The
University
Of
Sheffield.

Department
Of
Mechanical
Engineering.

# Development of a Novel Bearing Concept for Improved Wind Turbine Gearbox Reliability

Thomas Peter Howard

January 2016

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

# Abstract

The wind industry has experienced rapid growth in recent years as a result of ever increasing concerns over fossil fuel power generation and the associated impact this has upon climate change. Wind turbine technology has however displayed relatively poor reliability from the outset, largely due to the highly variable nature of the wind, which subsequently places system components under extremely harsh loading conditions. Such reliability issues are becoming increasingly problematic for the wind turbine operator, as wind farms – and indeed the turbines themselves – continue to be up-scaled thus increasing the complexity and cost of maintenance, particularly in off-shore environments. Gearbox failures have been found to be responsible for a large proportion of wind turbine downtime, and this is very often linked to the failure of bearings at certain locations within the gearbox. One such critical location is the epicyclic stage, where planetary support bearings are often found to exhibit damage to a localised portion of their inner raceways, corresponding to the location of the applied load.

A concept has been proposed by Ricardo UK Ltd. to extend the life of such bearings by periodically rotating the normally static inner raceway so as to avoid the build of damage to one localised region. The concept has been termed the "MultiLife$^{TM}$" mechanism and a key aim of this thesis was to establish proof-of-concept for this system. Initial analytical work was performed, through which it was identified that a five-fold enhancement to bearing life would be theoretically achievable. The concept was subsequently validated experimentally through testing on a bespoke test platform.

In addition to this, the ultrasound technique has been explored as a means to provide early warning of bearing failure and thus support the operation of the MultiLife system. This technique has previously been proven as a method to measure the thickness of the oil films that form within bearing

contacts. It was identified for this study that oil films would be too small to measure due to the low viscosity lubricants and high bearing loads utilised to accelerate bearing failures. Nonetheless, a general reduction in the amount of ultrasonic energy reflected from the rolling contact was observed during bearing failure, which was linked to a breakdown of the lubricating oil layer due to degradation of the rolling surface. In this case the ultrasound technique did not provide any advanced warning of bearing failure over more classical condition monitoring techniques; however, it was identified that the use of higher resolution ultrasound sensor arrays would enhance the capabilities considerably.

Currently, very few monitoring techniques are applied to wind turbine gearbox bearings, and as a further part of this work a condition monitoring system has been installed within an operational $600kW$ wind turbine. A high speed shaft bearing has been instrumented with a variety of sensors, including ultrasound instrumentation, to assess the potential of such a system in providing early warning of impending gearbox failures.

# Acknowledgements

Above all, thank you to Matt Marshall for his guidance and support, and for always having faith in me. Thank you to Rob Dwyer-Joyce for also being so supportive and encouraging throughout. I am extremely grateful for all the fantastic opportunities and experiences that I have had whilst working for you both, thank you for allowing me to achieve so much.

Thank you to all the folks at Ricardo whom I've had the pleasure of working with over the years. Sincerest thanks to Prof Jon Wheals for involving us in his MultiLife brainchild from the start. Thank you also to Hugh Freeman for all the technical support, and always making himself available to help out in times of trouble!

Thanks to Dave Butcher for the day-to-day technical support and for always being so willing to help out. Thank you for also investing so much time in teaching me to use a lathe and milling machine and actually make things myself. It's been extremely fun, and has made me a better engineer overall.

Jack Naumann has worked alongside me for the majority of this project. Thank you Jack for all the support and friendship, and for putting up with my often pedantic and no doubt frustrating ways of working!

It has been great to be part of the tribology group and have the friendship and support of so many fantastic people. I would like to extend particular thanks to Robin Mills, Giuseppe Tronci, Wenqu Chen and Ramon Fuentes for providing key expertise at various stages throughout this PhD.

Thanks to Phil Harper and his team at Tribosonics for all the advice and ultrasound knowledge.

Finally, thank you to Mum and Dad for always enabling me to pursue my goals and ambitions. And Laura, thank you for your love, support and patience throughout this, I couldn't have done it without you.

# Contents

# Chapter 1

# INTRODUCTION

In this chapter the Wind Turbine system is introduced and evidence is presented which points to increased adoption rates of such systems in the years to come. Reliability issues are highlighted, and bearing failures exposed as a particular root-cause, which has led to the development of a new bearing concept to alleviate such issues. The scope of this thesis is subsequently introduced, which focusses around the development of this proposed bearing concept.

## 1.1 The Growth of Wind Turbines

Worldwide investment in renewable energy has grown considerably in recent years in response to increasing concerns over future energy security and the clear implications that conventional fossil fuel power generation has upon global warming and climate change. The UK alone, as part of the *"EU Renewables Directive (2009/28/EC),"* is committed to a target of 15% of total energy consumption to be produced from renewable sources by 2020; estimated to represent around $216 - 225TWh$ in energy for heat, transport and electricity. This represents an increase of almost 5.5 times production levels at the beginning of 2008 [1], and similar targets exist for all European member states, as well as many other industrial nations worldwide.

Figure 1.1 shows the number of wind turbines installed within the EU as of 2011 in terms of total energy capacity, and also provides projections for further installations up to the 2020 "2009/28/EC" commitment deadline. Overlaid on the figure is the total EU electricity consumption provided by such wind turbine installations where, again, figures beyond 2011

represent an estimate based on average wind energy production at the time of publication. By 2020, production of $581 TWh$ is forecast, representing an almost threefold increase on 2011 production levels, and around 15.1% of total EU electricity consumption; thus contributing almost half of the total "energy from renewables" EU commitment [2].
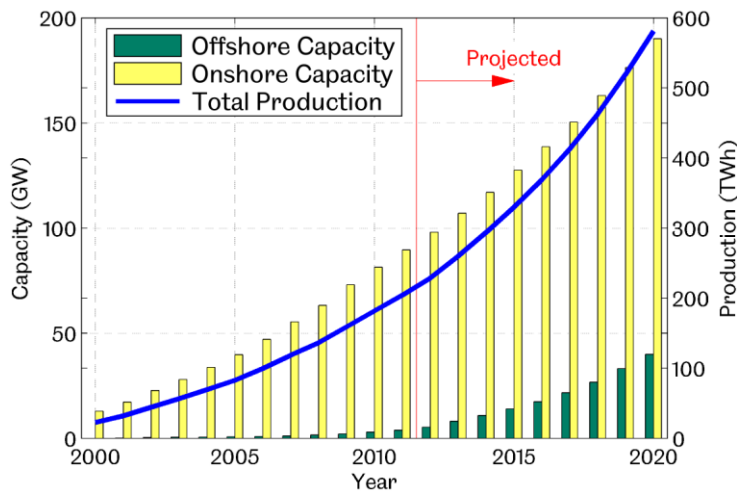


Figure 1.1: Total offshore and onshore installed wind turbine capacity and associated electricity production within the European Union as of 2011, and projected increases to 2020 (reproduced from [2])

Based on the above, it is clear that the wind industry will be relied upon heavily in not only the achievement of short term renewable energy targets, but also as a means to develop much longer term sustainable societies. Indeed, wind energy appears to be particularly promising, not only due to the abundance of wind as a resource, but also because of the relative maturity of the technology which has led to a reduction in the costs of installation over the last 20-30 years, thus influencing the continually rising uptake as suggested by Figure 1.1.

A limitation however is the relatively low power density of wind energy, meaning that large areas of space must be occupied by turbines in order to extract sufficient energy to meet demands. This has led to an upscaling in the technology of late to exploit higher wind speeds at greater altitudes and increase swept area through use of larger turbine blades, thus increasing power density. Consequently however, all operational loads upon the system become amplified [3]. Further to this, as observed in Figure 1.1, there is an increasing shift in focus towards offshore applications, where not only are wind speeds generally higher and more consistent, but also less spatial limitations exist, and thus larger wind farms can be installed for greater power production. This adds further

2

complexities however, due to the harsher environments within which the turbines must operate, and thus additional considerations must be made to factors such as corrosion, harsher temperatures and added sealing requirements.

As such, reliability issues are becoming much more prevalent, which results in increased costs to the industry. This is particularly significant for offshore applications, where the cost for maintenance teams to access sites can be high, and environmental conditions can often limit access for significant periods of time. As such, the downtime that results from a turbine failure can have further implications upon the ultimate cost of the technology, as operators suffer loss in revenue during such periods. General reliability issues, both on and offshore, result in lack of confidence, which is not only bad for public perception but also drives up insurance premiums, thus adding further costs for the operator.

It has been recently estimated that the cost of offshore wind energy in the UK is around £140/$MWh$, and this needs to be reduced to at least £100/$MWh$ by 2020 in order to maintain economic viability and meet UK renewables targets [4]. Operations and maintenance (O&M) costs have been found to account for 18% of the total cost of energy for offshore wind technology and, therefore, there is considerable scope to make significant cost reductions through resolving reliability issues alone. New technologies and advanced condition monitoring techniques will play a vital role in achieving this; enabling more effective, site-wide maintenance scheduling and a general reduction in total turbine downtime.

## 1.2 Wind Turbine Gearbox Reliability

Figure 1.2 presents data from a sample of around 650 wind turbines in Sweden recorded within the period 2000-2004 [5]. This provides a breakdown of failures and subsequent downtime per component recorded across the five year period as a percentage of the total system wide failure rates and downtime hours recorded respectively. The Wind Turbine Gearbox (WTGB) is particularly prominent here since, despite moderate failure rates, the downtime ensuing from such failures was found to be extremely high. Indeed, the average downtime resulting from a single gearbox failure was found to be equal to almost 257 hours, a significantly longer turnaround than that resulting from a control system failure, where total downtime was also similarly high, but an average single failure was found to have typically been resolved within around 185 hours by comparison.
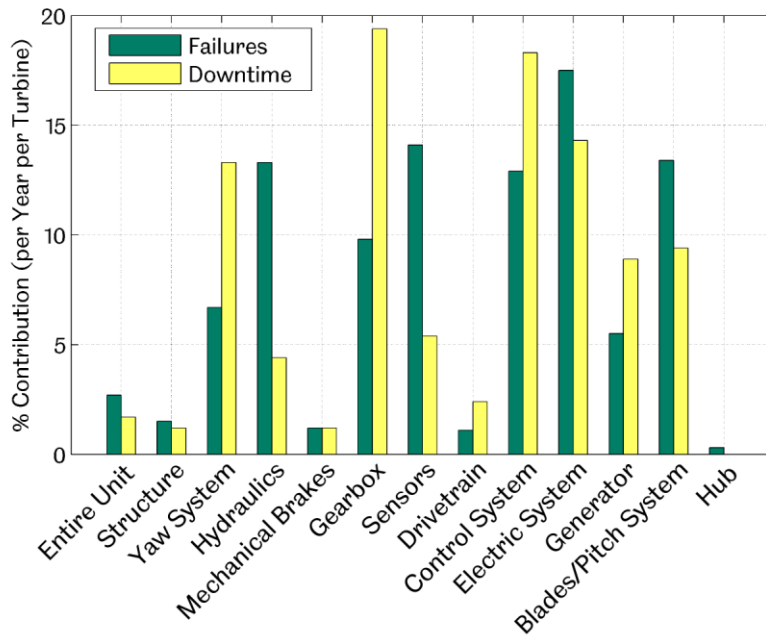
Figure 1.2: Distribution of failures and associated downtime per component recorded from a sample of Swedish wind turbines from 2000-2004 (adapted from [5])

As such, the combination of these two factors makes the gearbox particularly critical with regards to reliability and associated costs. This trend does not appear to have diminished in recent years; in the United States alone for example, 35.1% of insurance claims made within the wind industry in 2012 related to problems resulting from gearbox damage [6]. It is suggested in fact that gearbox replacements are typically required at least two to three times within the 20 year design life of a turbine, which can cost in excess of £200,000 per unit for the larger design variants [7].

## 1.2.1 Architecture

"Direct drive" or "gearless" wind turbine systems continue to be extensively explored in an attempt to mitigate the reliability issues highlighted above. Presently however, such systems are considerably more expensive and require a permanent magnet rotor of significant diameter in order to generate comparable electrical power at the very low rotor input speeds at which the wind turbine must operate (typically $15 - 20rpm$). Consequently, nacelle hardware is larger, heavier and generators require large amounts of costly rare earth magnetic materials. Commercial viability and sustainability of such systems therefore becomes questionable, and subsequently the technology has so far only been proven for the smaller, sub-megawatt turbine variants. There is currently

little knowledge of the capabilities of such technologies at a larger scale for offshore applications. As such, geared systems prevail, and high ratio gearboxes are employed to convert rotor input speeds in the order of $15 - 20 rpm$ into speeds of around $1500 rpm$, which can then be directly fed to a high speed generator so as to enable electrical power to be produced at grid compatible frequency $(50 Hz)$.

In order to achieve such high gear ratios, whilst maintaining low overall package size and weight, a gearbox with at least one epicyclic stage is generally employed, the typical architecture of which is shown in Figure 1.3. Rotor input rotation drives a planet carrier, which commonly holds three or four planet gears. Planet gears mesh with a stationary annulus causing planetary rotation, which in turn drives a central sun gear, subsequently transmitting rotation to either a secondary epicyclic stage, or directly to a series of parallel stages. A high speed output shaft then couples directly to the electrical generator.
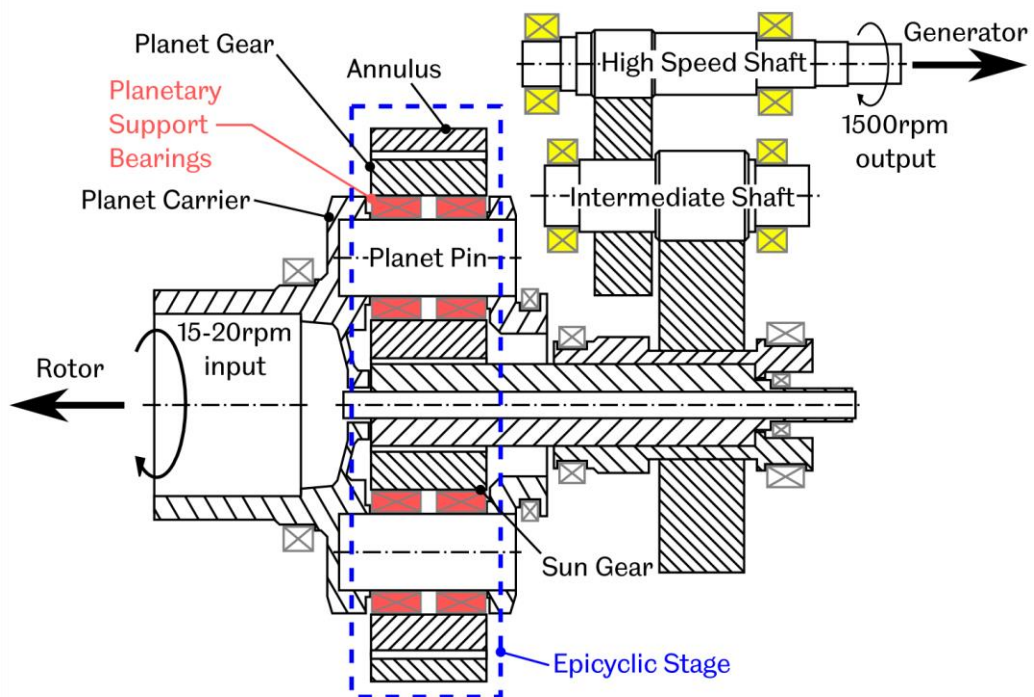
Figure 1.3: Cross section of a typical single stage planetary wind turbine gearbox illustrating critical locations for bearing failures (adapted from [8])

## 1.2.2 Root-Cause Failure Analysis

It has been observed that WTGB failures tend to ultimately arise due to the failure of bearings. As a result, bearing clearances are increased and

wear debris is introduced into the gearbox, which subsequently results in both misalignment issues and surface wear of gears themselves [9, 10]. Previous analysis has found that failures most frequently develop from planetary support bearings as well as intermediate and high speed shaft locating bearings [8, 10]. The location of the planetary support bearings are highlighted in red in Figure 1.3, and provide the key focus for this study. Here, cylindrical roller bearings are employed, where (as illustrated in Figure 1.4) inner raceways are static and mount directly to a planet pin, which rigidly couples to the planet carrier frame as shown in Figure 1.3. Subsequently, the bearing outer raceways rotate to permit rotation of the planet gears.
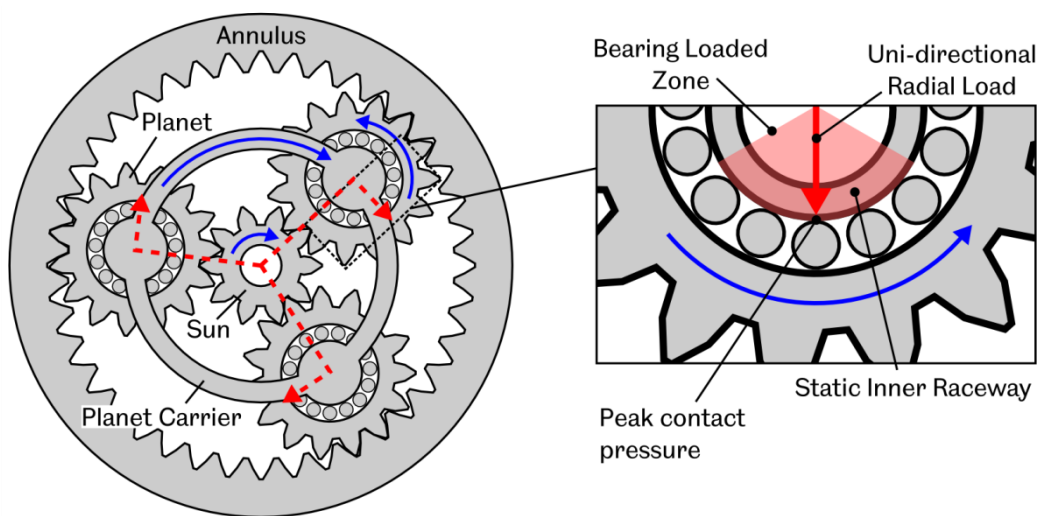


Figure 1.4: Illustration of the epicyclic stage of the WTGB and the loads induced upon the planetary support bearings from torque applied to the planet carrier

Figure 1.5 shows an ex-service planetary bearing inner raceway taken from a $600kW$ wind turbine gearbox. The failure mode here is typical of a planetary support bearing failure; where excessive spalling damage is observed across a localised portion of the inner raceway [11]. The localised nature of damage here is thought to be a consequence of such bearings having static inner raceways and applied load (resulting from the torque applied through the planet carrier) being purely radial and unidirectional (as shown in Figure 1.4). This is problematic, since the same portion of raceway material will always be under load, as dictated by the bearing loaded zone, and this localised raceway segment will therefore rapidly accrue a high number of rolling contact fatigue cycles as rollers repeatedly induce stresses in the material during bearing rotation. Effects are amplified here since contact between the inner raceway and rolling element is counter-formal and therefore stresses will be much higher

than those induced between the opposing outer raceway rolling element interface, despite both experiencing the same applied load.
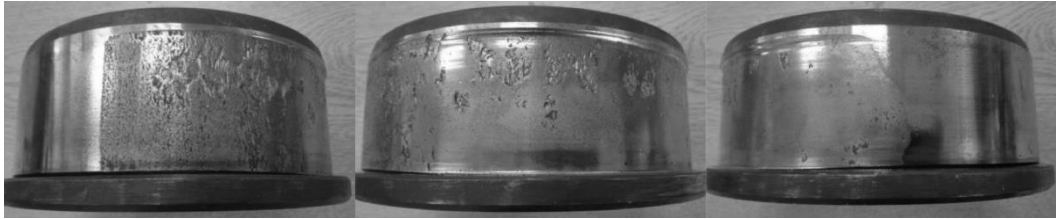


Figure 1.5: Planetary support bearing inner raceway from a 600kW WTGB illustrating localised spalling damage across the bearing loaded zone (courtesy of T Bruce)

However, whilst not being an ideal scenario, this mode of operation is not unusual for a bearing, and even if high radial loading is anticipated, a reasonable operational life would be expected prior to the occurrence of failure. Indeed, calculation of bearing fatigue life, through application of the widely accepted $L_{10}$ life rating procedure [12], has typically been found to significantly over predict the expected life of such bearings, suggesting that the failure mode is not a result of classical rolling contact fatigue [8, 13] and that load regimes are in fact vastly more complex than anticipated [14].

WTGB planetary bearing failures such as that observed in Figure 1.5 have been widely linked to an advanced rolling contact fatigue mechanism resulting in the formation of White Etching Cracks (WECs) within bearing materials [15]. The subsequent cyclic loading that then takes place during normal operation causes such cracks to propagate rapidly, resulting in a rolling surface failure via White Structure Flaking. Sub-surface initiation appears to be dominant for the case of planetary stage bearing failures specifically [16], and the initiation of WECs tends to be due to the presence of manganese sulphide (MnS) inclusions within the bearing steel [17].

Overloading is thought to be a key driver of this, arising from very sudden impulses through the drivetrain, generating torque spikes through the planet carrier and effectively resulting in hammering or impact to planetary bearing rolling surfaces. It has been identified that such overload events may arise in a number of ways and can take place very frequently in a WTGB. The transient nature of the wind alone will induce such effects; a wind gust for instance causes the rotor to accelerate dramatically, thus directly inducing a torque spike. Engagement and disengagement to the electricity grid is known to generate torque reversals through the gearbox, an occurrence that is understood to occur many thousands of times per year [13, 18] as network operators attempt to

manage electrical supply and demand. Normal and emergency stops have a similar effect, occurring through the combined application of a brake to the high speed shaft and feathering of blades [19]. Normal stops alone will take place very often as turbines cut in and out if wind speeds are either too high or low for optimum power generation.

However, while the prevalence of such events may well be high, there is little known about the magnitude at which such loads are induced, and this is thought to also be highly variable. Some evidence has suggested that, throughout the service life of a particular WTGB design variant, transient loading beyond the rated load of the gearbox may occur as little as 200 times, and the most extreme load cases may account for as little as 10 individual instances [16]. This therefore suggests that failures can be induced rapidly as a result of a very low number of impact cycles.

## 1.3   The MultiLife Concept

A solution has been proposed by Ricardo UK Ltd. to extend the life of such planetary WTGB bearings, and indeed any bearing whose inner raceway is static and which experiences a constant uni-directional radial load. The concept, termed the *"MultiLife$^{TM}$ Bearing,"* is a mechanism that adopts a modified inner raceway for such bearings and utilises an actuation mechanism to allow it to be periodically rotated or "indexed." As such, bearing life is extended since damage can be distributed around the entire raceway circumference, thus avoiding the significant build-up of wear to the localised loaded region as discussed above.

Figure 1.6 illustrates the MultiLife mechanism, which consists of a series of pistons within individual cylinders arranged on a ring which mounts axially to a planetary bearing inner raceway, modified with an axial timing mark to engage with a mating key on the actuating mechanism. Indexation is achieved through hydraulic actuation of the pistons, causing rotation of the MultiLife casing as well as the bearing inner raceway, relative to the stationary inner shaft. A sprag clutch on the inner bore of the actuating mechanism engages with the stationary shaft to restrict reverse slip between the actuation mechanism and shaft, thus ensuring that raceway indexation is achieved. Much effort was made in the development of the mechanism to allow for easy integration within current WTGB designs. As such, the hydraulic power required for actuation could feasibly be obtained from the existing gearbox hydraulic system. Furthermore, in the event of mechanical failure of the device, bearing raceways will

automatically be locked in a static position, and the system will thus default to that of a standard "non-indexing" bearing.
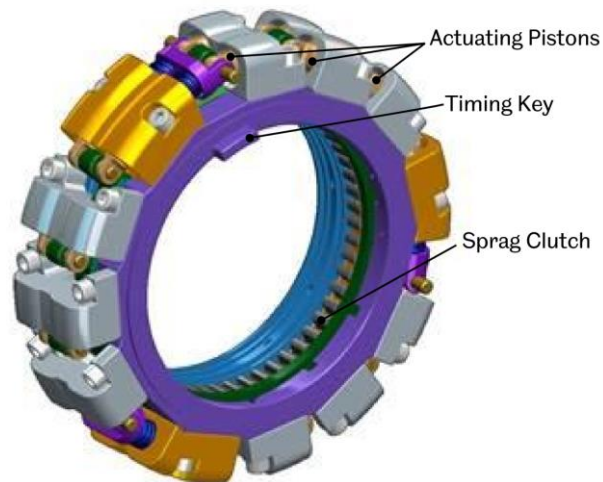


Figure 1.6: The MultiLife Mechanism (courtesy of Ricardo UK Ltd.)

