted by STLE Tribology Frontiers Conference, October 25-27 2015, Denver, USA.

Chen, W., Connolley, T., Collins, D., Reinhar, C., Barhli, S., Drakopoulos, M., Mills, R., Marshall, M., Dwyer-Joyce, R., Mostafavi, M. " Dynamic strain measurement in a ball bearing", accepted by BSSM(British Society for Strain Measurement) 10th International conference on Experimental Mechanics, September 1-3 2015, Edinburgh, UK.