## 1.4 About this Thesis

The work presented within this thesis was performed in collaboration with Ricardo UK Ltd. The key aims were to provide proof-of-concept for the MultiLife mechanism, and develop advanced sensor systems for real-time measurement of bearing condition to further aid and optimise operation of the MultiLife system. The sensor system utilised the ultrasound technique for oil film monitoring, as well as acoustic emissions (AE) measurements and signal processing techniques for detection of crack formation and propagation. Development of the latter was conducted as part of a parallel PhD project [20], and is therefore not discussed in detail within this work.

The project was initiated and originally funded through the *"Northern Wind Innovation Programme"* (NWIP), developed to promote innovation and technological enhancement to the Offshore Wind Industry. The work has since gained further funding from the *"Department of Energy and Climate Change"* (DECC) to focus on deployment of the MultiLife technology and sensor systems within a Scottish Power Renewables Wind Turbine on Barnesmore Windfarm in the Republic of Ireland.

### 1.4.1 Objectives

As discussed briefly above, the aim of this project was to provide proof of concept for the MultiLife mechanism and gain an understanding into the potential of such a system in enhancing bearing life. In addition, this project also aimed to develop ultrasonic measurement techniques to provide real-time indication of bearing condition and enable early warning of impending bearing failures. Objectives were as follows:

1. Develop a bearing life model for the analysis of indexing bearings to understand the theoretical life enhancements that are achievable and the optimum operational regimes for a MultiLife Bearing
2. Instrument a bespoke test platform developed by Ricardo UK Ltd, and identify and explore a series of bearing life cycle tests
3. Validate the MultiLife thesis through simulation of representative bearing failures and assessment of the reduction to propagation rates through raceway indexation
4. Analyse ultrasonic measurements with the aim to extract features pertaining to bearing operation and degradation
5. Industrialise sensor and data acquisition systems for application to an operational wind turbine

## 1.4.2 Key Contributions to Knowledge Arising from this Work

Through the work contained within this thesis a number of novel techniques have been established which provide original contributions to knowledge.

A bearing life model has been developed which considers a bearing raceway as a series of circumferential elements, subsequently enabling the life of each discrete element to be assessed. As such, the standard ISO bearing life rating principles have been built upon to also permit assessment of failure *location* based on the orientation of applied load.

Through the above an analytical model has also been devised to consider the life rating of a bearing whose nominally stationary inner raceway is periodically indexed; as per the MultiLife principle. Consequently the model has been utilised to consider a variety of "Indexing Strategies" and identify optimum zones of operation based on the distribution of fatigue across the entire circumference of a bearing raceway.

A test platform has been developed to enable life-to-failure testing of planetary stage wind turbine gearbox bearings at a scale representative of a $2 - 3MW$ wind turbine. Test bearings have been instrumented with bespoke ultrasonic sensors and a link has been established between the obtained ultrasonic measurements and the breakdown of the lubricating oil film during a rolling surface failure. The measurement technique has been further implemented in an operational wind turbine, illustrating that the measurement principles and systems are sufficiently robust to be applied in the field and could thus be viably utilised to support future improvements in wind turbine gearbox bearing reliability as well as MultiLife functionality.

### 1.4.3  Thesis Layout

This thesis is arranged as follows. Chapter 2 begins by introducing the bearing life rating technique and discusses its development to an internationally recognised standard. Following this, bearing failure mechanisms are briefly introduced, leading on to a presentation of technologies currently and historically used to measure and monitor bearing condition. Following this, a brief description of the monitoring techniques currently employed within wind turbines is presented. Chapter 3 introduces the ultrasound technique, its application to the measurement of oil film thickness in bearings, and the ways in which ultrasound can be generated, acquired and processed to facilitate such measurements.

Chapter 4 describes the development of the MultiLife test platform and the determination of appropriate operating conditions for accelerated life testing. Ultrasound implementation within the test rig is also presented and further consideration is also made to the general applicability of the ultrasound technique to a MultiLife bearing. Chapter 5 then describes the development of a theoretical model to assess bearing life enhancement through raceway indexation and applies this to the test platform, through consideration of the proposed operating parameters established in Chapter 4.

In Chapters 6 and 7 ultrasonic data obtained from the test platform is analysed and discussed. Chapter 6 deals with more general features of ultrasonic measurements, linking these to particular aspects of bearing operation. Chapter 7 then explores the relationship between ultrasonic reflection and bearing condition through assessment of data obtained through a full bearing life cycle test.

Chapter 8 discusses the application of ultrasound sensors in an industrial WTGB. The practical challenges associated with taking such technologies from a laboratory based setting to that of a real industrial system are discussed, including aspects such as instrumentation, data acquisition, data management and the development of autonomous software. Early data obtained from the system as well as that obtained during bench testing of the gearbox prior to installation are also presented.

Finally, conclusions of the presented work are drawn in Chapter 9.

# Chapter 2

# ROLLING ELEMENT BEARING FAILURES AND METHODS OF DETECTION

In this chapter the causes of bearing failures and the associated failure mechanisms are introduced with focus on the drivers of rolling contact fatigue. This leads to a presentation of the methods that have been developed over the last century to predict bearing fatigue life, resulting in the internationally recognised standard used widely in present day machine design. Following this, a range of techniques for detecting bearing damage are discussed and reviewed in terms of the relative ability of each technique to detect faults ahead of catastrophic failure. Finally, the techniques currently utilised for such detection in Wind Turbine Gearboxes are presented.

## 2.1 Analysis of Bearing Failures

A bearing is considered to have failed at the point at which it ceases to operate as intended, rather than the point at which it no longer functions altogether [21]. Elevated noise and vibration levels, as well as increased operating temperatures are the primary indicators of this. Damage may take a variety of forms, and ensue for a multitude of reasons, not only due to operational events, but also as a result of installation and manufacturing issues. The failure mode specific to the bearings under consideration within this study has already been presented in the previous chapter, however, it is important to understand the multitude of ways in which

bearings can fail and the reasons behind such failures in order to be able to develop mitigation strategies appropriately.

Bearing life is ultimately finite, and through normal operation, under "ideal" operating conditions, failures will eventually occur through yield of the bearing material via the process of rolling contact fatigue, which is discussed in more detail later in this chapter. Aside from this, all other failures are considered avoidable, and may arise at bearing rolling surfaces due to a range of application specific factors, namely [22]:

- Incorrect mounting or assembly and poor handling prior to installation
- Inappropriate lubricant type, insufficient supply, or overheating such that adequate film thicknesses are not generated
- Application of misaligned loads, or inappropriate selection of bearing type and/or size such that rated load and/or speed are exceeded
- Insufficient filtering of lubricant or inadequate sealing of bearings such that contaminants may be entrained into rolling contacts
- Exposure to transient operating conditions such as complex and fluctuating loads, highly variable speeds, rotational oscillations and vibration
- Exposure to electrical currents

Studies have suggested that up to 95% of bearing failures occur prematurely, as a result of any number of the above factors [23]. There exists a vast array of literature documenting bearing failure modes, their causes and alleviation strategies, as published by researchers and technology specialists, but also by bearing manufacturers themselves. Figure 2.1 provides a summary of some such documentation, linking specific failure modes to the operational factors listed above. Only those resulting in failure of bearing rolling surfaces are considered within this summary; however other failures are possible, such as those to rolling element cages, bearing end faces and inner bore and outer diameter surfaces. Each failure mode is discussed further within the following paragraphs.
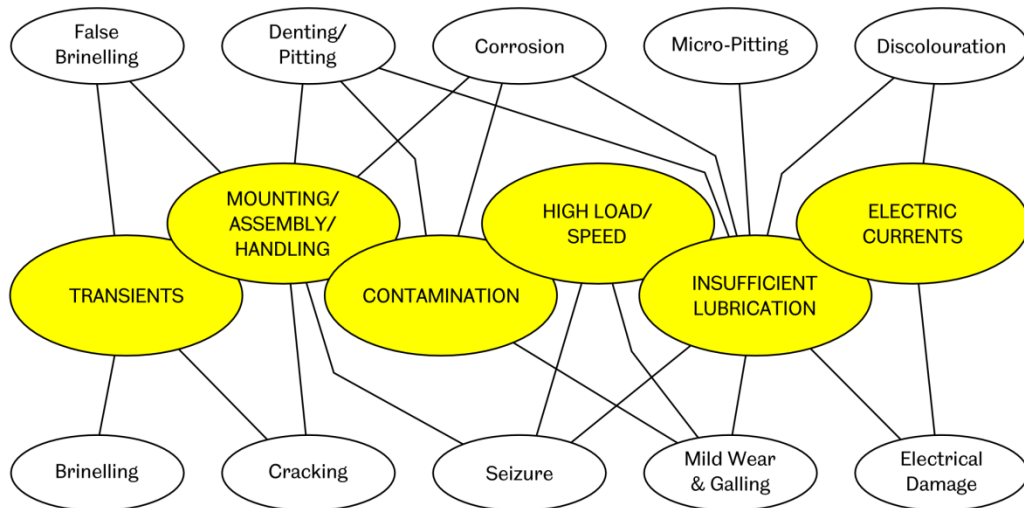
Figure 2.1: Typical bearing failure mechanisms resulting from associated off-design operational factors

## Brinelling and False Brinelling

Brinelling damage is that of plastic deformation to raceway surfaces due to severe and sudden overloads whilst bearings are stationary. Regularly spaced raceway indentations are induced at the points where the rolling elements are situated, thus resulting in uneven operation and noise [24].

False brinelling on the other hand may also occur in stationary bearings but as a result of vibration or very small oscillations and, by contrast, develops via an abrasive fretting wear mechanism [25]. Its appearance is similar to that of brinelling but wider indentations are typically induced due to the amplitude of vibration [22]. False brinelling can arise due to transportation, and is therefore also highlighted as a result of handling in Figure 2.1. This will tend to arise if bearings are transported in pre-assembled machines and are therefore under load but do not rotate. The ambient vibrations experienced during transit can be sufficient to induce such defects [26].

## Cracking and Fracture

Cracks can often be introduced during bearing installation as a result of hoop stresses induced when bearings are installed with tight interference fits [26]. Short duration high load impacts can cause sub-surface cracks within bearing raceways, which – if significantly high – will result in complete component fracture, but can often develop into surface spalls through propagation via rolling contact fatigue [25].

## Denting/Pitting

Denting and pitting are two terms often used interchangeably, but the latter is often also used to describe the manifestation of surface damage resulting from a corrosion type mechanism. The former tends to relate more to macro-scale plastic deformation of rolling surfaces; and is the intended target of discussion here. This is essentially a brinelling process but on a smaller scale and manifests as small plastic deformations to rolling surfaces due to the entrainment of hard particles within the rolling contact. Such particles may be introduced in the installation process, but will more commonly arise due to wear debris from other machine elements (as a result of inadequate bearing sealing) or be transported within lubricant which has not been sufficiently filtered. Denting is also linked to lubrication due to the fact that effects will be enhanced if lubricant films do not act to separate the rolling surfaces sufficiently. The process will also ensue from wear debris generated from the bearing itself, and is therefore often a secondary failure mechanism [26].

## Corrosion

Corrosion to bearing surfaces can arise prior to installation as a result of exposure to moisture and oil residues during handling. If such contaminants are able to penetrate the protective oxide layers that form around the bearing steels then chemical etching can take place. A bearing may also become exposed to moisture during operation through inadequate or damaged seals [27]. In addition to this, unsuitable lubricants and particular lubricant additives have been found to induce corrosion in some cases [23, 25]. Light corrosion may only affect the top surface of the bearing material and can often be removed with light abrasive paper and not subsequently effect bearing operational performance, more heavily corroded surfaces may also require re-grinding to remove such surface disruption. More advanced cases however, when left for some time, can progress below the bearing surface and result in the formation of pits, which may advance to spalling damage under continued normal operation via rolling contact fatigue [27].

## Seizure

Seizure may arise if a temperature gradient exists between inner and outer bearing raceways. This can act to cause a higher rate of expansion to one bearing component – potentially reducing radial clearances – thus inducing preloads such that loading of rolling contacts is increased and subsequently higher friction is induced, essentially causing restriction to

bearing rotation [22]. This can be linked to lubrication since temperature gradients will often arise if sufficient lubricant does not exist to dissipate heat appropriately. Seizure will logically also be induced as a result of extremely high operational loading, and also if bearings are installed with excessively high interference fits thus reducing radial clearance from the outset. The high localised friction and temperatures that are induced will effectively result in material softening, melting and deformation [28].

**Mild Wear and Galling**

Surfaces on a micro-scale are uneven and composed of many peaks and troughs, or "asperities." Bearings in particular are manufactured with precision to minimise such surface variation, however a certain degree of roughness is inevitable. Mild wear, in this context, refers to the removal of such asperities and flattening of rolling surfaces. Galling, on the other hand, is a more severe mode of wear and is associated with adhesion and macro-scale removal of material from the rolling surface [29]. All forms of wear generally arise when rolling elements begin to slide rather than roll across bearing surfaces.

Mild wear will typically occur under light loading, but where lubrication is insufficient to fully separate the contacting surfaces and a certain degree of sliding can occur [26]. The light nature of such wear will often cause rolling surfaces to become more polished in appearance, and this is generally not considered critical to bearing operation initially [22]. Continued light wear such as this however generates increased amounts of wear debris thus increasing the overall rate of surface damage via third body abrasion [30]. Eventually this may lead to increased bearing clearance, thus potentially resulting in misalignments of shafts and damage to surrounding machine elements.

More severe galling mechanisms will ensue as a result of, again, lack of sufficient lubrication, combined with higher temperatures and more severe operating conditions such that local temperatures at asperity junctions are sufficient to promote welding [26, 28]. This is generally observed via scoring and tearing to rolling surfaces.

**Micro-Pitting**

Micro-pitting is often also referred to as a severe form of surface distress [22], and is essentially a plastic deformation mechanism much like that of brinelling and denting, but on an asperity scale. This tends to occur under poor lubricant conditions [31], resulting in increased metal-metal contact

which subsequently places surface asperity peaks under high localised loads [26]. This results in surface damage in the form of very small cracks and pits/spalls, and can also cause a certain degree of surface polishing if sliding motion also takes place [31].

**Discolouration**

Discolouration is often a precursor to more severe forms of failure, and is typically an indicator of high temperature and insufficient lubrication. Discoloration to roller and raceway surfaces will result from burnt lubricant which subsequently stains the metal surfaces. Extended periods of metal-to-metal contact will also cause localised heating of the bearing materials, thus resulting in metal discoloration when temperatures reach significant levels [27].

**Electrical Damage**

In cases where bearings are exposed to electrical currents, potential differences can arise across bearings as a result of the electrical resistance of the lubricant films [26]. This can cause sparking between the two components, generating high localised temperatures, resulting in melting and pitting to both surfaces [27, 28]. This can often result in a "corrugation" or "fluting" effect to rolling surfaces which can sometimes be similar to – and difficult to distinguish from – the process of false brinelling, however, electrical fluting will often also show signs of discolouration from surface burning [30]. Electrically conductive greases are available and may be a more appropriate choice in situations where electrical currents will inevitably be applied to bearings [26].

## 2.1.1 Rolling Contact Fatigue and Spalling

The final mechanism of wear, which has been mentioned briefly in the above, and discussed in relation to WTGB planetary bearing failures in §1.2.2, but which has not yet been appropriately defined is "spalling." This process, often also referred to as "flaking," "peeling" or "pitting," is the process of surface failure via rolling contact fatigue (RCF). As alluded to at the beginning of this chapter, even when all means of operational best practice are adopted and rolling surfaces are smooth, well lubricated and operate under pure rolling, failure may still arise as a result of such material fatigue after a certain duration of operational time has elapsed [32]. This mode of failure is sub-surface initiated, arising from cyclic stressing of the volume of raceway material residing within the bearing loaded zone as rolling elements repeatedly move across it. As illustrated

in Figure 2.2(a), cracking tends to occur around material voids or non-metallic inclusions, usually at relatively steep angles to the rolling surface [17]. Under repeated loading cracks will grow and eventually propagate to the rolling surface, causing a segment of material to be removed, leaving a pit (or spall) [21]. As discussed in §1.2.2, it is thought that this mechanism may also be driven by transient loading, acting to essentially accelerate the crack formation and growth process. Debris generated during this process can often become entrained in rolling contacts if such bearings remain in-service, resulting in localised overloading and stress concentrations, and leading to further surface damage [24].
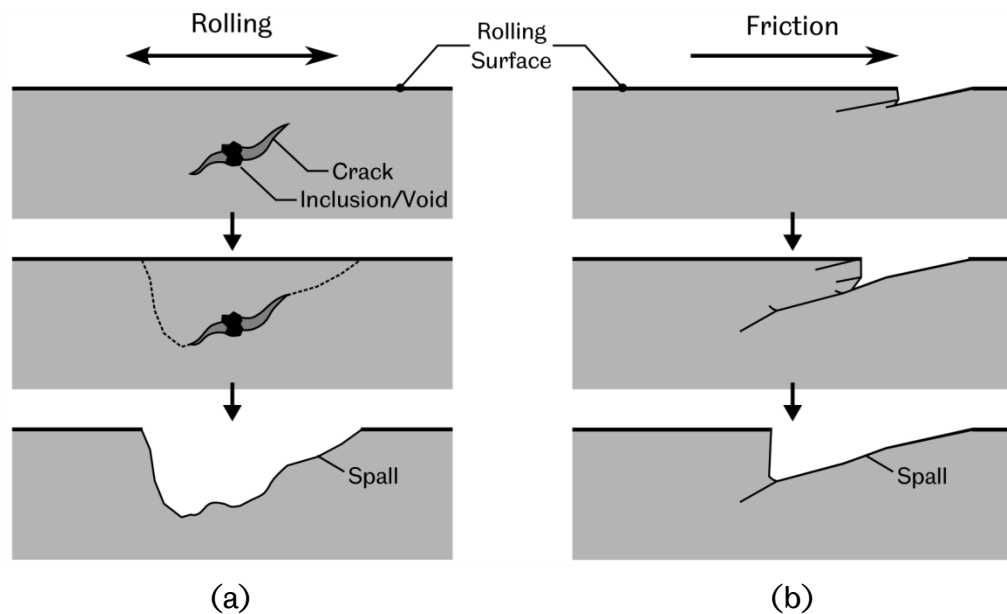


Figure 2.2: Illustration of (a) sub-surface and (b) surface initiated rolling contact fatigue resulting in surface spalling (adapted from [13, 17, 33])

As illustrated in Figure 2.2(b), RCF may also be *surface* initiated, which generally manifests as shallower spalls, but which can be just as detrimental to bearing performance. This can occur as a direct result of insufficient bearing lubrication, where raised surface regions adjacent to furrows and grinding marks begin to interact within the contact due to a lack of fluid to separate the surfaces. Cracks and pits subsequently form around such areas [33, 34]. Elevated friction due to sliding is also known to compound the problem, which is enhanced in the presence of additional surface defects resulting from alternative failure mechanisms such as denting/pitting, micro-pitting, brinelling or corrosion [22]. Furthermore, such additional failure mechanisms change surface geometry, form debris and can also cause misalignment and point loading, which all act to raise

surface stresses, thus again enhancing crack formation and growth and accelerating the RCF process [34, 35].

The processes of bearing failure via both surface and sub-surface initiated rolling contact fatigue are summarised in Figure 2.3. Based on the above discussion it is clear that RCF is a crucial factor in bearing operation, which has subsequently led to it providing the basis for standardised assessment of bearing life, the development of which is explored further in the following section.
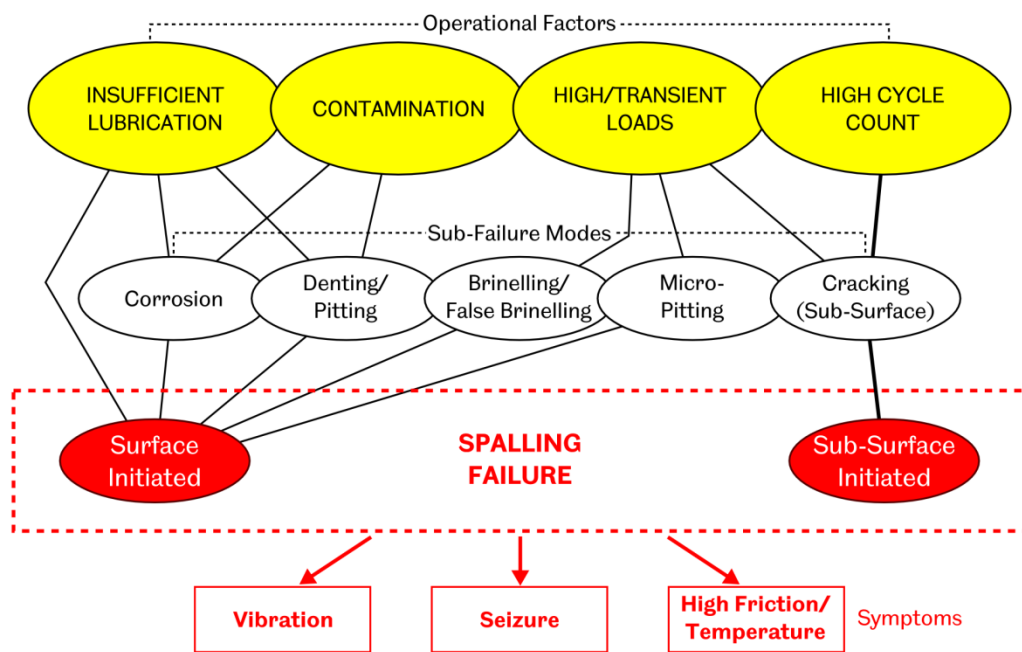
Figure 2.3: The process of bearing failure via rolling contact fatigue

## 2.2 The Establishment of a Bearing Life Rating

Bearing life rating calculations allow a prediction to be made to the fatigue endurance of a ball or roller bearing based on the loading conditions applied and the geometry of the bearing under analysis. Bearing life is defined as the number of revolutions (or hours, if the application is at constant speed) before the first evidence of fatigue spalling occurs in the material of the rings or rolling elements [12].

As bearing technologies have advanced over the last century the methods employed to perform such predictions have also developed and attempts

have continually been made to improve accuracy in such calculations. National standards have been created and maintained to provide a basis for life calculation of bearings in any given application. In addition to this, bearing manufacturers themselves often maintain methods specific to their own bearings based on in-house life cycle testing and analysis. General methodology in all cases is the same, yet life modification factors (as discussed in the following section) will often vary. The current section discusses the development of the bearing life rating procedure, from its inception to the most recent enhancements.

### 2.2.1  Early Analysis of Bearings and Loading

The earliest analysis of rolling element bearings was performed by Stribeck, who established an equation to relate the radial load on a bearing, $F_r$, to the resultant load on the single most heavily loaded element.

$$Q_{max} = \frac{5F_r}{Z}$$

(2.1)

This was applicable to bearings with zero internal clearance under a purely radial load, where bearings contained a total of $Z$ rolling elements (cited in [36]). Stribeck further defined the maximum load carrying capacity of a given bearing as:

$$C = \frac{kZd_e^2}{5}$$

(2.2)

For ball bearings (cited in [37]) and:

$$C = \frac{kZLd_e}{5}$$

(2.3)

For roller bearings (cited in [36]), where $d_e$ is the roller or ball diameter, $L$ the roller length, and $k$ is a constant relating to the bearing type.

Goodman was the first to introduce the concept of rolling contact fatigue, and extended the Stribeck definition of load carrying capacity to also consider bearing speed. These concepts were soon adopted by bearing manufacturers who began to develop their own calculations for maximum

bearing load, but which closely resembled those of Stribeck's with the added speed component [37].

## 2.2.2 The Basic Life Rating

The work of Lundberg and Palmgren, published in the 1940's, is considered the most significant contribution to roller bearing life theory, and still provides the basis for current bearing life rating standards [12]. Based on statistical methods established by Weibull for the failure of solids, they established a probability of bearing survival after $N$ operational cycles. This was defined specifically as the formation of the first spall arising from rolling contact fatigue (cited in [37]):

$$\ln\left(\frac{1}{S}\right) \propto \frac{N^e \tau_0^c V}{z_0^h}$$

(2.4)

Where $S$ is the probability of survival, $V$ is the volume of the raceway material that will endure stress and the exponents $e$, $c$ and $h$ are dependent on the shape of the Weibull distribution curve and are a function of material and bearing type.

As is apparent from Equation (2.4), Lundberg and Palmgren attributed crack formation and subsequent spalling failure (as described in §2.1.1) to the occurrence of a reversing maximum orthogonal shear stress, $\tau_0$. This was suggested since regions of the sub-surface material would effectively experience a maximum variation in such shear stress equivalent to $2\tau_0$, which would quickly take its toll on raceway material, considering such stress reversals would be repeatedly induced with every roller passage. Such stress is known to arise within the raceway material during the passage of a roller, and acts parallel and normal to the rolling direction at a depth $z_0$ below the surface [29]. The maximum orthogonal shear stress and the Lundberg-Palmgren hypothesis are discussed further in §2.2.3

From Equation (2.4) a simple equation was derived that linked bearing life to applied load, a full derivation of which can be found in [21], but which essentially leads to the following:

$$L_{10} = \left(\frac{C}{P}\right)^p$$

(2.5)

Here, $L_{10}$ is the life (quoted per million revolutions) that 10% of a sample of bearings would be expected to fail under the same applied equivalent load, $P$ (a combination of both axial and radial load components). $C$ was termed the "*dynamic* load capacity", representing the load at which 10% of the bearing sample would fail after $10^6$ revolutions. Lundberg and Palmgren originally developed equations to establish this for a variety of bearing types based on Hertzian contact theory. Current standards provide simplified calculations which are based on bearing parameters such as the rolling element diameter, contact angle, number of rolling elements and empirical factors to account for the variety of bearing geometries and material types [12].

$p$ is a load-life exponent which relates to the Hertzian contact mode at play between rollers and raceways for a given bearing type. Lundberg and Palmgren determined such exponents experimentally and found $p = 3$ to be appropriate for point contact scenarios such as that occurring within ball bearings. For roller bearings – where a line contact is made – the relation $p = 4$ was initially defined, but was later reduced to $p = 10/3$ as the design of roller bearings developed and most began to operate in a combined point and line contact mode [37].

### 2.2.3 The Lundberg-Palmgren RCF Initiating Stress Hypothesis

As discussed above, the basic life rating principle was, and still is based on the Lundberg-Palmgren assumption that maximum orthogonal shear stress acts as the key driver to RCF. Figure 2.4 illustrates such variation in the orthogonal shear stress $\tau_{yz}$ across the width of a rolling contact equivalent to that which will occur in a cylindrical roller bearing such as that of key relevance within this study. The depth at which the peak $\tau_{yz}$ occurs for a given point along the $y$ axis will vary, and the plot therefore represents the depth at which it is maximised ($z_0$) at each point in $y$.
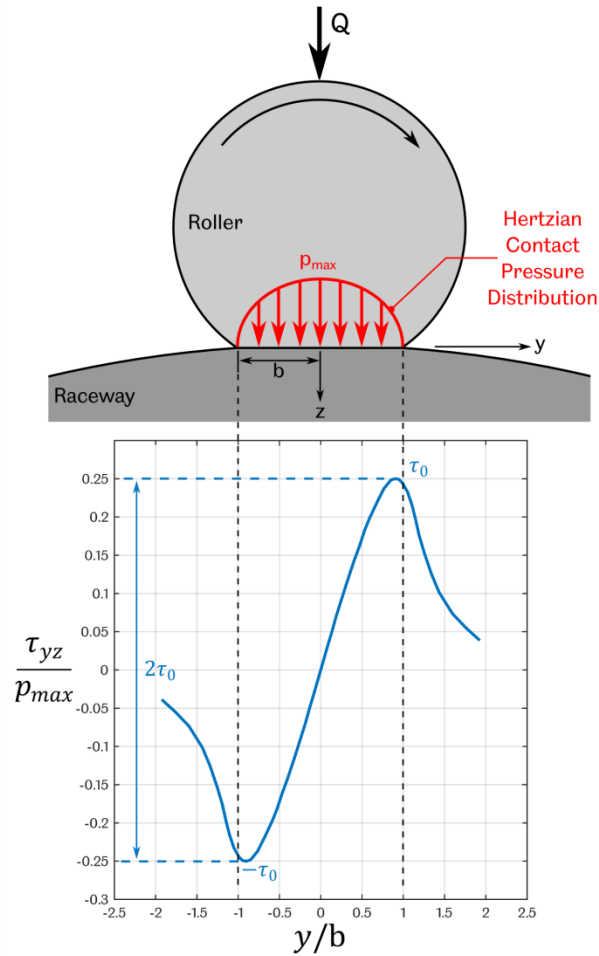
Figure 2.4: Maximum orthogonal shear stress for a Hertzian line contact occurring at a depth $z_o$ across the contact width, $b$ (adapted from [21])

Maximum orthogonal shear stress is therefore related to the maximum Hertzian contact pressure according to:

$$\tau_0 = 0.25 p_{max} \tag{2.6}$$

Where, as dictated by Hertzian theory, the pressure distribution across the width of the contact follows an elliptical profile, and is maximised at the centre of the contact such that:

$$p_{max} = \frac{2Q}{L\pi b} \tag{2.7}$$

Here, $Q$ is the normal load applied to the contact and $b$ is the contact half width, as illustrated in Figure 2.4. $L$ is the contact length, where it is assumed that the pressure distribution will be equal across this dimension. The contact half width can also be established from Hertzian theory:

$$b = \sqrt{\frac{4QR^*}{L\pi E^*}}$$

(2.8)

Where $R^*$ is the reduced radius of the roller (Body 1) and raceway (Body 2) contact and $E^*$ is the composite contact modulus, both of which are established from Equations (2.9).

$$\frac{1}{R^*} = \frac{1}{R_1} + \frac{1}{R_2}$$

$$\frac{1}{E^*} = \frac{1-v_1^2}{E_1} + \frac{1-v_2^2}{E_2}$$

(2.9)

Where $v$ is Poisson's Ratio.

As can be seen in Figure 2.4, the maximum orthogonal shear stress does not occur directly below the contact, but in fact will occur at $y = \pm0.9b$. Further to this, whilst not illustrated in the figure, $\tau_0$ is also known to be induced at a depth $z_0 \approx 0.49b$ [29].

More recent research has questioned the initial Lundberg-Palmgren hypothesis, since metallurgical studies of bearing materials after operation have shown plastic deformations to be observable somewhat deeper into the material sub-structure at around $0.75b$. This appears to coincide with the location of the maximum *principal* shear stress $\tau_1$, that is found if consideration is made to shear stress in any orientation within the $yz$ plane [38]. Maximum principal shear stress for a line contact is defined by Equation (2.10) and, as illustrated in Figure 2.5, occurs directly below the contact ($y = 0$), at a depth of $0.786b$.

$$\tau_1 = 0.3p_{max}$$

(2.10)
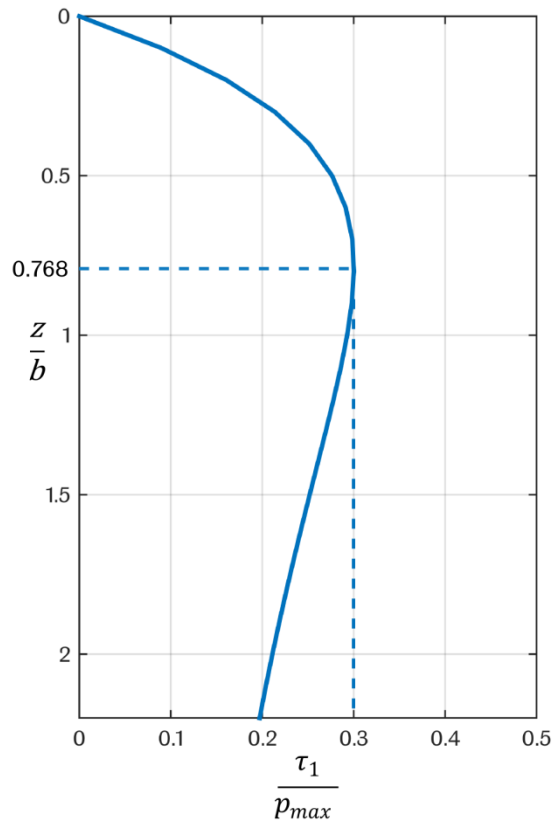
Figure 2.5: Maximum principal shear stress in the z axis ($y = 0$) at various depths below the surface (reproduced from [29])

This has also therefore been suggested as an appropriate metric in determining rolling contact fatigue life [39].

### 2.2.4 Bearing Life Modification Factors

Enhancement in the quality and cleanliness of bearing steels relatively soon after the widespread adoption of the Lundberg-Palmgren bearing life principle meant that bearing life calculations began to increasingly under-predict the fatigue lives observed in reality. This was a result of bearing materials being stronger and having fewer inclusions such that components could support higher stresses and less crack initiation points existed within the materials [37]. Furthermore, the Lundberg-Palmgren theory relied solely upon sub-surface fatigue initiation, and did not consider the possibility of surface initiated fatigue, as identified in §2.1.1 [32], nor did it consider the effects of lubrication on bearing life. Indeed, since the inception of the model, lubrication theory has improved

dramatically such that the effects to bearing life of different regimes of elastohydrodynamic lubrication (EHL) are now more widely understood [35].

As such, modification factors were applied to the original Lundberg-Palmgren equation (Equation (2.5)) to realign predictions with updated bearing endurance data.

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

(2.11)

The similarity of the above to Equation (2.5) is clear, however modification factors $a_1$, $a_2$ and $a_3$ were adopted which essentially acted to multiply the Lundberg-Palmgren or "Basic" rating life according to additional operational parameters.

### 2.2.4.1 Life Modification Factor for Reliability

The life modification factor $a_1$ allowed alternative probabilities of failure to be assessed other than the standard 10% failure rate associated with the basic $L_{10}$ life rating. For instance, if a greater certainty in bearing reliability was required, a machine designer might opt for assessment at a failure rate of 1%, that is, the life at which 99% of a sample of bearings would be expected to exceed the life rating (thus becoming $L_1$). The life modification factor for reliability was defined according to Equation (2.12), which was based on a Weibull distribution with a slope factor of 3/2, and where $S$ represented the percentage reliability [35].

$$a_1 = \begin{cases} 1, & S = 90 \\ 4.48 \left(\ln \frac{100}{S}\right)^{\frac{2}{3}}, & S > 90 \end{cases}$$

(2.12)

### 2.2.4.2 Life Modification Factor for Material

The life modification factor for material was employed to account for general improvements in bearing steels since the original model was introduced. This factor was also designed to account for a variety of improvements in production processes such as advanced heat treatment, and manufacturing methods [35].

### 2.2.4.3 Life Modification Factor for Lubrication and Operating Conditions

The final life modification factor, $a_3$, was initially defined in an attempt to compensate for the effects of lubricant film formation. This was based on consideration of the Lamda Ratio ($\Lambda$), which will be discussed further in §3.3.3, but is essentially the ratio between theoretical minimum lubricant film thickness and the composite roughness of the contacting surfaces [37]. This was later expanded to also take into consideration the internal load distribution within the bearing – which would vary based on factors such as the radial clearance, and thus affect the load applied to the roller and raceway contacts – and temperature. Both of which would ultimately influence lubricant film thickness formation [35].

## 2.2.5 Introduction of a Limiting Stress

Despite the introduction of the above modification factors, predicted bearing lives were still found to deviate from actual endurance data over a wide range of operating conditions. Indeed, it was subsequently found that for applied loads below a threshold, bearing fatigue life was essentially infinite [37]. Ioannides and Harris defined a new bearing survival relationship, similar to Equation (2.4), but introducing the concept of a fatigue limit stress [40].

$$\ln\left(\frac{1}{\Delta S_i}\right) \propto \frac{N^e(\sigma_i - \sigma_u)^c \Delta V_i}{z_i^h} \tag{2.13}$$

The initial notable difference to Equation (2.4) was that an elemental approach was employed, where subscript $i$ indicates assessment of an element of the bearing material such that survival of that particular element was based on a locally induced stress $\sigma_i$ within the elemental volume $\Delta V_i$. Secondly, a threshold was applied to the induced stress, such that if a fatigue limiting stress $\sigma_u$ was not exceeded in the material element, fatigue failure would not occur. It was also noted by Ioannides and Harris that any stress parameter could be used as the failure initiating stress, and the equation was not necessarily limited to use of the orthogonal shear stress $\tau_0$.

Acknowledgement was also made by Ioannides and Harris, to more recent observations that with the presence of contaminant particles within lubricants, the onset of fatigue failure would be accelerated due to such

particles becoming entrained within the rolling contacts. This, it was suggested, could be accommodated through application of stress concentration factors to the stress parameters in Equation (2.13), since surface stress raisers would effectively act to multiply sub-surface stress fields in a similar manner.

## 2.2.6 The Current ISO281 Bearing Life Rating Procedure

The procedures for defining the life rating of rolling element bearings are now standardised under ISO281:2007 [12], and the modified life rating equation currently assumes the form:

$$L_{nm} = a_1 a_{ISO} \left(\frac{C}{P}\right)^p$$

(2.14)

### 2.2.6.1 Life Modification Factor for Reliability

Current values of $a_1$ deviate slightly from those previously defined according to Equation (2.12), and have been adjusted to reflect the most up to date observations. Figure 2.6(a) compares current $a_1$ factors against those derived in [35] according to Equation (2.12). Extrapolation of the ISO $a_1$ data permits assessment of bearing life at higher probabilities of failure. While such extrapolation is not recommended within the standard, it will be of relevance further on in this study to obtain a basic understanding of approximate bearing life at higher certainties of failure.
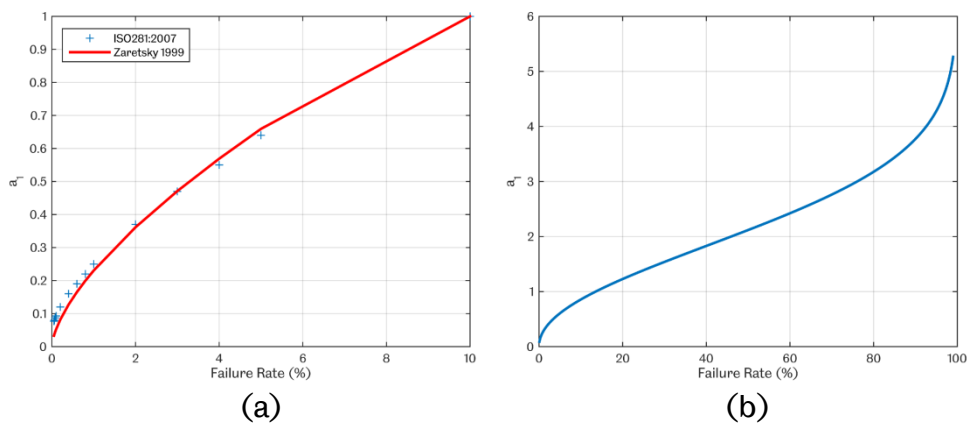


Figure 2.6: (a) Bearing life modification factors for reliability according to [12] as compared with earlier definitions, (b) extrapolation of $a_1$ values for failure rates up to 99%

As such, an extrapolated curve has been produced in Figure 2.6(b) to indicate the likely progression of $a_1$ values beyond failure rates of 10%.

### 2.2.6.2 Life Modification Factor for Systems Approach

Individually defined life modification factors for consideration of material and lubrication, as discussed in §2.2.4.2 and §2.2.4.3 respectively, were eventually found to be misleading, since it was recognised that both factors were in fact co-dependent [37]. In response to this, bearing manufacturers began to define a single modification factor to accommodate the effects of both. This was adopted and expanded – in combination with the Ioannides-Harris theory of a fatigue stress limit – to provide a more comprehensive life modification factor to accommodate a range of operational elements [12].

$$a_{ISO} = f\left(\frac{e_C C_u}{P}, \kappa\right)$$

$$(2.15)$$

To simplify calculations for the machine designer, $a_{ISO}$ designates a fatigue *load* limit, $C_u$ which relates to an induced Hertzian contact stress limit of around $1500 MPa$, deemed applicable to current high quality bearing steels. This is typically defined by the manufacturer and draws still upon the Lundberg-Palmgren hypothesis of the orthogonal shear stress being the key driver of RCF. This is largely based upon bearing load distribution and therefore draws upon the Stribeck equation (2.1) for the maximum internal bearing load, and subsequently equates the maximum shear stress emanating from this.

Lubricant contamination levels are defined by the contamination factor, $e_C$, which can range from $0 - 1$. This is generally applied according to grading criteria for oil cleanliness and is also dependant, to a certain degree, on bearing size. Lookup tables are provided within the standard for determination of this, but equations are also provided for more accurate determination.

Finally, $\kappa$, is a measure of the lubricant and its effectiveness to generate sufficient oil film thicknesses within the bearing to sufficiently separate the rolling contacts.

$$\kappa = \frac{\upsilon}{\upsilon_1}$$

<div align="right">(2.16)</div>

This is termed the viscosity ratio, and compares actual kinematic oil viscosity at the expected bearing operating temperature, with a reference minimum viscosity $\upsilon_1$ required by the specific bearing in question. This essentially relates to the Lamda ratio (Equation (3.17), §3.3.3) and can also be determined according to:

$$\kappa \approx \Lambda^{1.3}$$

<div align="right">(2.17)</div>

Upon definition of the factors $\frac{e_C C_u}{P}$ and $\kappa$, the life modification factor $a_{ISO}$ can be obtained either through lookup charts or specific equations, both of which are detailed within the standard. A range of values can be obtained such that $0.1 \leq a_{ISO} \leq 50$, depending on the operating regime.

## 2.3 Detection of Bearing Failures

Having explored the key mechanisms of bearing failure as well as the means to establish statistical bearing life, the following section explores the sensor technologies available to assess bearing health and detect incipient failure.

### 2.3.1 Temperature

Historically, the measurement of bearing operating temperature has been widely employed in machines and systems to assess operational stability. Generally, under steady state operation, bearing temperatures will remain relatively constant, but will begin to rise if operational factors deviate from the norm and become detrimental to bearing performance. This is linked to breakdown of the lubricant, which causes rolling contacts to move from EHL to boundary lubricated regimes, resulting in increased asperity interactions, generating bearing friction, and resulting in high local (flash) temperatures [41].

Resistance temperature detectors (RTDs) or thermocouples are typically used to facilitate such measurements in industrial systems. Ideally such instrumentation should be placed as close to the rolling surfaces as

possible, so as to provide accurate measurement of the flash temperatures occurring at contact interfaces. In reality however, this is often difficult to achieve and requires modifications to be made to bearing components, risking compromise to their structural integrity and load carrying capacity. As a consequence, sensors are typically installed as close as possible to the component of interest, usually on bearing outer surfaces or within housings. Measurement of the lubricating oil upon exit from a bearing is also commonly employed to estimate the temperatures occurring within [23].

As a consequence, temperature measurements typically suffer from an inherent lag due to the time taken for high contact temperatures to propagate through to the point of measurement. Generally, temperature rises in the order of $10°C$ beyond that of the normal steady state operating level would be considered significant enough to raise alarms [23], which may take a significant time to occur, thus providing opportunity for considerable bearing damage to develop in the meantime.

## 2.3.2 Oil Analysis

Lubricating oils play a vital role in bearing operation, and consequently information can be gleaned from them relating to the health of a bearing. Over a period of time lubricants may become degraded and their ability to restrict the metal-to-metal contact of bearing components is impeded. A single oil supply is typically used to feed numerous bearings and other components within a system, therefore the presence of wear debris or breakdown of lubricant can have significant effects on an entire mechanical system. With regards to the monitoring of bearing health specifically, the presence of debris in a lubricant may either be directly linked to wear of the bearing under consideration, or due to wear of other components upstream of the lubricant supply. The latter can therefore provide indication of impending bearing failure should wear debris be entrained within the bearing contacts, and may therefore be alleviated through a lubricant change.

As such, a key factor in the analysis of lubricating oils is the determination of the size and quantity of foreign particles that reside within them. Indeed, an international standard has been developed to classify such particle contamination levels across a universal scale [42]. In addition to this however, analysis is also performed to establish changes to the oil properties, the status of the additives that provide advanced lubricating performance, and the levels of water or solvent ingress [23]. Such analysis

is typically laboratory based and is therefore performed offline, where samples of bearing lubricant are taken at regular intervals and then analysed through use of various methods such as microscopy, filtration techniques, atomic absorption spectroscopy, and Scanning Electron Microscopy (SEM) coupled with Energy Dispersive X-ray Analysis (EDAX) [43].

More immediate measurement techniques have been developed in recent years which can often be applied in-situ and therefore enable real-time measurement of oil condition. Particle counters are most common, utilising white light, or lasers, coupled with detectors to measure the scatter of the light source as oil moves between the source and detector. Fourier Transform Infrared Spectroscopy (FTIR) permits the measurement of a range of lubricant properties such as oxidation and additive levels through use of infrared radiation. Various chemical components will absorb such radiation at specific frequencies, and therefore the measurement of the absorption spectrum provides indication of the relative changes in lubricant composition when compared with the spectra of a clean, unused sample of the same oil, an example of which is shown in Figure 2.7. While this is widely used in a laboratory setting, the technique has also been proven to be applicable as an on-line condition monitoring tool [44].



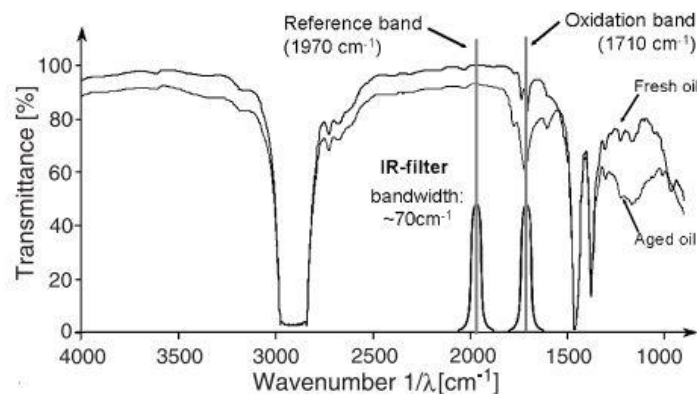Figure 2.7: Comparison of Infrared absorption spectra for a new and aged oil (reproduced from [44])

The electrically insulating nature of oils has also been exploited to detect the presence of metallic wear particles as well as acids, ions and broken oil molecules which all act to increase the electrical conductivity of the lubricant, a concept which has also shown success as a further on-line condition monitoring tool [45].

Overall however, given the fundamental nature of oil analysis, features will only become detectable within a lubricant subsequent to the occurrence of bearing wear or oil degradation, and therefore the technique does not provide a particularly early warning of such failures. Furthermore, as highlighted above, since a single oil supply will typically feed an entire system of bearings and components, the ability to locate the root-cause of failure can often be difficult without necessitating a full system strip-down.

## 2.3.3 Vibration

Bearing vibrations will be observed even in entirely healthy bearings and will arise through the cyclic passage of rolling elements through the bearing loaded zone, thus altering the internal load distribution and generating periodic vibration in both the rotating structure and bearing housing [46]. Vibration can also arise as a result of manufacturing faults such as increased surface roughness and waviness of rolling surfaces. While such faults are rare – given the ever increasing precision to which bearings are manufactured – even the slightest deviation from manufacturing tolerances can have a profound impact upon the smooth running of bearings in-service. Increased levels of vibration will also be observed as a consequence of poor installation, which can lead to misalignments, or the precession of bearing raceways in housings due to inadequate interference fits. Vibration will also increase due to the progressive degradation of bearing raceways as defects are induced in rolling surfaces that are subsequently impinged upon by rolling elements during rotation.

The detection of bearing vibration is facilitated through the use of vibration sensors or accelerometers which are ideally mounted directly to the bearing of interest. This can however often be impractical for bearings within systems such as gearboxes or motors, where access is limited and signals cannot be easily routed out to an external data acquisition system. Often therefore it is necessary to mount sensors to the outer casing of the machine under consideration, in which case multiple sensors should typically be employed, located at or near to specific areas of interest [47]. Orientation is also an important consideration, where sensors should be aligned to the plane of the expected vibrations, although multiple axis sensors are also common to satisfy all possible vibrational directions.

The detection of bearing failures through vibration monitoring has been the subject of large volumes of research over the past decades. Research

has generally been conducted through continuous monitoring of bearings under controlled operation, where failures have typically been induced in one of two ways [48]. One such method has been to modify the key bearing operational factors that influence fatigue life (as discussed in §2.2) such that the onset of rolling contact fatigue is significantly advanced (see [49] for example). Alternatively, faults have been introduced into bearing rolling surfaces through machining processes such as wire erosion (discussed further in §4.2.1), surface indentation or grinding (all of which are employed in [50] for example). The "Fault Seeding" approach has the advantage of being less time consuming and also permits control of defect size, such that vibration response can be characterised according to damage severity, as achieved in [51].

It has been found however that successful detection of such progressive bearing failures, particularly to obtain early warning indicators, tends to largely depend upon the data analysis techniques employed. While a detailed review of the multitude of available techniques is beyond the scope of this study, a comparison of the most common of such techniques is provided in the following paragraphs.

At a basic level, vibration signals can be analysed in the time domain through analysis of the root mean square (RMS) of the signal and identification of a general increase beyond a known "undamaged" baseline level [49]. The ratio of the maximum signal peak to its RMS level, termed the "crest factor," is found to exhibit greater sensitivity than basic RMS monitoring, and has also been proven to be less sensitive to general operational changes that would also act to increase general acceleration levels [52]. Indeed, most statistical parameters will ultimately vary with the operating conditions of the machine or bearing, and thus parameters such as speed and load must also be considered so as to be able to associate statistical changes appropriately. An exception to this is the "kurtosis" or "peakedness" of a vibration signal, which is said to be almost universally equal to 3 for an undamaged bearing regardless of operating conditions, and thus any increase beyond this can be attributed solely to the onset of damage [49].

Generally, there are limitations to the insight that can be gained from the above procedures however, particularly if sensors are not mounted to the bearing of key interest. For applications where sensors are situated on outer casings of complex systems, and are thus providing measurement of the system as a whole, it is difficult to attribute an observed rise in vibration level to any specific bearing or component.

Alternatively, frequency domain analysis of vibration signals can be used to detect particular operational features and associate vibration and noise with specific bearings within a machine. Impingement of a defect by a passing roller in a bearing will generate short duration high frequency vibration bursts (with frequency content up to around $100kHz$ [53]). Such bursts will occur periodically, at a frequency dependent upon the bearing geometry and speed of rotation, and will also change based on whether the defect resides on the inner or outer raceway or even the rolling element surface. Such inner and outer race ball pass frequencies ($BPF_I$ and $BPF_O$ respectively) as well as the ball spin frequency ($BSF$) are defined in Equations (2.18) [48]. Evidently then, when monitoring the vibration of a machine or bearing system, the observation of such characteristic frequencies may enable an analyst to pinpoint the source of increased machine noise or operational instability.

$$BPF_O = \frac{ZN_r}{120}\left(1 - \frac{d_e}{p_d}\cos\phi\right)$$

$$BPF_I = \frac{ZN_r}{120}\left(1 + \frac{d_e}{p_d}\cos\phi\right)$$

(2.18)

$$BSF = \frac{p_dN_r}{120d_e}\left(1 - \left(\frac{d_e}{p_d}\cos\phi\right)^2\right)$$

Here, $N_r$ is the shaft rotational speed, $p_d$ is the bearing pitch diameter (the mean of the inner and outer raceway contact diameters), and $\phi$ is the raceway-roller contact angle. All frequencies are quoted in $Hz$.

Clearly, the measurement of bearing or system vibrations can be significantly more informative and immediate than the alternative condition monitoring tools discussed so far with regards to detecting bearing failure. Indeed, elevated vibration levels should be detectable at the instance at which rolling surface damage initiates, and therefore with sufficient real-time monitoring and the employment of appropriate signal processing techniques, automated shutdowns can be integrated within control systems so as to cease operation and thus the further propagation of damage. Such changes can often be subtle however, and may be masked by the much more dominant noise and vibrations emitted from the overall machine, thus requiring more advanced signal processing techniques to extract the relevant information. The opportunity for defects to be detected early in the failure process is therefore application

specific, and may be difficult without advanced knowledge of the system under consideration.

## 2.3.4 Acoustic Emission

Acoustic emissions (AE) are transient elastic waves which propagate through solids, and are generated as a result of permanent structural deformation or impact. AE is generally associated with sound waves at frequencies beyond the audible range ( $20kHz$ ), typically between $100kHz - 1MHz$. In bearings, acoustic emissions may arise – in a similar manner to that of vibration – as a result of impingement of surface defects by rolling elements during rotation, where the induced elastic waves propagate through the bearing materials to the point of measurement. What sets this apart from vibration however is that acoustic emissions may also be released within the bearing materials during the formation and propagation of sub-surface cracks, and therefore there is potential for the technique to provide indication to the onset of rolling contact fatigue, prior to the occurrence of surface damage [54].

A limitation of the technique is the relatively high rates of acquisition required to capture acoustic signals with sufficient resolution, where sample rates in the order of $2MHz$ are typically employed. This subsequently results in a requirement for more advanced and costly data acquisition systems. Further to this, large volumes of data can be generated extremely quickly, and therefore – for long term installations – capture regimes must be chosen so as to manage data rates appropriately.

Processing techniques similar to those utilised with vibration data are often employed to analyse AE signals and, as shown in Figure 2.8, defects can similarly be associated with specific bearing components based on the periodicity of the observed acoustic energy bursts according to Equations
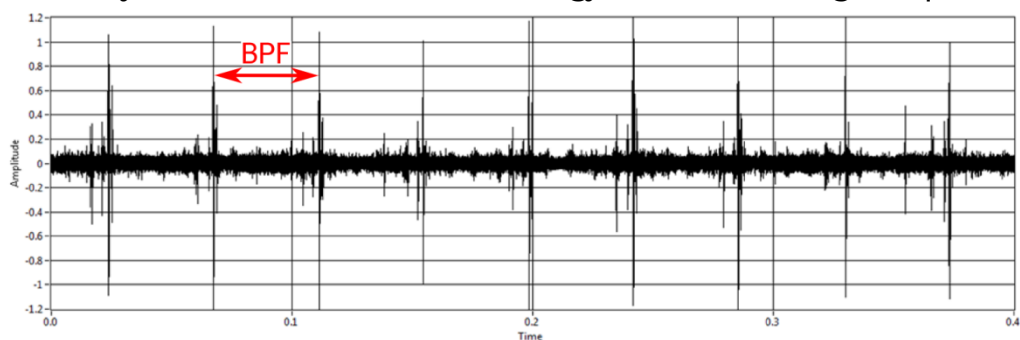


Figure 2.8: Raw time domain AE signal obtained from a rolling element bearing illustrating the occurrence of periodic acoustic bursts associated with a characteristic defect frequency (courtesy of J Naumann)

(2.18). However, when comparing similar processing techniques applied to both AE and vibration data obtained from the same bearing system, AE has often been found to be significantly more sensitive in providing early detection of incipient failure [55-57].

Alternative processing and analysis techniques applied to AE signals have also been used to provide further insight into bearing faults; [58] for example, demonstrated the relationship between AE burst duration and defect size. Detection of defects has also been shown to be further enhanced through use of multiple acoustic emissions sensors. In situations where sensors can be distributed around a component and all the signals acquired simultaneously it is possible to compare the relative times at which an AE burst is received at each sensor, and thus determine the specific location of the acoustic source through triangulation [59]. Such calculations can be difficult in components with complex geometries however, leading to significant errors; particularly if the speed of sound through the medium is approximated, or varies due to the component being made up of multiple materials. Such methods have been optimised however through calibration procedures performed on the component of interest, where artificial acoustic signals have been generated and measured around the structure prior to it being commissioned. The calibration data then acts as a lookup table, which can be used to more accurately locate real defects when in operation, without the need for speed of sound information [60]. This has been applied to bearing raceways with great success, an example of which is shown in Figure 2.9 illustrating the location of an artificial acoustic source based on the differences in the time of arrival at three raceway mounted AE sensors.

Overall, the advantages of the acoustic emissions technique over vibration measurements have been demonstrated through much earlier indication of incipient bearing failure, as well as the apparent relationship between acoustic emission energy and defect size. In addition, AE presents further advantages due to the potential to detect stress waves within bearing materials induced due to the formation of sub-surface cracks. This however, may be difficult in large systems where general noise and vibration levels are high. Furthermore, the technique is optimised through installation of sensors directly to the component of interest, when this is not practical however, and sensors must be placed on outer casings and further away from the acoustic source, sensitivity will inevitably be reduced. Indeed, this may restrict detection entirely if there is a break in the sound propagation path between the source and sensor. Therefore,

the applicability of the technique in such cases remains to be seen and is the subject of ongoing research.
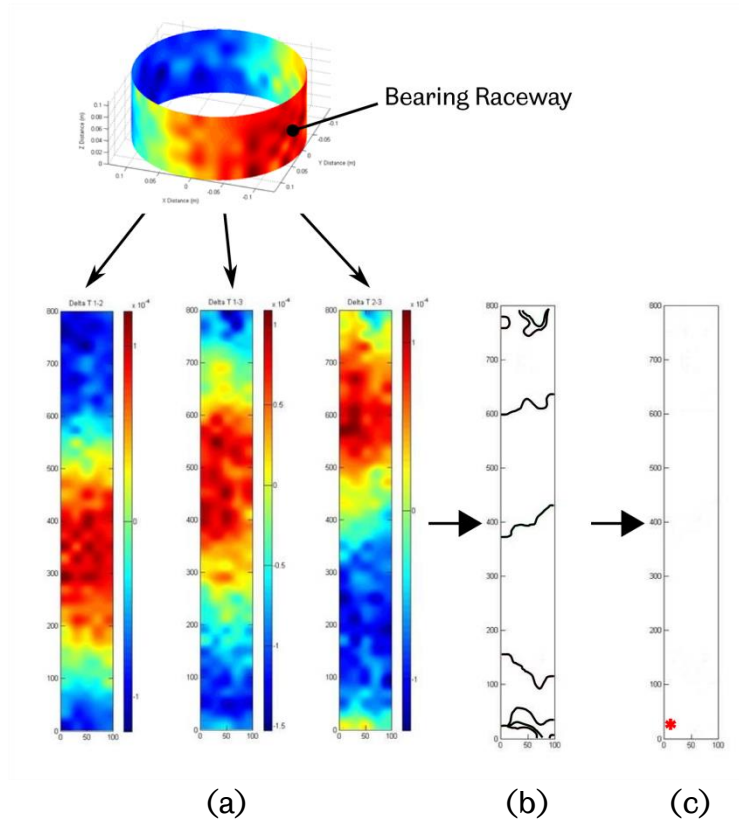


Figure 2.9: Location of a defect on a bearing raceway instrumented with three AE sensors (a) Intensity maps showing the difference in arrival times between sensor pairs 1 and 2, 2 and 3, and 1 and 3 (b) Possible locations based on signals from each sensor pair (c) ultimate defect location due to convergence in (b) (courtesy of J Naumann)

## 2.3.5 The Potential of Ultrasound

Ultrasound is investigated within this study as an additional means of detecting the onset of bearing failure. The technique – as will be discussed in more detail in the following chapter – has previously been proven for its ability to determine the thickness of the oil films that are formed within bearing contacts under EHL lubrication. Its applicability to the detection of early bearing faults however has not yet been extensively explored.

The measurements obtained through the application of ultrasound will relate to extremely localised regions of bearing surfaces, and will therefore focus specifically on the changes occurring at a contact level. It

is hypothesised that the onset of damage – which subsequently results in spalling to the rolling surface – will cause oil films to breakdown thus leading to metal-to-metal contact of bearing components. This should therefore be detectable by the ultrasonic sensor at the instance at which disruption of the oil film takes place thus allowing detection prior to the onset of noise and vibration. Alternatively however it may be the case that the breakdown of oil films takes place as a precursor to bearing failure, in which case the opportunity for early fault detection will be high, perhaps more so than all other techniques presented above.

## 2.4 Detection of Bearing Failures in Wind Turbine Gearboxes

A lot of the aforementioned techniques for bearing monitoring have not yet found widespread adoption in wind turbines, despite the high gearbox failure rates observed in the field and the strong evidence suggesting bearing failures to be the root cause (as discussed in the previous chapter). This is largely due to the costs associated with implementation of more advanced bearing monitoring techniques, as well as the fact that space within the gearbox is limited, thus making it difficult to directly instrument bearings in certain locations. The following section discusses the monitoring technologies currently utilised by the wind industry and the capabilities of such technology in detecting gearbox bearing failures.

Early wind turbines were installed with no remote monitoring capabilities whatsoever, but as adoption rates grew, and operators began to employ increasing numbers of turbines per site, the need for basic, centralised monitoring and control of various operational parameters became necessary. Supervisory Control and Data Acquisition (SCADA) systems were subsequently devised and installed, which now feature in most modern day wind turbines. SCADA systems have been standardised and therefore generally utilise similar control algorithms and monitoring techniques worldwide [61]. Data from each wind turbine on a farm is usually fed to a central control room where operational parameters and alarms can be viewed and various parameters can be controlled remotely. Given the high data throughput, particularly for larger wind farms, the data which is stored and output from each individual SCADA system must be minimised and therefore measurements are typically averaged over 10 minute windows to reduce data volumes [62].

Various parameters relating to overall turbine performance are monitored, including rotor and generator rotational speed, power output and general nacelle temperatures. Measurements specific to the gearbox usually only cover oil temperature, pressure and high speed shaft bearing temperatures, where alarms are also typically triggered based on any of these parameters exceeding a threshold [61]. Monitoring is therefore somewhat limited, which has led to research into more advanced system modelling techniques which can be applied to SCADA data to enhance fault detection and early warnings of failure on a component level.

Physical models of wind turbine drivetrain components were constructed in [63] to assess the energy balance and thus the expected heat output under normal operation. This was then used to define expected operating temperatures of components whose temperatures are monitored as part of the standard SCADA system. As such, temperature thresholds were then established which, if exceeded would indicate abnormal operation. These thresholds were applied to historic SCADA data to retrospectively predict the onset of a failure, and advanced warnings within the range of one month to two years were demonstrated.

Various researchers have applied artificial neural network (ANN) models to SCADA data to characterise the normal "healthy" behaviour of a wind turbine system [62, 64, 65]. Such techniques are used to "learn" the standard response of the system in terms of the available sensor outputs, such as temperatures and pressures. Once the baseline level is mathematically characterised, any deviation of an actual measured output from the expected model response can be used as an indication of failure. This has been widely demonstrated as being successful in providing an early warning of incipient system failure far in advance of that detected by the SCADA threshold-based warning system. It is generally noted however, that determination of the root-cause of a failure is not possible without prior knowledge of how measured outputs would typically respond to such specific component failures.

Overall, SCADA-based techniques are attractive since no additional data acquisition hardware is necessary and therefore the cost of adoption is negligible to the wind turbine operator. However, while general success in providing early warning of failure has been demonstrated, the methods can ultimately only go so far in determining the root-cause of failures due to available data having such low sampling rate and measured parameters being few and far between [66].

More advanced Condition Monitoring (CM) systems are commercially available for wind turbine applications but are always installed in addition to – and independent of – the traditional SCADA systems. A key drawback is the high cost, which has limited the uptake of such technologies in recent years [66]. These systems provide advanced monitoring capabilities through the installation of additional sensors upon specific subcomponents of the wind turbine including the generator, main rotor bearing, blades and gearbox. Additional sensor techniques applied specifically to the gearbox consist of [61, 66]:

- Bearing mounted thermocouples
- Oil particle counters
- Vibration sensors at certain bearing locations

Figure 2.10 illustrates the typical location of such sensors as part of a common CM system.
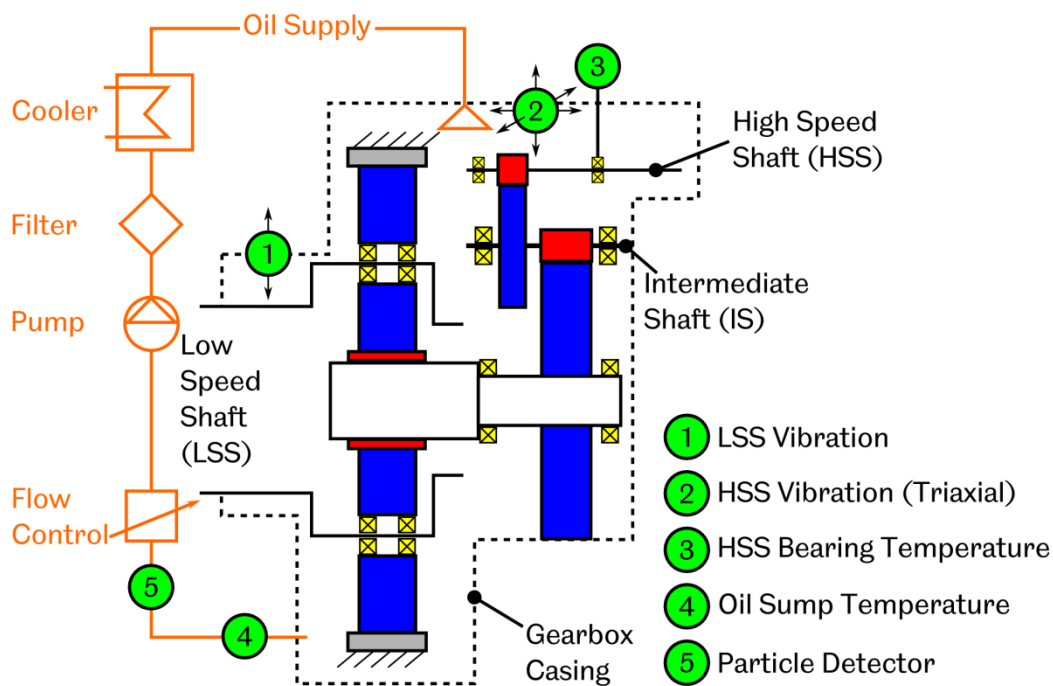


Figure 2.10: Standard WTGB sensors utilised in an advanced CM System (adapted from [61])

One of the key difficulties when dealing with more advanced sensor technologies in wind turbine applications is the highly variable operating conditions that gearboxes tend to experience. This means that signals such as vibration measurements will experience a high degree of variability associated with such fluctuating operating conditions alone, making it much more difficult to identify features corresponding to

bearing damage [66]. As such, it is essential that CM systems work in conjunction with the conventional SCADA system so as to associate signal changes to operating conditions or bearing damage appropriately, and reduce the occurrence of false alarms.

As is clear from Figure 2.10, even in advanced monitoring systems, placement of sensors is still limited to the gearbox outer casing and a relatively low number of sensors are generally employed. This is likely due to the fact that systems are often not installed solely for the purpose of *gearbox* monitoring, but for monitoring of the wind turbine system as a whole. As such, the addition of further gearbox sensors would only add to the overall costs of the CM system and place further demands on the acquisition hardware. Overall however, this will ultimately restrict the capabilities of such systems with regards to the detection of critical gearbox bearing failures (as discussed in §1.2.2) and reduce overall effectiveness. Indeed, it has been shown through use of more extensive arrays of gearbox casing mounted vibration sensors that detection of failures to specific sub-components can be successfully performed [67], thus potentially highlighting the importance of dedicated gearbox monitoring systems in the long term.

The additional costs associated with more advanced CM systems is clearly a limiting factor, but further difficulties arise with the shear volumes of data that such systems will generate. Indeed, the wind industry is somewhat unique in that any given wind farm must utilise a large number of individual turbines in order to achieve sufficient throughput. This in turn means that CM systems must be employed within a large percentage – if not all – of the turbines on any given site in order to make a significant impact on overall site performance. Clearly then, the development of robust and autonomous data analysis techniques will also be essential in order to deal with such "Big Data" appropriately; ensuring that critical events are correctly flagged, whilst minimising recurrent false alarms.

In parallel to the above, the applicability of more advanced sensing techniques such as AE and Ultrasound also need to be explored further, and thus the latter technique forms a key part of this particular work.

## 2.5 Conclusions

- Bearing failures can arise due to a range of operational factors however, if all means of best practice are employed, bearing failure may still arise due to material degradation in a process known as

Rolling Contact Fatigue (RCF). This is driven by the occurrence of sub-surface shear stresses within the bearing material and can cause cracks to form around material inclusions which eventually propagate to rolling surfaces.

- Bearing life rating calculations ($L_{10}$) allow the rolling contact fatigue life of a bearing to be predicted based on anticipated operating conditions. This is a statistical parameter which estimates the time for 10% of a sample of bearings to fail. Such calculations are ultimately based on the occurrence of a peak orthogonal shear stress induced within the bearing material, calculated based on classical Hertzian contact theory.

- Various methods are available for detecting bearing failures. Such methods involve the use of sensors to measure bearing operational parameters such as temperature, vibration, and the status of lubricating oils. Sensors can also be used to detect the emittance of high frequency sound waves within bearing materials which are generated in the event of crack formation and propagation (AE).

- AE monitoring has the potential to provide early warning of bearing failure, but the technique has yet to be utilised on a wide scale as an industrial condition monitoring technique, largely due to cost and the requirement for advanced data processing techniques.

- Despite the high rates of bearing failures observed in wind turbines, most modern systems do not currently employ advanced condition monitoring techniques that specifically focus on the detection of bearing faults at critical gearbox locations.

# Chapter 3

# THE MEASUREMENT OF OIL FILM THICKNESS USING ULTRASOUND

Having introduced current methods of bearing health monitoring in Chapter 2, this chapter introduces the Ultrasound Technique, which can be used to determine the thickness of the oil films that form within bearing contacts. As such, this technique offers a potential alternative to the previously introduced bearing condition monitoring methods, and may provide advanced warning of incipient bearing failure. Within this chapter the theoretical principles of ultrasonic reflection and its application to the determination of oil film thickness is described. The means of generating, acquiring and processing ultrasonic signals are introduced, with particular reference to the specific hardware used within this study. Finally, various examples are provided to illustrate previous applications of the Ultrasonic Technique for the determination of oil film thickness in various engineering components including rolling element bearings.

## 3.1 The Basic Principles of Ultrasound

Fundamentally, sound is the mechanical vibration of particles and the subsequently transferral to adjacent particles thus facilitating propagation. As such, sound requires a medium to travel through, whether that is of solid, liquid or gaseous form, and is therefore not supported in a vacuum. *Ultrasound* refers to such mechanical vibrations which occur at a frequency beyond that detectable by the human ear, specifically above around $20kHz$.

### 3.1.1 Propagation

All substances are comprised of particles interlinked via elastic or "electrostatic" bonds, which can be likened to springs as shown in Figure 3.1(a). A sound wave is generated in a medium through the excitation of a group of such particles via an external mechanical displacement. The elastic bonds attempt to resist such motion and force the particles back to their initial positions, which subsequently results in particle oscillations given that the bonds will inevitably have a finite stiffness. A sound wave is induced though a structure as a result of the transferral of such oscillations to adjacent particles via the elastic bonds.

The stiffness of the bonds determines the rate at which the sound wave is transmitted through the structure, where stiffer bonds transfer energy more efficiently and thus a faster rate of bulk sound transmission is achieved. This uniform transmission of vibrations through a group of particles in unison is referred to as an elastic wave, which implies that the particles eventually return to their initial positions and are therefore not permanently dislocated.

In solids, sound waves will propagate via four modes of oscillation; longitudinal, shear, surface and plate waves, the former two of which are generally used in ultrasonic testing. Within this study only longitudinal wave propagation is utilised, and the following discussion therefore relates solely to this particular wave mode. Here, particle oscillations occur in the same direction as the propagation of the bulk wave, and thus particle bonds experience tensile and compressive forces, as shown in Figure 3.1(b).



(a)                                    (b)
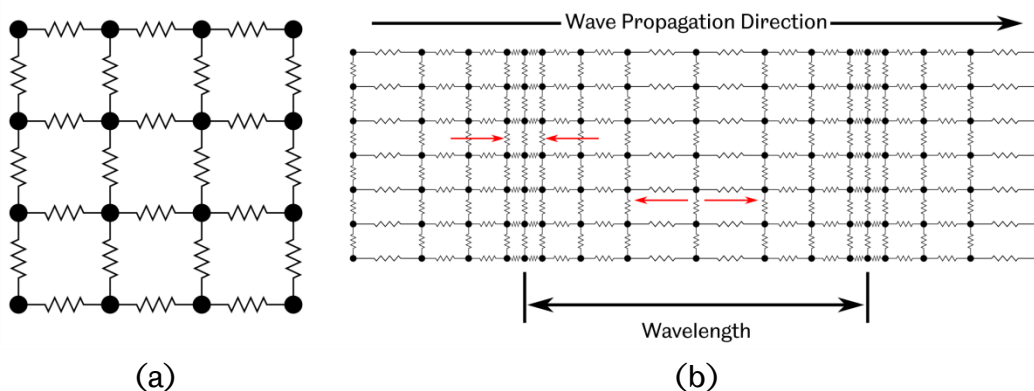
Figure 3.1: (a) Schematic of the inter-particle "spring-like" bonds of a structure (b) Illustration of the longitudinal propagation of a sound wave through particle oscillations

### 3.1.2 Material Properties Governing Ultrasonic Propagation

The propagation of ultrasound will vary according to the properties of the host material. The rate of propagation, the ability for the sound waves to penetrate the material, and how much ultrasonic energy is lost as a function of distance travelled are the most notably affected parameters. These are discussed in more detail within the following sections.

#### 3.1.2.1 Speed of Sound

As discussed above, the stiffness of the inter-particle electrostatic bonds within a structure will affect the rate of propagation of longitudinal ultrasonic waves. Speed of sound is temperature dependant, where speed will reduce as temperature increases. This is observed to a much greater degree in liquids and gases, whereas variation is much lower in solid materials. Speed of sound also varies with pressure, again to a much lesser degree in solids, where pressures in the order of $GPa$ are usually required to induce such changes.

Speed of sound can be determined through a "time-of-flight" experiment where ultrasound is transmitted through a material of known thickness. Through observation of the time taken for the sound wave to reverberate through the material the speed of sound, $c$, is calculated according to the basic velocity equation.

$$c = \frac{2x}{t}$$

<div align="right">(3.1)</div>

Here a pulse-echo measurement is assumed, and thus twice the material thickness ($x$) must be considered since the ultrasonic wave must travel through the material and back to the ultrasonic emitter/receiver.

Knowledge of the speed of sound then permits determination of the ultrasound wavelength, $\lambda$, at a given frequency, $f$.

$$\lambda = \frac{c}{f}$$

<div align="right">(3.2)</div>

### 3.1.2.2  Acoustic Impedance

Acoustic impedance is a measure of the ability of a sound wave to enter a material from another. This governs how much an ultrasound wave will be reflected at an interface, which will be of key relevance to this study, and will be discussed in more detail later in this chapter. Acoustic impedance is typically denoted $z$, and is a function of the speed of sound within a material and its density, $\rho$.

$$z = \rho c$$

(3.3)

Acoustic impedance is measured in $Ns/m^3$, also termed the $Rayl$, where $\rho$ and $c$ are measured in the SI units of $kg/m^3$ and $m/s$ respectively.

### 3.1.2.3  Attenuation

Attenuation is the rate at which the amplitude of an ultrasonic wave is reduced as it travels through a material. Such reduction in ultrasonic energy may arise in a number of ways.

**Radiation**

Radiation is a reduction in energy density due to divergence of an ultrasonic wave as it travels away from its source. As energy is emitted into a material from an ultrasonic source and is subsequently reflected back from the opposing interface the ultrasonic beam will spread, meaning that less energy is received across the area of the emitter/receiver on return.

**Absorption**

As discussed earlier, the propagation of ultrasound through a material relies on the oscillation of particles. Not all vibrational energy is transferred from one particle to the next however, and some is lost as frictional heat, which is said to have been "absorbed" by the host material. This is frequency dependant, where higher frequency sound waves will experience higher attenuation via absorption. This is due to the inevitable increase in particle velocity at higher frequencies, thus increasing the frictional heating that takes place.

**Scattering**

Scattering of ultrasonic waves will also occur as sound travels through a material containing flaws and grain boundaries. These will be randomly distributed through a material, and thus the scatter of an ultrasound wave is unavoidable and will vary in all structures, regardless of the composition. However, scattering generally only becomes an issue when the size of a discontinuity is comparable to or greater than the wavelength of the incident ultrasound wave. In the case of this particular application, ultrasonic waves of $10MHz$ centre frequency will be utilised extensively and applied to bearing steels with a speed of sound in the order of $5900m/s$. This therefore equates to an ultrasound wavelength of around $0.5mm$ according to Equation (3.2). It is considered unlikely that voids of this magnitude will exist within a high quality bearing steel, and therefore the effects of scattering should be relatively small.

# 3.2   Generating Ultrasound

The generation of ultrasound in a material requires the initiation of particle displacement at an appropriate oscillating frequency. This can arise from impacts or in the event of crack propagation. Indeed, the latter represents what was termed an "Acoustic Emission" in the previous chapter, which is essentially a passive ultrasound technique whereby sensors "listen" for the emittance of ultrasonic energy from within a material. In the application of what is herein termed the "Ultrasound Technique" however, it is necessary to stimulate such acoustic energy artificially and in a more controlled manner. Furthermore, the two examples stated above are destructive initiation mechanisms, where the propagation of the emitted ultrasonic wave will still be elastic, but such energy has ultimately emanated from a local region of the host material where permanent deformation has taken place.

Alternatively, the piezoelectric effect provides the means to generate such excitation without the associated local damage.

## 3.2.1  The Piezoelectric Effect

The piezoelectric effect is a phenomenon that occurs in certain materials and involves the generation of electrical charge when such a material is placed under mechanical deformation. The reverse is also true such that elastic deformations can be induced in materials under electrical

excitation. This effect tends to occur in materials with a crystalline or semi-crystalline structure, including crystals and various ceramics, but is also known to occur in human bone.

Consequently, controlled and repeatable displacements can be generated effectively through application of an oscillating electric field to a piezoelectric material. Given the reversible nature of the phenomenon, piezoelectric materials can be used to both emit an ultrasonic wave and measure the corresponding reflections received back from the opposing material interface.

### 3.2.2 Ultrasonic Transducers

A standard commercial ultrasonic transducer is illustrated in Figure 3.2.
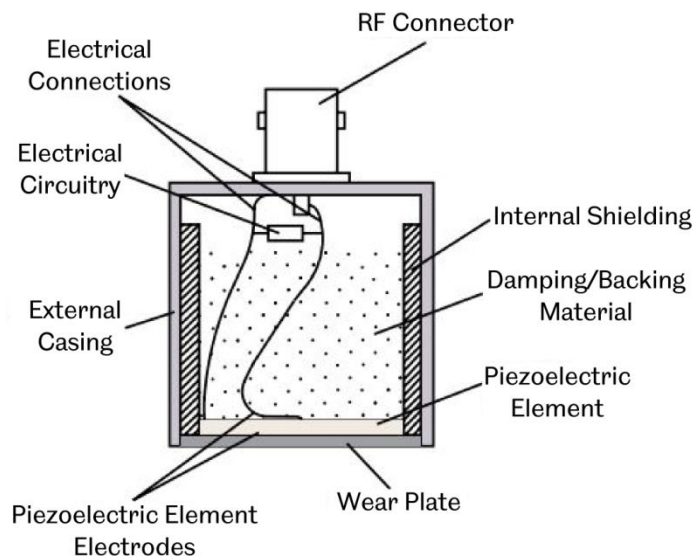


Figure 3.2: Composition of a standard commercial Ultrasound transducer (reproduced from [68])

All electronics are housed in a casing, typically metallic, to protect the system components and provide electrical shielding. A radio frequency (RF) connector allows for extraction of the transducer signal to an appropriate data acquisition system, where the same signal cable is also used to supply the initial electrical current to energise the piezoelectric element. The piezoelectric element itself is situated on the bottom face and ultrasonic energy is thus generated downwards with reference to the orientation in Figure 3.2. The element is mounted to a wear plate to protect it from the external environment.

A backing material is used to dampen the oscillations of the piezoelectric element so that it does not continue to resonate for long periods following application of the excitation pulse. Backing materials are usually highly attenuating and dense so as to restrict any reverse propagation of ultrasonic energy into the transducer itself.

Such transducers have found widespread adoption for a variety of engineering applications including flaw detection, thickness gauging and the assessment of weld integrity; however there are a number of drawbacks with regards to their use in bearing applications such as that considered within this study. Firstly, transducers are usually relatively large, and therefore require significant modifications to be made to bearing housings in order to accommodate them. Furthermore, transducers require a coupling medium to enable the effective transmission of ultrasonic energy into the bearing material. While both of the above have been achieved successfully in laboratory settings ([69] for example), this can often be impractical in industrial applications and on bearings residing within complex machines. Finally, transducers usually have limiting operating temperatures in the order of around $50°C$, which restricts their use in many industrial bearing applications.

### 3.2.3 The Use of Bare Piezoelectric Elements

The above limitations of commercial transducers can be avoided through use of the bare piezoelectric elements alone, which are small, can be bonded directly to bearing raceways – thus providing a permanent ultrasonic coupling – and have operational temperature limits imposed only by the piezoelectric material itself, which is usually in the order of hundreds of degrees Celsius.

Figure 3.3(a) illustrates the $10MHz$ wrap-around type piezoelectric elements used within this work, which are of $7.1mm$ diameter and $0.2mm$ thickness. The wrap-around designation means that the lower electrode wraps around the piezoelectric element onto the top surface such that external wiring can be connected from the top face only. Since the piezoelectric material requires a voltage differential to be applied across it, the active region of a wrap-around type element will not cover the entire diameter and will equate roughly to the region as designated in light blue in Figure 3.3(b). The piezoelectric elements may be reduced in size to control the effective width of the emitted ultrasonic wave-front. This can be achieved provided that both electrodes are maintained, and

therefore dictates that cuts are made in the orientation illustrated by the red lines in Figure 3.3(b).



<div align="center">(a)                (b)</div>
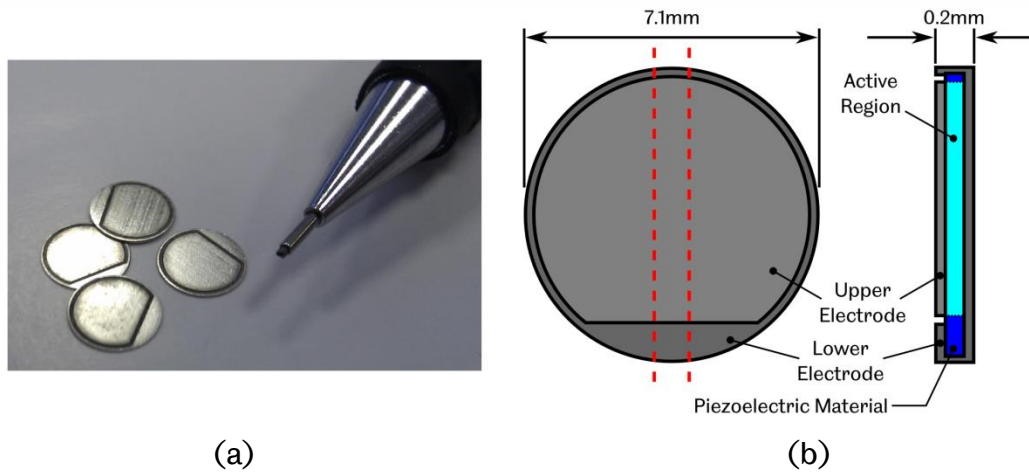
Figure 3.3: (a) Photograph of the $10MHz$ piezoelectric elements used within this work, (b) associated dimensions and illustration of the active region

Piezoelectric elements can then be bonded to a test specimen, either in their full disc form, or as sliced sections. It was identified in previous research [70] that high temperature strain gauge adhesives work particularly well to provide a good acoustic bond of the element to the host material, which is well maintained at high temperatures. Once appropriately bonded, coaxial cable is then soldered to the element to permit application of the excitation signal and extraction of the subsequent voltage response. The element is then typically coated or "potted" in a high temperature epoxy to provide protection for the element and soldered junctions, a certain degree of strain relief for the cabling, and to dampen the oscillations of the piezoelectric element. Through this, the key functions of a commercial ultrasonic transducer (as discussed in the previous section) are replicated but on a much smaller scale.

## 3.3  Ultrasonic Reflection from an Interface

When ultrasound is applied normal to an interface comprising two media of different acoustic impedances a portion of the incident ultrasonic wave will be reflected, and the rest will be transmitted through. Indeed, for a "perfect" interface, the ultrasonic reflection is characterised by Equation (3.4) [71].

$$R = \frac{z_1 - z_2}{z_1 + z_2}$$

<div align="right">(3.4)</div>

Where, as shown in Figure 3.4(a), $z_1$ and $z_2$ represent the acoustic impedances of the two materials within which the ultrasound is supported. $R$ is termed the "Reflection Coefficient," and represents the fraction of the incident ultrasonic energy $I$, that is reflected from the interface, which can evidently exist within the range $-1 \leq R \leq 1$, where $R = 0$ denotes a perfectly bonded, identical material couple, which would not induce any variation in the incident ultrasonic energy. A positive or negative sign indicates the phase of the reflected energy, the former being equal to and the latter being 180° out of phase of the incident wave [72]. It follows that the proportion of ultrasonic energy transmitted through the interface, $T$, is thus equivalent to:

$$T = 1 - R$$

<div align="right">(3.5)</div>
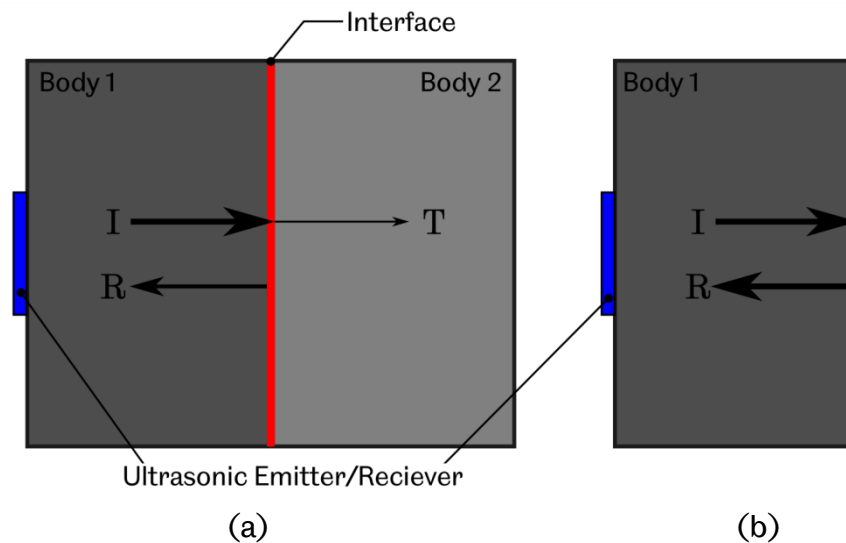


Figure 3.4: (a) Ultrasonic reflection and transmission from a perfect interface (b) Complete reflection from the boundary of a medium and a vacuum

The scenario illustrated in Figure 3.4(b) represents a case where $z_2 = 0$ and is therefore a vacuum, inducing full reflection of incident energy at the interface. It is useful to obtain a measure of reflected energy in such cases for real ultrasonic measurements, since this provides an indication of full ultrasonic energy available at a surface of interest, as emitted and

subsequently received back by an ultrasonic element. Then, in the presence of the measurement interface (as per Figure 3.4(a)), $R$ is established as the fraction of the reference energy, according to Equation (3.6), and thus requires no prior knowledge of the incident ultrasonic energy level $I$, or the level of attenuation through the host medium.

$$R = \frac{R_{(a)}}{R_{(b)}} \equiv \frac{R_{meas}}{R_{ref}}$$

(3.6)

In reality, the situation in Figure 3.4(b) is extremely difficult to obtain for practical measurements, and a condition where incident ultrasonic energy strikes an air boundary is deemed sufficiently close to approximate the full ultrasonic reference state. Indeed, for the purposes of this study it is assumed that this scenario is in fact achieved with a steel-air interface. In reality, this actually represents a reflection coefficient of $R \approx 0.99998$ (according to Equation (3.4)), that is; 99.998% reflection of incident energy, which is generally considered sufficient when applying the ultrasound technique in practical engineering scenarios.

### 3.3.1  Application to Real Engineering Contacts

The interface illustrated in Figure 3.4(a) will never truly exist due to real surfaces never being perfectly flat and being comprised of asperities, as introduced briefly in the previous chapter. As such, contact between the two surfaces will be intermittent, and will consist of regions of direct metal-to-metal contact as well as voids containing the surrounding fluid, typically air or oil. This is demonstrated in Figure 3.5(a), where surface
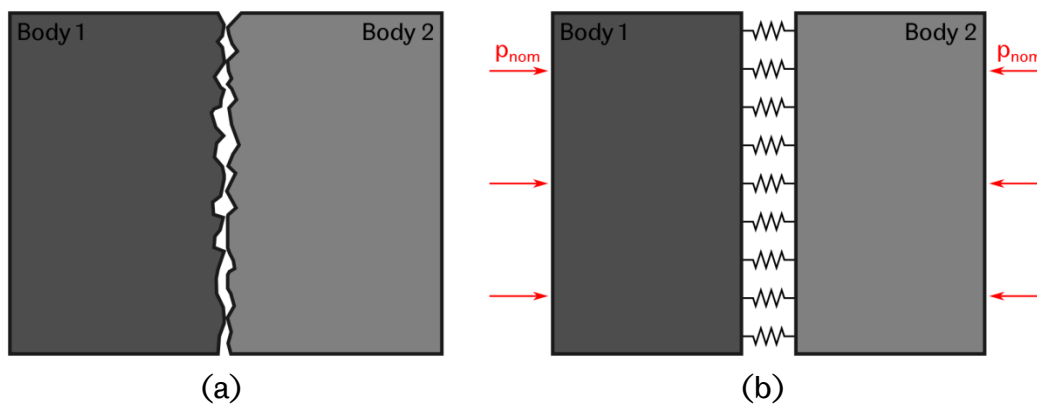


Figure 3.5: Illustration of (a) a real contact scenario, where surface asperities result in only partial contact across the interface, (b) the spring like behaviour of the interface under an applied normal contact pressure

profiles have been exaggerated for illustrative purposes.

This results in the interface displaying a spring-like behaviour when a normal pressure is applied, as illustrated in Figure 3.5(b). The effective "interfacial stiffness" is dictated by the number of discrete asperity interactions within the contact and the deformations of such asperities under load, which in turn causes a rate of approach of the two interfacing surfaces.

In reality then, the ultrasonic reflection from an interface is not truly represented by Equation (3.4), but is in fact found to be characterised as a function of the effective interfacial stiffness, $K$, in combination with the acoustic impedances of the contacting bodies and the characteristics of the incident ultrasonic wave [71].

$$R = \frac{z_1 - z_2 + i\omega\left(\frac{z_1 z_2}{K}\right)}{z_1 + z_2 + i\omega\left(\frac{z_1 z_2}{K}\right)}$$

(3.7)

This relationship is termed the "spring model" and forms a fundamental part of the ultrasonic technique as applied to the contact of engineering components. Here, $\omega$ represents the angular frequency of the ultrasonic wave incident upon the interface, where $\omega = 2\pi f$. The interfacial stiffness is derived from the change in contact loading $\delta Q$ and the approach of the mean lines of roughness of the two surfaces, $\delta u$ [73].

$$S = -\frac{\delta Q}{\delta u}$$

(3.8)

Which can be expressed as a function of the nominal contact pressure, $p_{nom}$, such that the interfacial stiffness per unit area $K$ $(GPa/m)$ becomes:

$$K = -\frac{\delta p_{nom}}{\delta u}$$

(3.9)

Indeed, it can be observed that as an interface moves closer to that of a "perfect" contact, as illustrated in Figure 3.4(a), the interfacial stiffness tends to infinity, and Equation (3.7) becomes Equation (3.4). Clearly, Equation (3.7) contains an imaginary component $i$, and thus represents both the amplitude and phase of the ultrasonic reflection vector. Physical

measurements will therefore relate to the magnitude of the reflection coefficient vector, $|R|$, established from trigonometric rules such that [74]:

$$|R| = \sqrt{\frac{(z_1 - z_2)^2 + \left(\frac{\omega z_1 z_2}{K}\right)^2}{(z_1 + z_2)^2 + \left(\frac{\omega z_1 z_2}{K}\right)^2}} \qquad (3.10)$$

Furthermore, if the two materials comprising the interface are identical, such that $z_1 = z_2 = z$, then the magnitude of $|R|$ becomes:

$$|R| = \frac{1}{\sqrt{1 + \left(\frac{2K}{\omega z}\right)^2}} \qquad (3.11)$$

## 3.3.2 Application to Oil Films

The above relations were also found to be applicable to contact interfaces separated by thin oil film layers [75]. Here, opposing surface asperities would no longer directly interact due to the separation resulting from the oil film, and ultrasonic reflection was found to relate to the stiffness imposed by the fluid instead, such that Equation (3.7) was still valid. In such contact scenarios the effective stiffness, $K$, would therefore relate to the compressibility of the oil according to Equation (3.12).

$$K = \frac{B}{h} \qquad (3.12)$$

Where $B$ is the bulk modulus of the oil layer and $h$ is the lubricant film thickness. Under ambient conditions the speed of sound through a lubricant film is said to relate to the fluid density and the bulk modulus according to [76]:

$$c = \sqrt{\frac{B}{\rho}} \qquad (3.13)$$

And therefore, the lubricant stiffness (Equation (3.12)) can be expressed in terms of easily definable fluid parameters and the lubricant film thickness:

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

(3.14)

It should be noted that the deviation of Equation (3.13) when conditions stray from that of ambient temperature and pressure complicate the determination of $K$. Additional considerations and measurements will often be required, as will be discussed later in this work.

Nonetheless, substitution of Equation (3.14) into both Equations (3.10) and (3.11) and rearranging, leads to the definition of oil film thickness, $h$, in terms of the measured ultrasonic reflection.

$$h = \frac{\rho c^2}{\omega z_1 z_2} \sqrt{\frac{|R|^2 (z_1 + z_2)^2 - (z_1 - z_2)^2}{1 - |R|^2}}$$

(3.15)

For the case when the oil layer separates two different materials, and:

$$h = \frac{2\rho c^2}{\omega z} \sqrt{\frac{|R|^2}{1 - |R|^2}}$$

(3.16)

When the materials either side of the lubricant layer are identical ($z_1 = z_2 = z$).

### 3.3.3 Spring Model Assumptions

In application of the spring model to lubricated contacts, two key assumptions are made:

1. The lubricant layer acts to completely separate the two bounding surfaces such that no metal-to-metal contact occurs.
2. The surfaces either side of the lubricant layer – and indeed the lubricant layer itself – are parallel across the ultrasonic wavefront.

It has already been discussed that, in fact, surfaces are rough and therefore the latter of the two above points will never truly be achieved. Often however, the oil film thickness is significantly larger than the undulations in the surrounding surfaces such that, in a relative sense, the interface will appear smooth to the incident ultrasonic wave.

Indeed, the degree of surface separation achieved by a lubricant in an EHL contact may be assessed through use of the Lamda Ratio, which considers the ratio between the theoretical film thickness and the composite roughness of the two surrounding surfaces.

$$\Lambda = \frac{h}{\sqrt{R_{q1}^2 + R_{q2}^2}}$$

(3.17)

Here, $R_{q1}$ and $R_{q2}$ represent the root mean square roughness of the two contacting bodies. Therefore, with regards to the first of the above assumptions, the spring model is applicable for scenarios where $\Lambda$ is significantly higher than unity. For $\Lambda \leq 1$, the contact is said to be boundary lubricated (also termed a "mixed" lubrication regime), and as $\Lambda$ tends to 0, asperity interactions are increased and metal-to-metal surface interactions dominate the contact. In such cases, the spring model for lubricated contacts will not be applicable.

### 3.3.4 Application to Solid and Mixed Contacts

Previous work has demonstrated applicability of the spring model to dry contacts, leading to determination of contact pressure based on the relationship identified in Equation (3.9) [77]. A calibration experiment is initially required however to relate the change in interfacial stiffness to absolute contact pressure, and thus ultrasonic measurements must initially be recorded from an interface of the same material couple under known load and constant contact area, such that the obtained reflection coefficient can be used to infer interfacial stiffness (according to Equations (3.10) or (3.11)) and thus relate this directly to pressure. Conveniently, it was found that the relationship between contact pressure and interfacial stiffness is linear for moderate contact loads [78].

In a mixed regime, film thickness is insufficient to fully separate the contacting surfaces, and the contact will consist of regions of direct metal-to-metal contact, but also regions where the surface is fully separated by

the lubricating oil. It was shown in [79] that the spring model may still be applied in such cases, but where a composite stiffness is utilised, $K_T$, which was defined as the sum of both the solid, $K_S$, and lubricant layer, $K_L$, stiffnesses:

$$K_T = K_S + K_L \qquad (3.18)$$

Or, given Equations (3.9) and (3.12):

$$K_T = \frac{B}{h} - \frac{\delta p_{nom}}{\delta u} \qquad (3.19)$$

Further analytical work in [80] found the solid stiffness component of a mixed regime contact to contribute much less to the overall composite stiffness $K_T$, owing to the very high effective stiffness of the lubricant layer due to it being extremely thin. Ultrasonic measurements taken within the same study were used to extract both stiffness components from mixed regime contacts, and corresponding film thickness values were found to agree well with theoretical results.

# 3.4  Application of the Oil Film Thickness Technique to Bearings

With the theoretical principles introduced, the remainder of this chapter discusses the application of ultrasound to bearings, and the steps required to determine oil film thickness.

## 3.4.1  Ultrasonic Reflections within a Bearing Raceway

To facilitate measurements in rolling element bearings, piezoelectric elements are bonded directly to bearing raceway outer surfaces to generate ultrasonic waves that will propagate through to the rolling contact surface. This sets up a sequence of reverberations within the raceway body as the elastic wave reflects from the rolling surface to the ultrasonic source element and back again. The number of reverberations that are observed by the source element is limited by the attenuation of ultrasonic energy through the raceway material.

Figure 3.6(a) illustrates a measurement instance where the roller-raceway contact is not in line with the ultrasonic element. As such, a steel-air interface is observed, and it is assumed that all incident energy is reflected from the rolling surface back to the source element. Generally, it is sufficient to only observe one single rolling surface reflection, where the first is typically chosen due to its optimal signal-to-noise ratio (SNR).



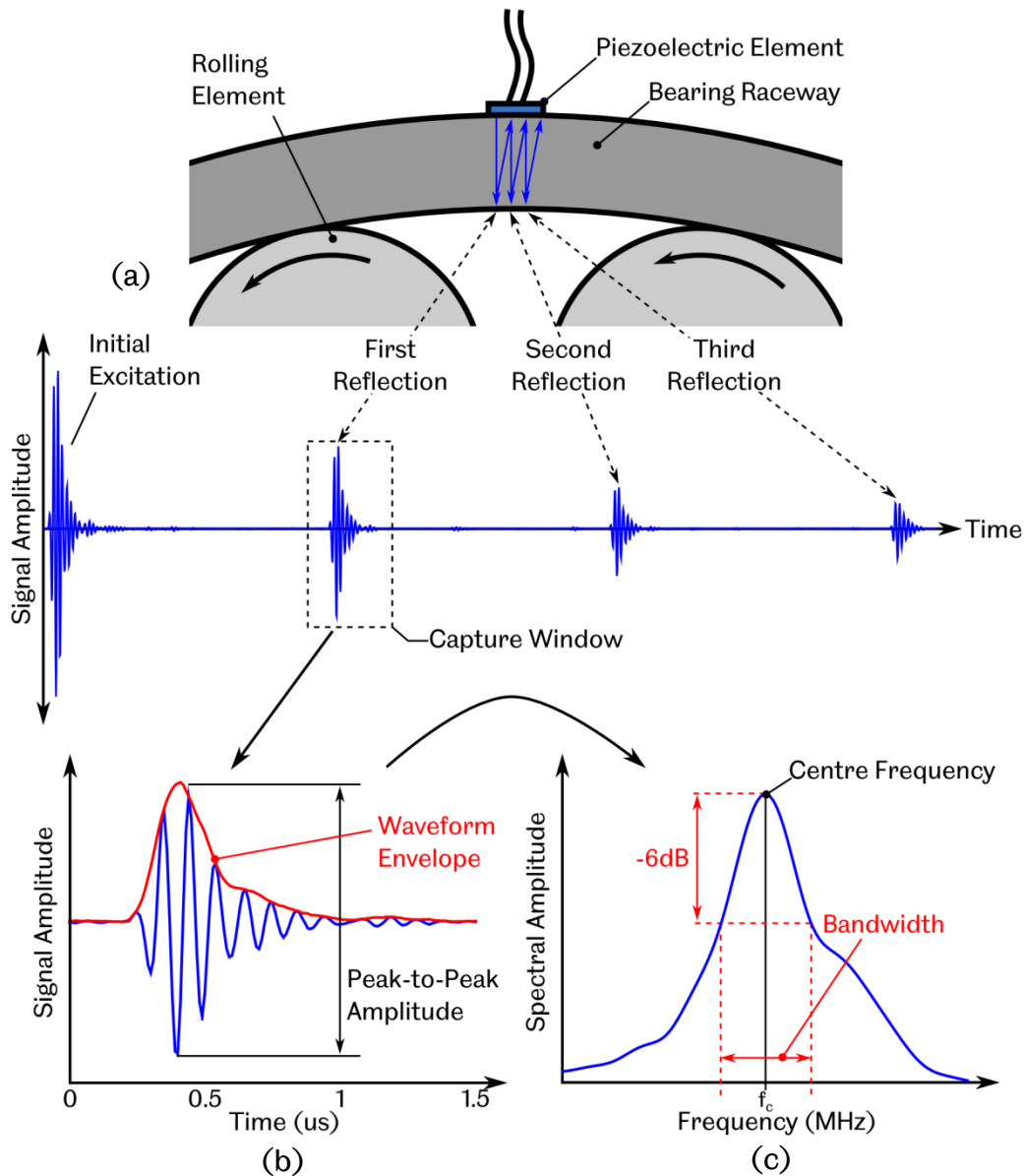Figure 3.6: (a) Illustration of the ultrasonic reflections generated through a bearing raceway and further detail of the first reflection in both the (b) time and (c) frequency domain

The first reflection is illustrated in more detail in Figure 3.6(b) and (c) in both the time and frequency domain respectively. The peak-to-peak amplitude of the time domain pulse is a measure of the overall signal

60

strength and can be used to determine the reflection coefficient, $|R|$. However, the change in the peak-to-peak amplitude of the signal in the presence of an oil film will represent the average response across the entire bandwidth of the sensor. Given the frequency dependence of Equations (3.15) and (3.16), it is usually more appropriate to analyse the frequency content of the reflection pulse and observe the change in spectral amplitude at a given frequency. In the frequency domain it can be observed that the reflected pulse has a narrow bandwidth, where frequency content is centred around the central frequency of the piezoelectric source element, $f_c$. As shown, the bandwidth is generally defined by a $-6dB$ drop in signal amplitude either side of the resonant peak.

### 3.4.2 The Ultrasonic Measurement System

#### 3.4.2.1 Hardware

A PC based ultrasonic data acquisition system was used to perform the measurements within this study. The system, termed a "Film Measurement System," or "FMS," (*Tribosonics Ltd.*) comprised of two bespoke PCI cards to facilitate ultrasonic measurements as shown in Figure 3.7.
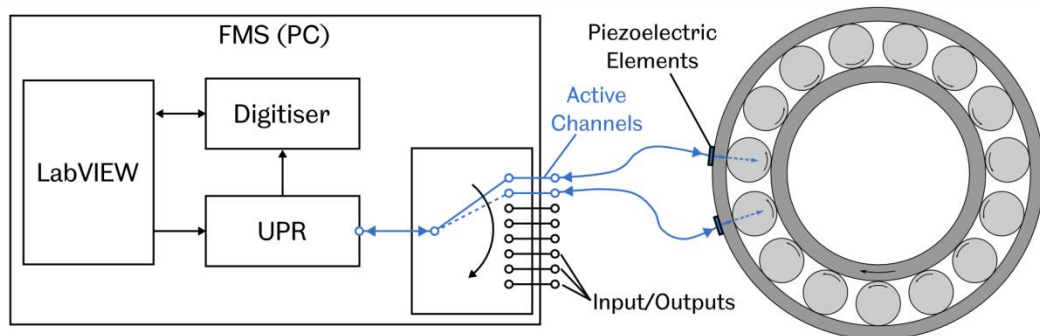


Figure 3.7: Schematic of the ultrasonic data acquisition hardware

Voltage pulses were generated by the Ultrasonic Pulser Receiver (UPR) to excite the piezoelectric elements and induce ultrasonic sound waves. These short duration pulses could be generated at a software selectable rate of up to $80kHz$. This is termed the "Pulse Repetition Rate" (PRR), and essentially controls the rate at which ultrasonic pulses are generated within the test specimen. The UPR could deliver excitation pulses at up to $300\,Volts$ and the frequency, or "Pulse Width" (PW), of the excitation

signal could be varied to suit the resonant frequency of the piezoelectric element being used $(1-20MHz)$. The UPR was capable of delivering such excitation pulses to up to 8 individual piezoelectric elements at any one time. However, the excitation could not be performed simultaneously on each and would essentially be switched between all "active channels" sequentially, meaning that the $PRR$ would be split based on the number of channels that were active (i.e. a PRR of $80kHz$ with 2 active channels would equate to a real pulse rate of $40kHz$ per channel).

The additional function of the UPR was to receive the subsequent voltage responses from the piezoelectric elements and pass these to the analogue-to-digital converter, or "Digitiser" for storage on the PC. The digitiser would sample the ultrasonic waveforms at $100MHz$ to a resolution of 12 bits. The digitised waveforms were then passed to the control software for viewing and storage.

### 3.4.2.2 Acquisition and Storage

Control of ultrasound generation and data acquisition was achieved through a *LabVIEW* control interface. This allowed for manipulation of general UPR settings such as the PRR and excitation voltage, as well as channel-by-channel signal configurations through adjustment of the PW, as well as "Range" and "Delay" which could be selected so as to "window" the appropriate waveform corresponding to the reflection of interest, as illustrated in Figure 3.6(a). Furthermore, gains could be applied to each channel individually so as to amplify waveforms for clarity. This was essentially a multiplier applied to the acquired data and would therefore also act to increase signal noise, this could however be used to achieve comparable peak-to-peak amplitudes across all active channels.

Given the high PRR's that could be applied, as well as the high digitisation rates, large data volumes could be generated rapidly. In order to enable all ultrasonic waveforms to be stored at the highest applicable pulse rates, waveforms were stored temporarily within hardware buffers upon acquisition, which could then be viewed and removed from the buffers at a much slower rate by request via the acquisition software. As such, ultrasonic waveforms would be obtained in a stacked time series or "data stream," and knowledge of the PRR and signal range would be required in order to easily separate out the data and apply appropriate timing information in post-processing (Figure 3.8).
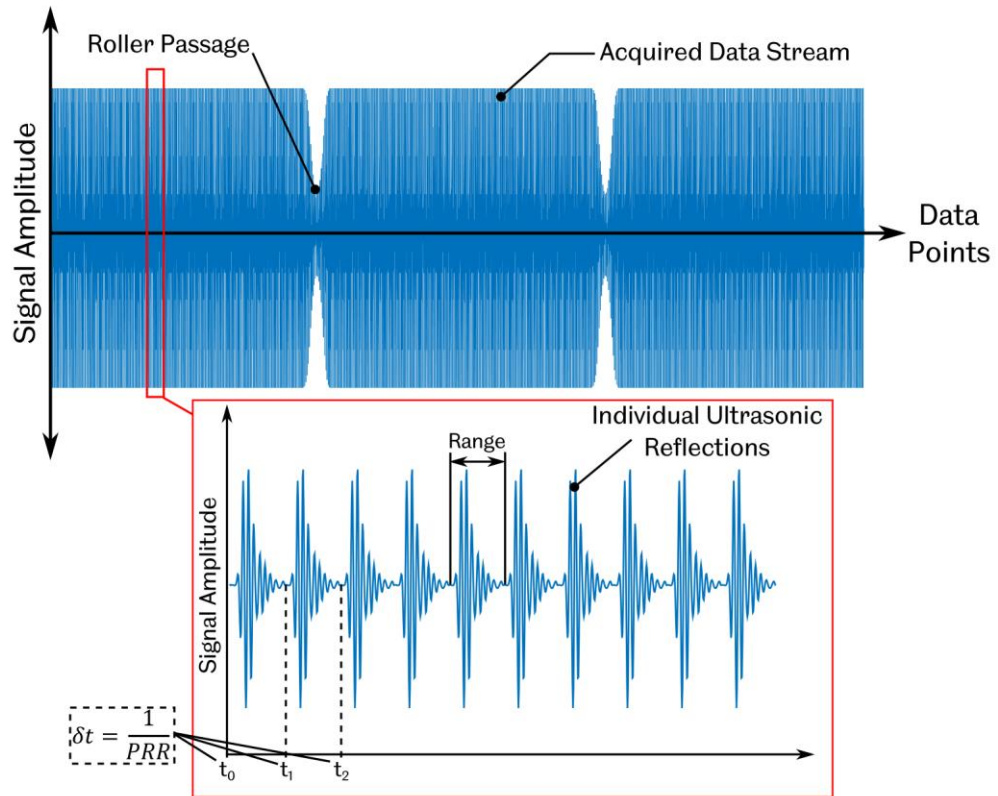
Figure 3.8: Illustration of a stacked ultrasonic data stream when stored to disc by the data acquisition system

### 3.4.3 Applying the Spring Model to Measurement Data

As discussed in §3.3, application of the spring model essentially requires determination of the reflection coefficient from the measurement data which is established according to Equation (3.6). This is reproduced in below to reflect the physical quantities of the ultrasonic reflection pulse used to determine the reflection coefficient $|R|$.

$$|R| = \frac{A(f)_{meas}}{A(f)_{ref}} \tag{3.20}$$

Here, $A(f)_{meas}$ is the spectral amplitude of the ultrasonic reflection from the oil layer, and $A(f)_{ref}$ is the *reference* spectral amplitude, which is the amplitude of the ultrasound wave if it were completely reflected at the measurement interface ($R = 1$). As such, $A(f)_{ref}$ represents the total ultrasonic energy available at the rolling surface. As discussed in §3.3, this is only truly achieved in a vacuum, but is said to be represented with

reasonable accuracy by a steel-air interface, which is indeed much more attainable in practice.

In fact, in applications such as that of rolling element bearings where an ultrasonic sensor will only ever observe an oil film intermittently as rolling elements move across its line-of-sight, a steel-air reference condition is always readily available. This means that, provided data streams are sufficiently long to capture reflections from the measurement interface as well as those when the measurement interface is not visible (as illustrated in Figure 3.8), a *reference* waveform will always exist alongside a *measurement* signal. This *real-time reference* is particularly useful, since any changes in the sensor response that may arise as a result of temperature variations during operation will automatically be compensated for due to the fact that variations will exist in both the *measurement* and *reference* signals.

### 3.4.3.1  Obtaining the Modal Reference Waveform

The most practical way to extract an appropriate reference from a given data stream is to determine the "modal waveform" of the data set. This process is illustrated in Figure 3.9, and is performed by analysing the signal amplitude at a given time step in each ultrasonic pulse within the data
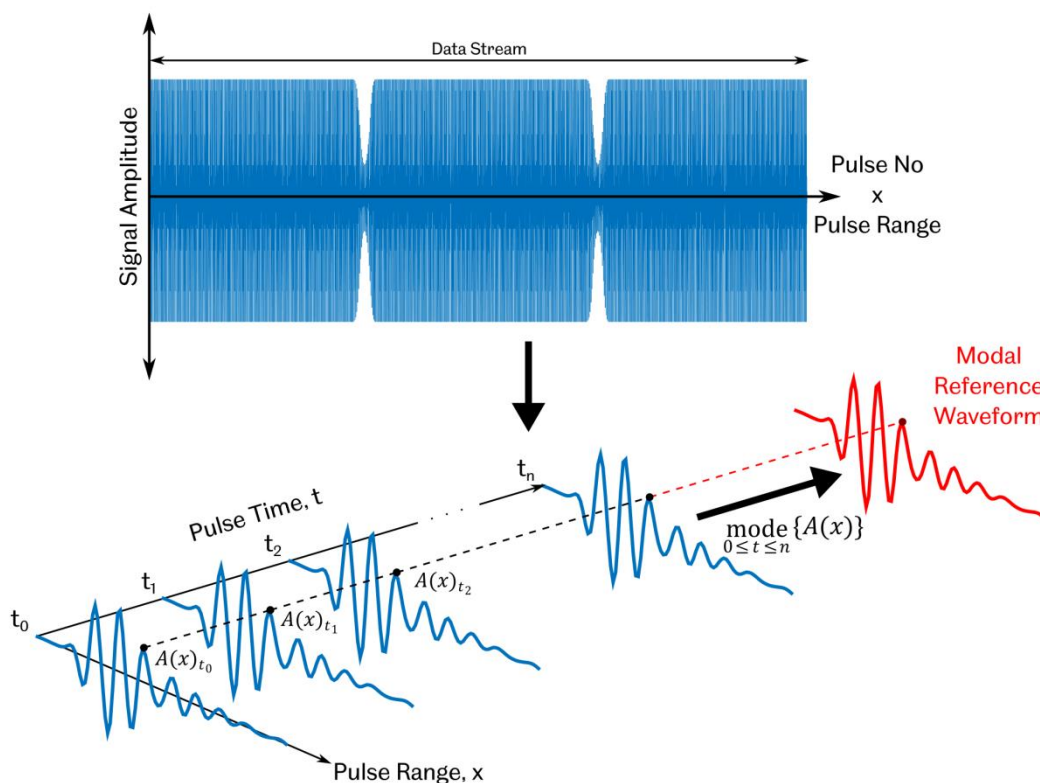


Figure 3.9: Obtaining the modal reference waveform from an ultrasonic data stream

stream and determining the single most frequently occurring value (mode) for each such time step. From this, a single waveform can be obtained which reflects the most commonly occurring reflection pulse within the entire data stream.

However, in situations where the measurement interface is observed very frequently, the PRR of the data is relatively low, or the measurements are particularly noisy, this procedure may not generate an accurate representation of the reference reflection. In such cases it may be necessary to apply a filter on the peak-to-peak amplitude of each reflection pulse initially, so as to only utilise those reflections that fall within a certain amplitude range when determining the reference signal (as described in [72]). In this study however, the "Modal Reference" method was found to be sufficient and was thus utilised throughout.

### 3.4.3.2 Calculating the Spectral Reflection Coefficient

Once the reference waveform is obtained from the ultrasonic data stream, this can then be compared to the ultrasonic reflections in the presence of the oil film, and a reflection coefficient obtained according to Equation



Figure 3.10: Application of Equation (3.6) from an ultrasonic measurement signal obtained from an example bearing measurement

(3.6). This procedure is illustrated in Figure 3.10.
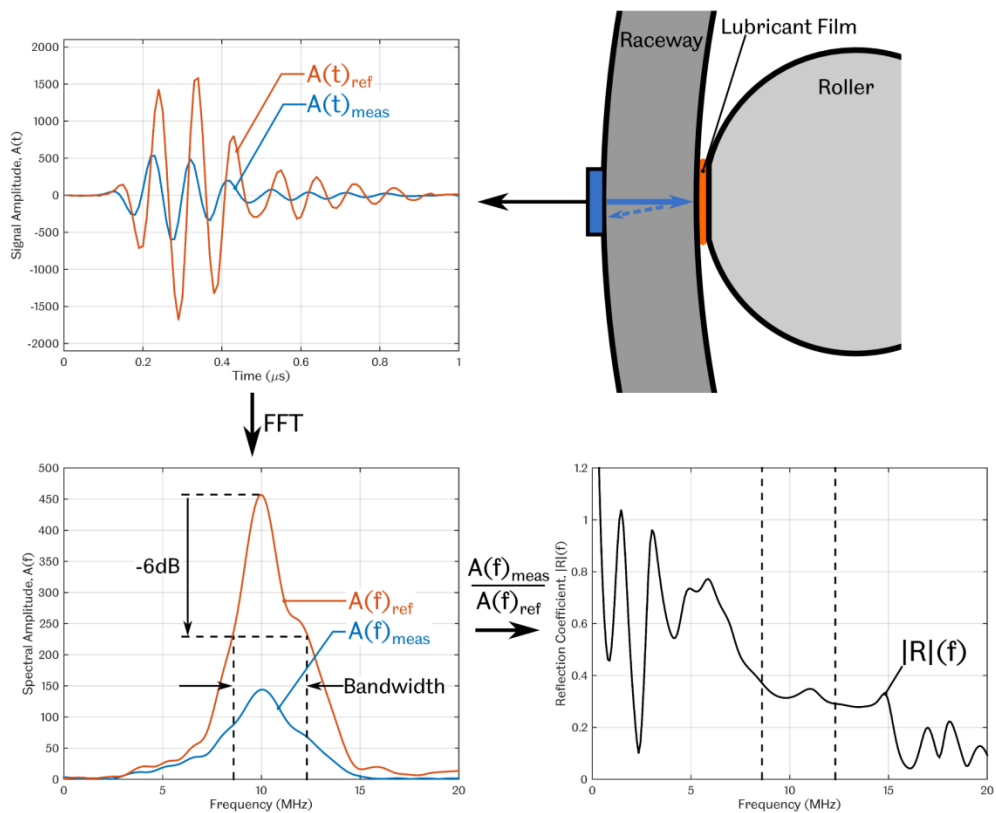
It is often necessary to zero-pad time domain signals before converting to the frequency domain due to reflections typically having very few data points. This has the effect of increasing the resolution of the resultant frequency spectra as calculated by the Fast Fourier Transform (FFT) algorithm. Following this, Equation (3.6) can be calculated point-by-point from the frequency spectra of the reference and measurement signals, resulting in definition of the reflection coefficient, $|R|$, as a function of frequency. Those reflection values lying within the bandwidth of the signal can then be converted to film thickness, according to Equations (3.15) or (3.16), based on material and lubricant properties.

### 3.4.4 Spring Model Applications and Limitations

Application of the ultrasound technique for the measurement of lubricant film thickness was first applied in [75] to layers of water formed between two glass plates. A shim of known thickness was used to separate the plates at one end, thus resulting in the formation of a fluid wedge, the thickness of which could be easily determined based on the geometry of the setup. Ultrasound waves were applied to the fluid wedge and thickness determined based the induced ultrasonic reflections and application of the spring model. Figure 3.11 shows the spring-model results as compared with the geometric prediction of the fluid wedge. It should be noted that the spring model was only applicable to films below $25\mu m$ in this case, due to the frequency content of the incident ultrasonic wave. Beyond this a continuum model prediction was utilised, a discussion
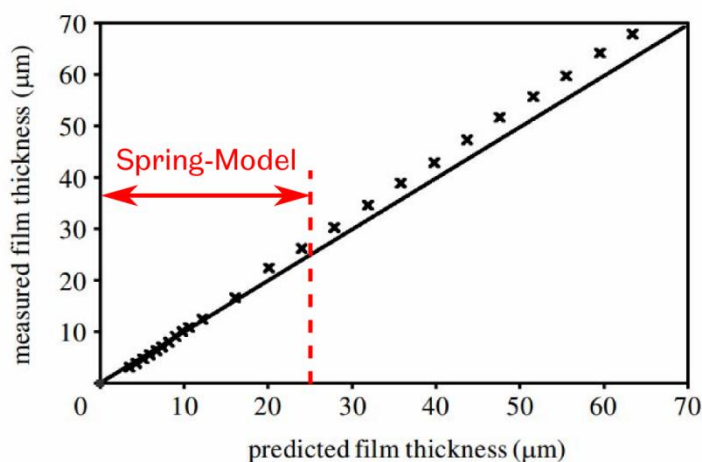


Figure 3.11: Comparison between measured film thickness, as predicted by the spring model, and predicted film thickness, as calculated based on geometry of the fluid wedge (reproduced from [75])

66

of which is beyond the scope of this work. Below $25\mu m$ however, the spring model was found to agree well with the geometrically predicted fluid film thicknesses as shown.

Indeed, the range of film thicknesses that can be feasibly assessed through use of the spring model depends very much on the frequency content of the incident ultrasonic wave. Figure 3.12 shows the anticipated reflection coefficient from oil film layers within the range $0.1 - 30\mu m$ according to Equation (3.11), based on the response of incident ultrasonic waves at frequencies of $1, 5, 10$ and $20MHz$. In this case, reflection coefficients have been calculated according to a steel-oil-steel interface, with oil properties of a typical hydraulic oil (VG32) at room temperature applied. It can be observed that increasing the frequency of the ultrasound wave reduces the magnitude of oil film thicknesses that can be observed as $R$ tends to 1 at much lower film thickness values. Indeed, for a $10MHz$ ultrasound wave, $R = 0.99$ for $h = 6.9\mu m$ , suggesting the limiting measurable film thickness at this particular frequency. Generally, it is advantageous to select an ultrasonic emitter with a frequency range such that significant variation in $|R|$ will be induced across the expected film thickness range, since this will reduce the likelihood of error in spring-model calculations.
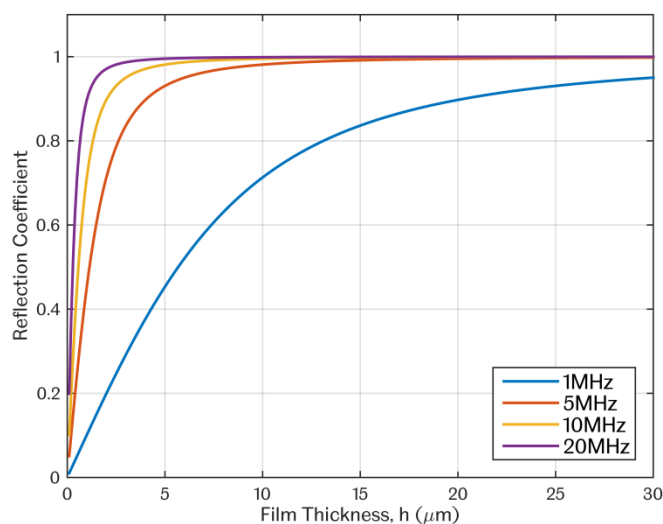


Figure 3.12: Variation in ultrasonic reflection for sound waves of $1, 5, 10$ and $20MHz$ centre frequencies, incident upon a range of oil film layer thicknesses

Following the work in [75], the ultrasonic technique was further applied to EHL films formed in a ball-on-flat contact, as well as in an operational rolling element bearing [81]. Results of the latter were found to agree well with theoretical oil film thickness predictions, albeit with a certain degree

of scatter to the measurement data. The technique has since been further applied to the characterisation of oil film thickness in a range of engineering applications such as journal bearings [82], mechanical seals [83], the piston ring and cylinder interface in reciprocating engines [84], and numerous rolling element bearing applications [69, 80, 85].

In [85], ultrasonic measurements were utilised to monitor the failure of ball bearing lubricating oil through introduction of sand, water and acetone. It was found that the ultrasonic technique was capable of detecting such failures in advance of more classical temperature and vibration monitoring techniques as introduced in the previous chapter.

# 3.5 Conclusions

- Ultrasound is an elastic wave that is transmitted through a structure at a frequency beyond that audible by the human ear ($> 20 kHz$). Propagation occurs via the transferral of particle oscillations via inter-particle "electrostatic" bonds.
- Ultrasound can be generated through use of piezoelectric materials which displace under an applied voltage and can therefore induce the required particle oscillations within a material.
- Piezoelectric elements can be bonded to bearing surfaces to facilitate the generation of ultrasound through a bearing raceway.
- Ultrasound will be reflected from boundaries between media due to a mismatch in acoustic impedance of the two. Thin oil layers separating contacting surfaces (such as those arising in bearing contacts) will also reflect ultrasonic waves, and the proportion of energy reflected is a function of the thickness of the layer, which is characterised by the "Spring Model."
- The ultrasonic technique has been applied to many lubricated contact scenarios and the spring model utilised to determine the thickness of the lubricant layer. Results have shown good agreement with theoretical film thickness calculations.
- The procedure for the measurement of ultrasonic reflection and determination of film thickness has been demonstrated for rolling element bearing applications.

# Chapter 4

# EXPERIMENTAL DESIGN

This chapter presents the test rig used for the experimental work within this study. Test conditions are defined to accelerate bearing life, and a method of fault seeding is proposed that would appropriately simulate the impact events known to occur to planetary WTGB bearings in-service. Finally, the ultrasonic instrumentation of test bearings is described, and an experiment is discussed that was used to map the ultrasonic energy at the rolling surface and subsequently establish the effective measurable area of the installed sensor.

## 4.1  Test Rig

A bespoke test rig was designed and commissioned by Ricardo UK Ltd. to provide validation of the MultiLife bearing concept. Throughout the design phase the author was responsible for the integration of previously mentioned AE and Ultrasonic sensors within the rig so that test bearings could be continually monitored. Upon commissioning, the test rig was received by the University and housed in the Mechanical Engineering Lea Laboratory for the duration of the test programme.

The rig was designed to replicate the operational mode of planetary stage WTGB bearings through application of uni-directional and purely radial load upon a stationary internal bearing shaft. Bearings were bore-mounted and rotated at constant speed via their outer raceways, thus replicating the planetary bearing operating mode accordingly. The rig was designed to accommodate test bearings of the *NU2244* cylindrical roller (CRB) type; selected to be representative of that utilised within a typical

$2.5 - 3MW$ WTGB with respect to both type and size; thus all testing was performed at full scale. Key specifications of the test bearings are outlined in Table 4.1.

| | |
|---|---|
| Bore diameter, $d$ | $220mm$ |
| Outer diameter, $D$ | $400mm$ |
| Width, $b$ | $108mm$ |
| Inner raceway thickness, $t$ | $19.5mm$ |
| Dynamic load rating, $C$ | $1600kN$ |
| Fatigue load limit, $C_u$ | $250kN$ |
| Number of rolling elements, $Z$ | 15 |
| Rolling element diameter, $d_e$ | $54mm$ |
| Rolling element length, $L$ | $82mm$ |
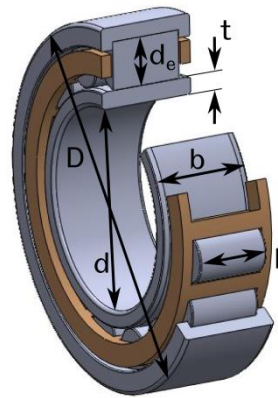
Table 4.1: Specification and illustration of NU2244 test bearings

The objectives for testing were threefold; as outlined below.

1. Confirm that representative failures could be simulated in test bearings through fault seeding and subsequent steady state operation at a predefined constant load and speed (selected to accelerate failure rates based on bearing life rating models).
2. Confirm that inner raceway indexation, following fault seeding, would act to extend bearing life (as per the MultiLife Thesis).
3. Obtain Ultrasound and AE data throughout the testing to identify any signal changes pertaining to bearing condition. Assess the applicability of such data as a means to provide a condition monitoring tool for wind turbine gearbox bearings.
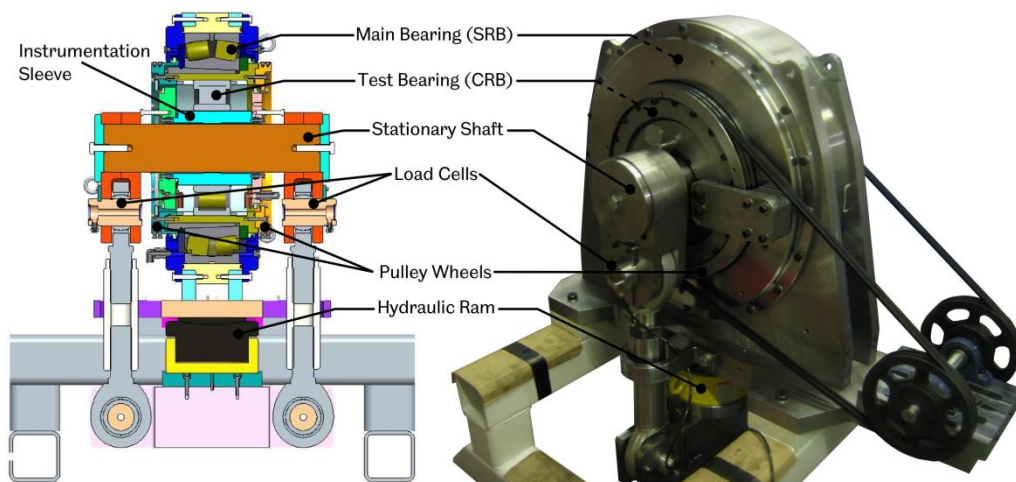
Figure 4.1: The MultiLife test rig

The test rig is shown in Figure 4.1 alongside a cross-sectional CAD drawing. Bearings were loaded via a hydraulic cylinder, which induced tension in two load arms connected to either side of the stationary bearing shaft. Two shear mode load pins coupling the load arms to the shaft provided indication of tensile force, thus permitting closed-loop feedback control of applied load via adjustment of hydraulic ram pressure. Hydraulic ram capacity of $410bar$ permitted maximum applicable radial loads of around $1500kN$.

Test bearing outer raceways were driven by a $7.5kW$, $1450rpm$ electric motor via a $4.5\colon 1$ gearbox reduction and $2.9\colon 1$ pulley stage. A variable frequency inverter (VFD) enabled a range of bearing speeds to be achieved up to a maximum of approximately $100rpm$. Test bearings were mounted within a larger double row spherical roller bearing (SRB) to permit the desired test bearing outer race rotation.

## 4.1.1  Establishing Robustness in the Test Setup

Given the time frames involved in observing full bearing life cycles, continuous 24 hour operation was desirable and it was therefore assumed that tests would often be carried out under minimal supervision. Consequently, robust control mechanisms were required to ensure the test conditions, as defined in §4.2, were met and maintained appropriately, as well as ensuring a high level of safety was upheld at all times. As a key part of the rig commissioning phase, the author was responsible for the development and implementation of all the measurement and control systems outlined in the following sections.

### 4.1.1.1  Sensors for Rig Control and Monitoring

To facilitate automated control of operating conditions, the test rig was run via a desktop PC and LabVIEW software. This also allowed for a range of test parameters to be measured and recorded, and safe shutdown procedures to be implemented when required. A schematic block diagram illustrating all test inputs and outputs is shown in Figure 4.2.

Bearing speed and load were both closed-loop feedback controlled to ensure that set points were maintained throughout each test. A standalone electric pump supplied pressure to the lock nut style hydraulic cylinder used to apply load to the test bearings. The pump was modified to receive either "Pump" or "Relieve" commands via $5V$ digital signals, thus allowing control of applied load based on feedback from the load cells.
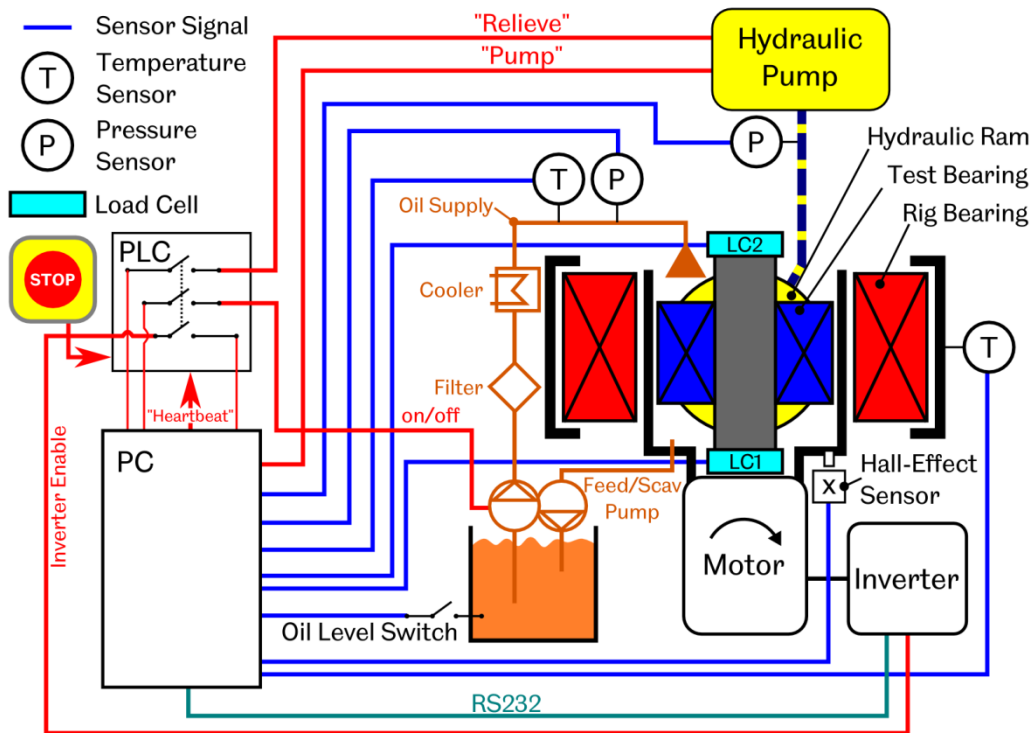
Figure 4.2: Schematic diagram of all test rig input and output signals for control of test parameters

A Hall-Effect sensor was mounted to the outer rig casing to sense the passage of a steel stud mounted to the rotating outer bearing sleeve. A once per revolution pulse would thus be observed, providing a count of bearing revolutions, as well as an indication of rotational speed. Motor speed was then controlled electronically via a serial communication (RS232) between the control PC and the VFD, therefore allowing bearing speed to be appropriately matched to that of the demand set point.

Pressure sensors were installed at the test bearing lubricant oil inlet point on the rig casing, and the input line to the hydraulic ram. K-type thermocouples were installed in three locations to measure temperature at the oil inlet to the test bearing, the rig casing, and temperature of the ambient air.

Lubricant was supplied to the test bearing continuously throughout the duration of a test. The hydraulic circuit consisted of a $0.37kW$ single-phase electric motor simultaneously driving in-line feed and scavenge gear pumps. Oil was pumped via a 25 micron filter into a single inlet port on one side of the test rig at 60° from bearing bottom-dead-centre (BDC) and at a radius of $140mm$ from the bearing centre axis so as to feed directly between the inner and outer raceways. A cooling radiator was installed ahead of the oil inlet port to enable passive cooling of the supply

via ambient air. A fan was also installed to provide enhanced airflow across the radiator so that more aggressive "active" cooling could be implemented if necessary. The scavenge pump drew oil back to tank through outlet ports situated either side of the rig at bearing BDC. Supply power to the motor was fed via a solid state relay, allowing on/off control of the hydraulic system via a $5V$ input signal. A level switch in the oil reservoir was installed to output a $5V$ digital signal to the control PC in the event of oil levels falling too low.

Control software was configured to store all input data in a log file at a rate of $1Hz$. Data was stored in comma separated value (csv) format in a single data file per test run. This was setup to append to the log file at 10 second intervals, so as to ensure that only a maximum of 10 seconds worth of data would ever be lost in the event of a power failure or computer crash.

### 4.1.1.2 Safe Shutdown Procedures

Test control software was configured to automate a safe shutdown in the event of predefined thresholds on various test parameters being exceeded. Table 4.2 details all shutdown criteria assigned to test variables and Figure 4.3 illustrates the shutdown procedure implemented once any of the defined thresholds were crossed.

| Test parameter | Shutdown Criterion | Description |
|---|---|---|
| Test bearing oil inlet temperature | > 90°C | Abnormal frictional heating; damage to pump? |
| Test bearing oil inlet pressure | < 0.5 bar | Potential leakage/loss of oil; insufficient supply to bearing |
| Hydraulic ram pressure | > 400 bar | Restriction in oil feed line, excessive pressure to ram |
| Rig casing temperature | > 80°C | Abnormal rig heating; potentially increased friction/resistance to motion due to damage on rolling surfaces |
| Test bearing speed | < 80% of demand | Belt slip due to increased resistance to motion, risk of belt damage |
| Oil level | < 1 litre | Potential oil leak; insufficient supply to bearing |

Table 4.2: Rig shutdown criteria

Figure 4.3: Rig shutdown sequence

A programmable logic controller (PLC) was also implemented (as illustrated in Figure 4.2) to provide overriding control in the event of a computer crash or power outage. In the absence of a PC "heartbeat" signal, or mains power, PLC relays would close to automatically cease bearing rotation, disable lubricant supply and release any applied bearing load.

## 4.2 Definition of Test Conditions for Accelerated Bearing Life

As discussed earlier, testing was performed at a scale representative of a $3MW$ wind turbine; that being one of the largest design variants currently utilised commercially.

As discussed in §1.2, failures of WTGB bearings can be observed up to two to three times within the 20 year design life of a given turbine. In-service this is clearly an unacceptable rate of failure, in the context of this bearing life testing however, this clearly represents a significant test time and it would therefore be necessary to induce failures within much shorter and repeatable time frames. As such, a key requirement prior to testing was to establish operating conditions that would permit observation of full bearing life cycles in reasonable time frames. A target range of 1-2 weeks was defined, considered to be suitably short to allow a sufficient data set to be obtained whilst long enough to ensure operating conditions were not extreme, and failure mechanisms were still representative of those observed in-service.

The life rating procedures outlined in §2.2.6 were therefore utilised and the following test variables were determined, found to give an $L_{10(h)}$ life rating of 69.4 hours for the test bearing data set.

- VG32 oil
- $100rpm$ bearing speed
- Oil temperature $\approx 60°C$

- Applied radial load $\approx 1140 kN$

Assessment of higher failure rates was also performed based on the Weibull cumulative distribution curve for the life modification factor $a_1$, as presented in Figure 2.6. It was established from this that by 336 hours (2 weeks), failure rates of 97.9% were expected, thus providing reasonable certainty that bearing failures would be induced in most of the bearing set within this timeframe.

The above analysis therefore satisfied the upper limit of the bearing life target range of 1-2 weeks with a sufficient degree of certainty. It is important to note however, that a purely fatigue based failure was assumed. As discussed in §1.2.2, while surface failure of planetary support inner raceways has been found to be driven by a cyclic fatigue mechanism, it is understood that initiation is typically a consequence of an impact event. The above assessment was therefore based on a "worst case" scenario for a standard bearing under stable operating conditions. In fact, most tests were expected to reach a conclusion much sooner due to the presence of seeded stress concentrations that would be introduced into the bearing inner raceways, thus applied operating conditions would merely act to enhance propagation of the desired failures.

## 4.2.1 Simulation of Impact Events through Geometric Fault Seeding

As discussed in Chapter 1, failure-inducing impact events occurring in-service may well take place only a handful of times, and therefore must be significant in magnitude to generate the observed failures. This is particularly the case for the larger bearings utilised in the megawatt WTGB variants such as the NU2244 type used within this study, where resistance to fatigue failure will be upheld at extremely high load (apparent from the dynamic load rating of such bearing as defined in Table 4.1). The actual magnitude of impact energy required to seed sufficient sub-surface faults within the bearing raceways is difficult to define, and would typically require sectioning of the material post-impact to assess material status and quantify the degree of induced damage.

A series of drop tests were thus performed, where a test bearing rolling element and inner raceway were struck together under incrementally increasing levels of instantaneous impact energy. This, it was hoped, would lead to identification of an energy level at which plastic deformation within the inner raceway would extend to the rolling surface. With this

knowledge, the impact energy level prior to this occurrence would then be used as the baseline for sub-surface fault seeding of bearings that would subsequently be run on the test rig. However, no such observations could be made up to the $500J$ maximum capacity of the impact rig.

To expedite testing and permit seeding of faults that were known, easily measurable and repeatable, initiation points were subsequently induced into bearing raceways through geometric changes to the rolling surface. Whilst not being entirely representative of the true mode of initiation, simulation of the ultimate failure mode was considered a higher priority than simulation of the impact event itself, and it was supposed that – by such methods – representative localised spalling failures could still be induced, but via a surface – rather than sub-surface – RCF initiation mechanism. As such, it was proposed that artificial defects would be introduced to bearing rolling surfaces via Wire-Cut Electrical Discharge Machining (EDM), thus only subtle changes to the rolling surface could be achieved. Indeed, based on the typical width of a wire electrode, a defect with a width of as little as $200\mu m$ could be induced across the entire width of the raceway (normal to the rolling direction) where depth could be controlled through the degree by which the tool electrode was incurred upon the stationary race. A brief study identified that depths of as little as $0.1mm$ were achievable, and this was used as the standard geometric fault for the testing (Figure 4.4).

It has since been identified that surface yield can be induced through compressive loading of a roller against the raceway under a $1250kN$
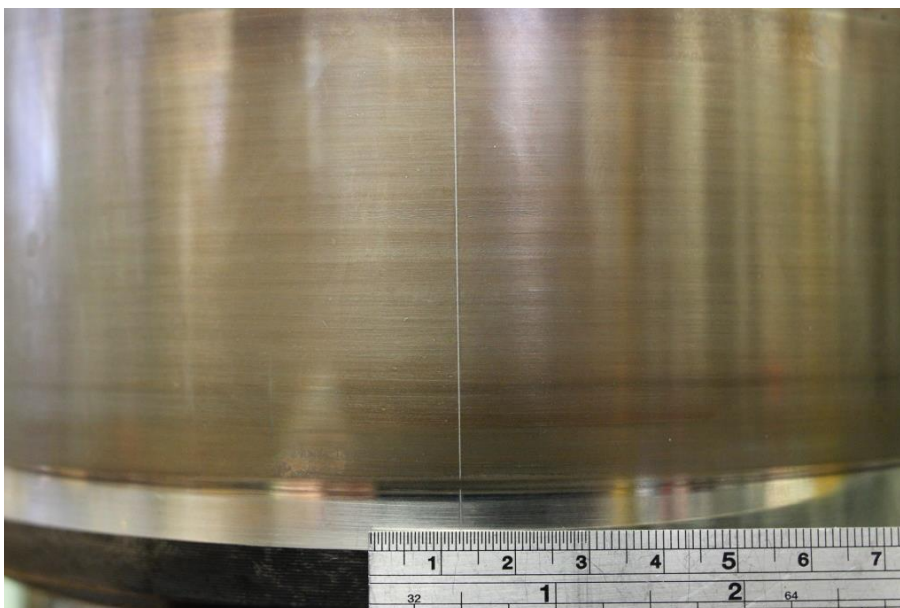


Figure 4.4: A test bearing inner raceway seeded with a geometric surface fault induced via wire-cut EDM

compressive load. Sub-surface crack initiation appears to have been identified during this through measurement of acoustic emission bursts which were found to occur as early as $800kN$ into the compression ramp. Bearings with such seeded faults have yet to be tested on the rig and these investigations are beyond the scope of this current work, but are the focus of additional studies [86].

## 4.3  Ultrasound Implementation

Instrumentation of test bearings was a key consideration from early in the rig design phase. It was important that both ultrasonic and acoustic emissions sensors were mounted to test bearing inner raceways, this being the critical component in terms of failures. A sleeve arrangement was therefore devised within which the sensors systems could be housed. As shown in Figure 4.1, this provided an intermediary between the stationary inner shaft and the test bearing inner raceways. Figure 4.5 shows a detailed view of the central sleeve component and illustrates the location of both AE and ultrasound sensors within it.

The application of AE sensors and associated signal analysis is not discussed in detail within this work; this being the focus of a parallel project [20]. It is of interest to note at this point however that three AE
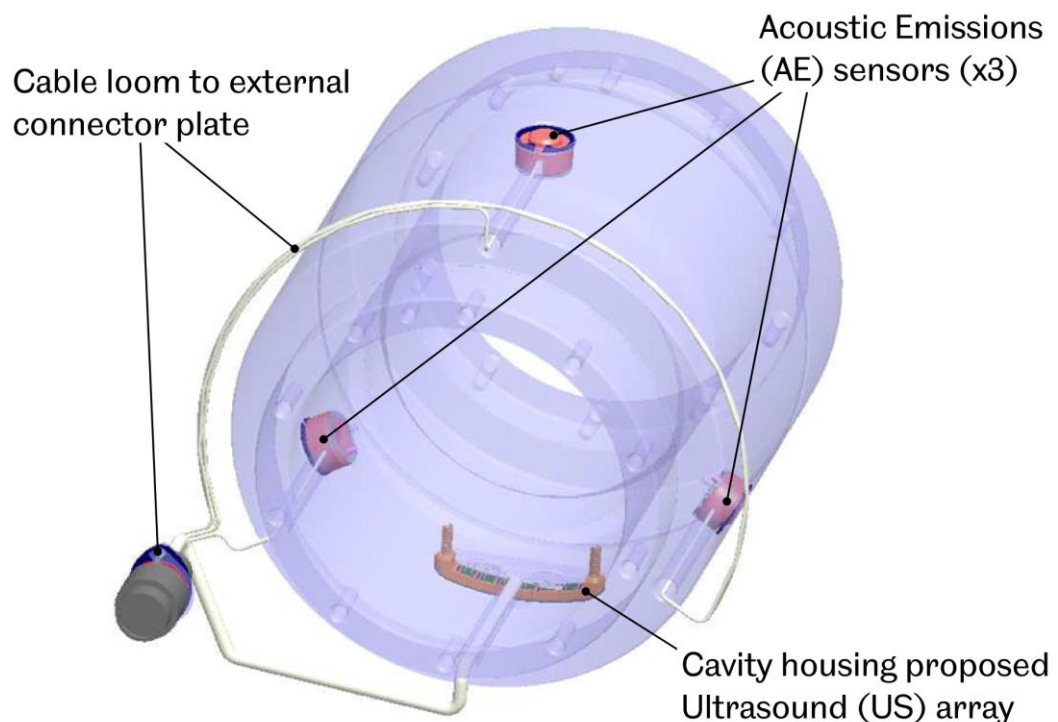


Figure 4.5: Test bearing internal sleeve to house condition monitoring instrumentation

sensors were utilised which were bonded to individual steel carriers that were held within the central sleeve and spring loaded against the inner raceway bore. Sensors were equally spaced at 120° intervals around the raceway circumference, which was sufficient to permit measurement of acoustic signals generated across the entire raceway, and would also allow defect locations to be determined through triangulation of acquired signals (as discussed in §2.3.4).

With regards to this study however, accommodation was made for an array of ultrasound sensors that would permit visualisation of the inner raceway-rolling element interface at the point of maximum bearing load, as well an adjacent 15° arc either side of bearing BDC. It was essential that ultrasound instrumentation was capable of monitoring any changes in the roller-inner raceway contact across the maximum loaded zone; that being the area within which failures were expected to initiate.

The application posed a significant technical challenge regarding ultrasound sensor installation. Where typically, as discussed in §3.4.1, piezoelectric elements would be permanently bonded to a rolling element bearing on the outer surface of the raceway of interest, it was desirable in this case to seek a mechanism for non-permanent coupling. Clearly, this represented the most practical solution for a MultiLife bearing to ensure that – following inner raceway indexation – the ultrasonic sensor would remain stationary in order to maintain visualisation of the critical, high load region of the bearing. As such, various sensor carrier systems were initially investigated whereby ultrasonic sensors would be mounted to, or within, an intermediate body, that could be positioned, held within the central sleeve and spring loaded against the inner raceway to facilitate measurement of the desired interface.

### 4.3.1 Maintaining Ultrasonic Coupling in a Non-Bonded Setup

The key challenge with such a system was the requirement for the sensor carrier-bearing interface to be adequately coupled in order to eliminate air gaps that may exist between the two that would consequently result in complete reflection of incident ultrasonic waves. In a bonded installation, the bond layer itself serves as the coupling medium, and therefore, for a non-bonded application, a means to ensure adequate coupling of the interface was required. An ultrasonic coupling rubber was thus identified, whose properties were unique in that high frequency sound waves would be transmitted with low attenuation if sufficient pressure was applied to

the layer. It was shown in [87] that a pressure of $0.3MPa$ was sufficient to achieve such transmission.

## 4.3.2 Sensor Carrier Design Investigation

Figure 4.6(c) provides further detail of the space allocated within the central sleeve component of Figure 4.5 to house ultrasound instrumentation. This was primarily designed to accommodate an array, or "carrier," of the general geometry illustrated in Figure 4.6(b). This design was adopted such that up to seven ultrasound sensors could be utilised, there was sufficient space for cable routing within the cavity, and the array could be coupled to the inner raceway via springs applied to either side, whose stiffness's were sufficient to generate the $0.3MPa$ interface pressure required by the coupling rubber.
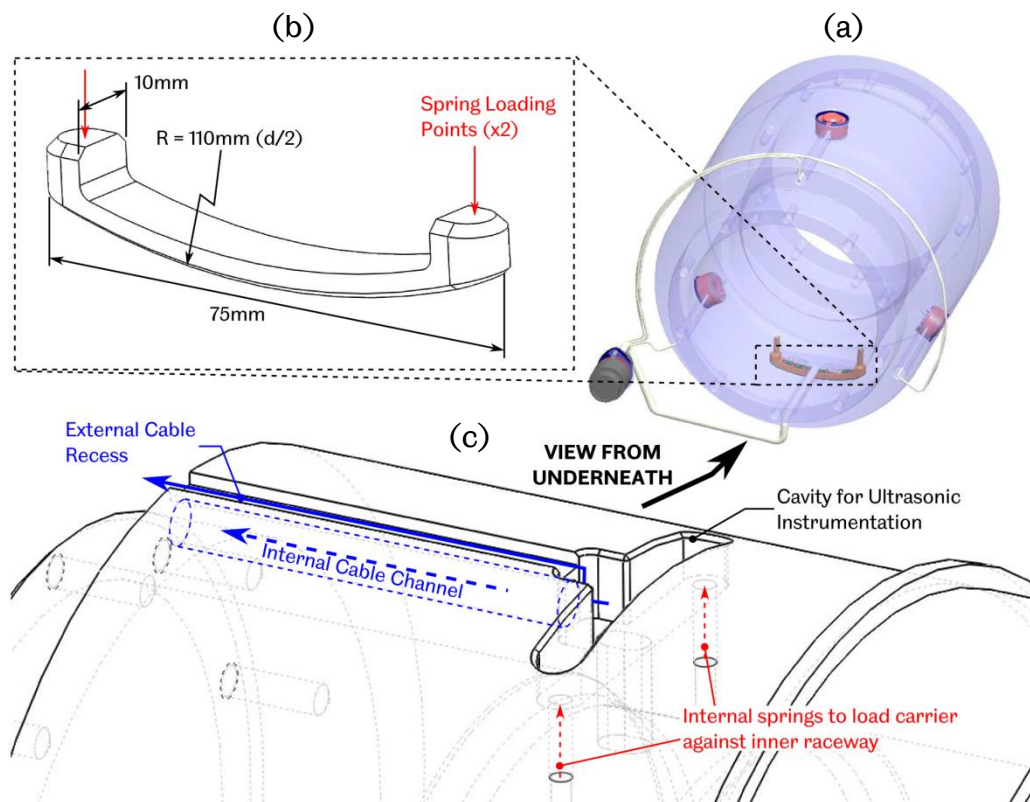


Figure 4.6: (a) Test bearing instrumentation sleeve as shown in Figure 4.5, (b) detail of sensor carrier design, (c) machined features in the bearing sleeve component for instrumentation

### 4.3.2.1 Design Iteration 1: Through Transmission Array



Figure 4.7: 1st iteration "Through-Transmission" ultrasound array

The initial design concept is shown in Figure 4.7, where piezoelectric elements were bonded to a Perspex carrier. This setup required that ultrasonic energy be transmitted through three discrete interfaces, and as such, consideration of acoustic mismatch and the associated ultrasonic reflection at media boundaries (as per Equation (3.4)) was essential to ensure that sufficient energy could be transmitted to the measurement interface. Given acoustic properties of the rubber layer as defined in [88] ($c = 957m/s$ and $\rho = 1420kg/m^3$); acoustic impedance was established, according to Equation (3.3), as $z_r \approx 1.4MNs/m^3$.

| Material | Acoustic Impedance, $z\ (MNs/m^3)$ | Ultrasonic reflection at media-rubber interface* |
|---|---|---|
| Steel | 46.0 | 0.94 |
| Aluminium | 17.4 | 0.85 |
| Perspex | 3.3 | 0.40 |
| PEEK | 3.4 | 0.42 |
| Polycarbonate | 2.8 | 0.33 |
| GFRP | 6.2 | 0.63 |

Table 4.3: Acoustic impedance comparison of potential carrier materials and associated reflection of ultrasonic energy (*assuming a "perfect" interface)

As shown in Table 4.3, too great an acoustic mismatch was found to exist between typical engineering metals and the rubber layer. Thermoplastics were found to provide a more satisfactory match, yet were typically highly attenuating. PEEK and Perspex were both found to have similar acoustic properties with good acoustic matching to the rubber layer, and both were subsequently explored further. Ultrasonic transmission tests of the two however found PEEK to again be too highly attenuating, leading to the selection of Perspex as the most appropriate carrier material.
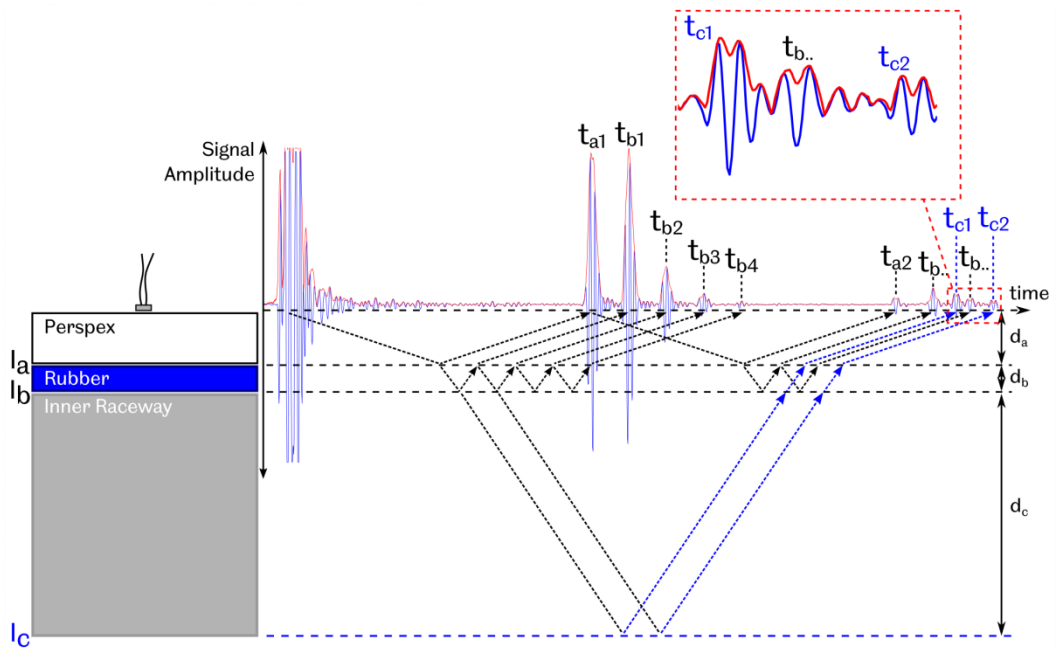
Figure 4.8: Ultrasonic reflections from a setup representative of the through transmission sensor array

Figure 4.8 shows the ultrasonic reflections that would be generated from a setup representative of the array shown in Figure 4.7, where an ultrasonic signal is transmitted through Perspex, rubber and steel layers of similar relative thicknesses. As illustrated, despite optimisation via acoustic matching; the overall thickness and multi-layer nature of the transmission path consequently resulted in relatively low amplitude signals received back from the measurement interface. It was also necessary to ensure thickness of the Perspex layer was accurately defined so as to ensure that pulses received from the measurement interface ($I_c$) were out of phase with adjacent rubber-steel ($I_b$) reverberations. However, despite signal amplitudes being sufficient to provide adequate measurements at room temperature, attenuation was found to increase significantly at test temperatures, owing to the low glass transition temperature of the Perspex (around $100°C$). As a result, signal-to-noise ratios were often found to reach unity at temperatures of around $70°C$.

As such, the overall concept was deemed inappropriate and alternative design concepts were sought.

### 4.3.2.2   Design Iteration 2: Potted Array

In order to address the restrictions imposed by the original sensor carrier design in terms of material requirements, and to minimise media

boundaries within the ultrasound transmission path, an alternative design was developed as shown in Figure 4.9. Here, the same carrier geometry was utilised, but piezo-electric elements were potted within individual holes drilled through a steel carrier at the desired sensor locations. By this method, piezo-electric elements could be coupled directly to the rubber layer; thus eliminating the need for a highly attenuative plastic layer entirely. Since ultrasound was no longer transmitted through the carrier body, acoustic matching was not necessary, and a more substantial engineering material could be employed.



Figure 4.9: 2$^{nd}$ iteration potted ultrasound array

Figure 4.10 shows the ultrasonic reflections generated by a single sensor within the array when coupled to a test bearing inner raceway. Clearly, time-of-flight to the measurement interface was considerably reduced with this setup, thus resulting in higher amplitude pulses received from the measurement interface, as well as higher signal-to-noise ratios for a given voltage applied to the piezo-electric element (in comparison to the through transmission array signals). However, without the Perspex layer acting as a delay line for the ultrasonic signal, reverberations through the rubber layer were found to dominate. As such, reflected pulses from the measurement interface ($I_b$) were always superimposed by reverberations from the rubber-raceway interface ($I_a$). It was initially assumed that – given no change was expected at this point – the $I_a$ reverberations could simply be subtracted from the full ultrasonic waveform to reveal the measurement signal. However, once commissioned, a wide degree of variability within the rubber-raceway interface itself was observed, thought to be due to ingress of oil between the inner raceway and central sleeve which could consequentially flow across the rubber layer. As such, significant noise was induced upon the ultrasonic $I_b$ interface pulses,

leading to difficulty in decoupling variations due to roller-raceway effects from those induced at the rubber-raceway interface.
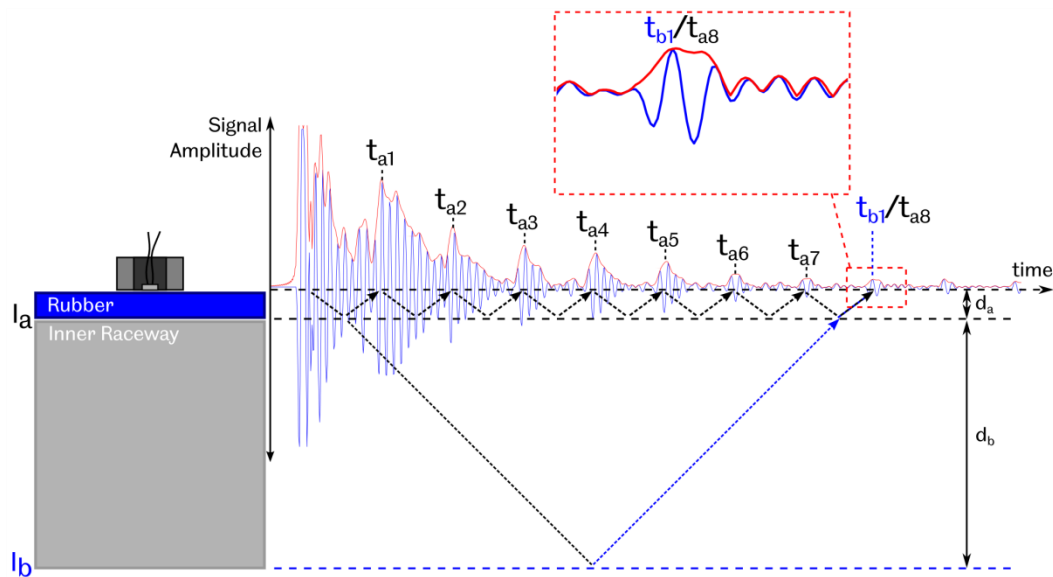


Figure 4.10: Ultrasonic reflections from a single sensor within the 2nd iteration sensor array

### 4.3.2.3 Summary

It was identified through the above investigations that a significant degree of complexity would be introduced into the implementation of ultrasound on test bearings when a non-permanent coupling mechanism is required. Clearly, the key priority within this study was that of data acquisition and assessing the applicability of ultrasound as a condition monitoring tool. This therefore took precedence over the determination of commercial applicability of such measurement techniques and therefore the investigation of non-permanent coupling mechanisms was not explored further.

Generally, the implementation of the non-permanent coupling methods investigated exposed a number of challenges, such as those discussed within the previous two sections, but also some additional factors as detailed below. Potential solutions are thus discussed as well as some points of further work that would be necessary in order to satisfy commercial viability.

- ▪ Ultimately, degradation of the coupling rubber layer was observed at high temperature ($\sim 80°C$) over extended durations. Silicone

variants are known to exist and would provide greater suitability over an extended temperature range.

- Variations observed in measurement pulses as a result of changes at the rubber-raceway interface may be decoupled through simultaneously monitoring the last pulse received from the rubber-raceway interface before the measurement pulse is received (pulse $t_{a7}$, Figure 4.10). Since this represents variation at the rubber-raceway interface only, the percentage change in signal amplitude could be assumed to be the same for the next reverberation (variations are constant within a $1.5\mu s$ window), and thus could be compensated for within the measurement pulse in real-time.

- In order to eliminate the need for a rubber layer entirely, coupling could also be achieved via a standard commercial liquid couplant (of which high temperature variants are also readily available [89]). This would require appropriate sealing of the interface to ensure retention and may also require means of injection should leakage occur, to ensure that the liquid chamber could always remain flooded.

### 4.3.3 Bonded Installation

Given the issues outlined in the previous section, a bonded installation was deemed the best practical solution. It was accepted however, given that bearings would need to be re-instrumented following each raceway indexation, that the utilisation of an array of sensors was unfeasible. Further to this, modifications would need to be made to the central sleeve component to allow inner raceways to be installed with sensors pre-attached; therefore it was deemed sensible that modifications to this component be kept to an absolute minimum. As such, installation of a single sensor at bearing BDC was considered the best approach.

Figure 4.11 shows a test bearing inner raceway instrumented with both a single piezoelectric element and k-type thermocouple. A $9 \times 5mm$ groove was machined into the central sleeve component (as shown in Figure 4.6(c)), such that instrumented raceways could be installed onto the sleeve, and cables could be routed to the pre-existing connector plate on the rig casing. The $7.1mm$ diameter piezoelectric discs were cut to strips of $1.5mm$ width to roughly match the expected contact width at the roller-inner raceway interface at BDC.
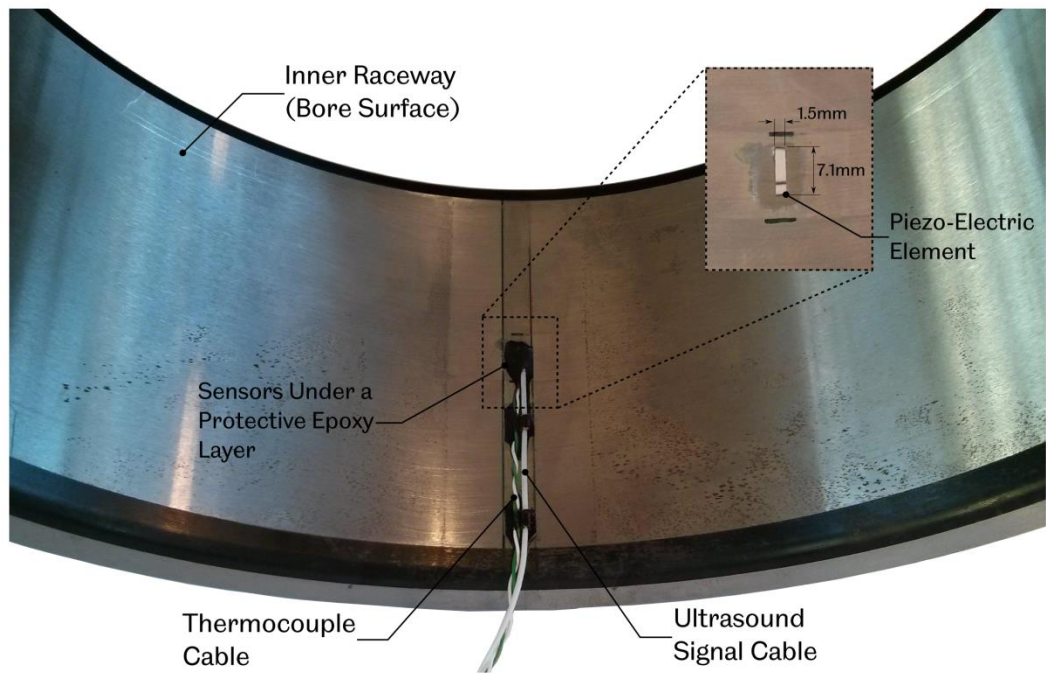
Figure 4.11: Test bearing inner raceway instrumented with an ultrasonic element and thermocouple

Indeed, dimensions of piezoelectric elements are typically defined such that the area of excitation is smaller than the expected area of contact at the measurement interface. However, for applications concerning rolling element bearings such as this; where lubrication regimes are elastohydrodynamic and contact areas are generally extremely small, this is often impractical. It was considered, however, that the real area of observation of a given piezoelectric element may in fact be much smaller than its excitation area due to divergence of ultrasonic energy through the transmission medium, as illustrated in Figure 4.12. Here, due to beam spread through the bearing raceway, a large proportion of reflected energy is diverted away from the ultrasound transducer and, consequently, only energy reflected from a very small area on the rolling



Figure 4.12: Effects of ultrasonic beam divergence on real measurement area

surface would be received across the width of the transducer.

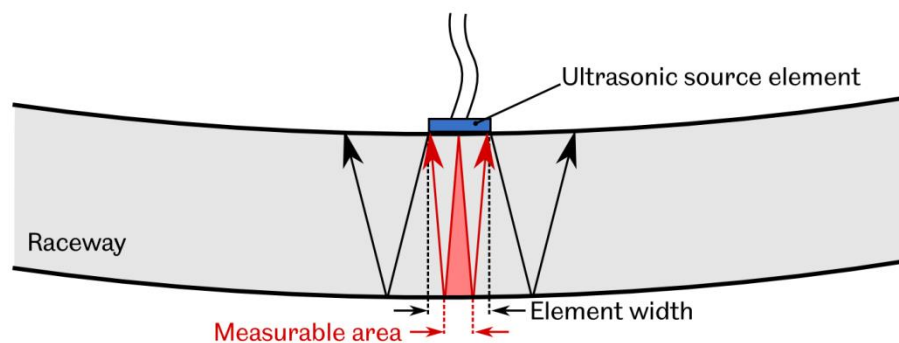Clearly, knowledge of the effective area of observation across the bearing rolling surface was of great interest, but required knowledge of the divergence angle of the ultrasonic beam. While standard calculations for beam divergence exist, these are generally empirical, and consider frequency and radius of typical commercial ultrasonic transducers, as well as associated sound velocity through the transmission medium [90]. Divergence of ultrasonic energy emitted from bare piezoelectric elements will vary by application, and will be influenced by numerous factors, the most significant of which being the bond layer between the element and the parent material, as well as natural variations that will evidently ensue from slicing discs into strips of a desired width.

### 4.3.3.1  Mapping of the Ultrasonic Wave-Front

An investigation was therefore carried out to assess the divergence of ultrasonic energy through the bearing raceway shown in Figure 4.11. This was achieved via the setup shown in Figure 4.13, where a fine, ball-tipped probe with a piezo-electric element mounted to the back-face was swept across a $20 \times 20mm$ area of the rolling surface, roughly central to the piezo-electric element mounted on the inner raceway bore. The probe was precision actuated via stepper motor and ball screw arrangements in two axes to permit measurement across a two-dimensional area. An FMS (as described in §3.4.2) was configured in "pitch-catch" mode to pulse the ultrasonic element installed on the bearing, and receive corresponding signals transmitted through to the rolling surface via the ultrasonic probe. The geometry of the probe tip was such that the area of contact formed between the two components was extremely small, thus the ultrasonic wave-front could be mapped with high spacial resolution. Stepper motor control and ultrasonic waveform acquisition were performed simultaneously via LabVIEW control software on the FMS, which was configured to acquire ultrasonic signals at $0.02mm$ steps in both the $x$ and $y$ axis across the entire measurement area.
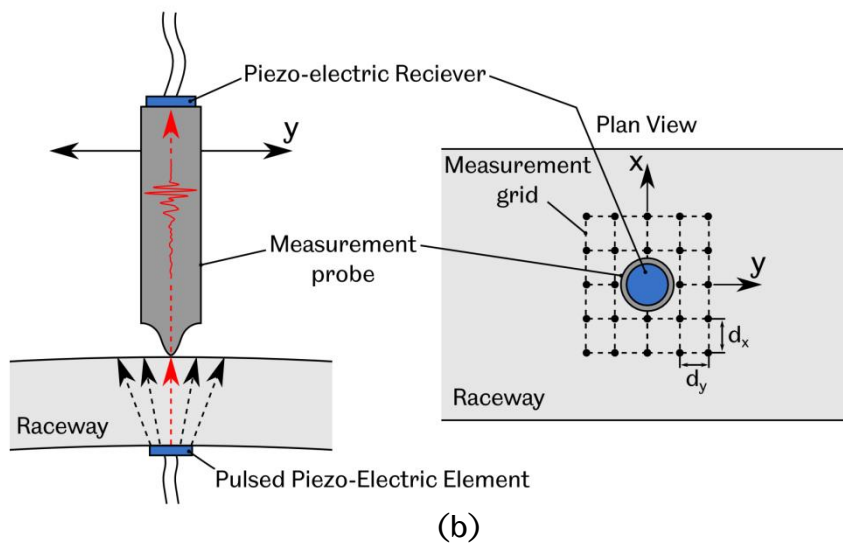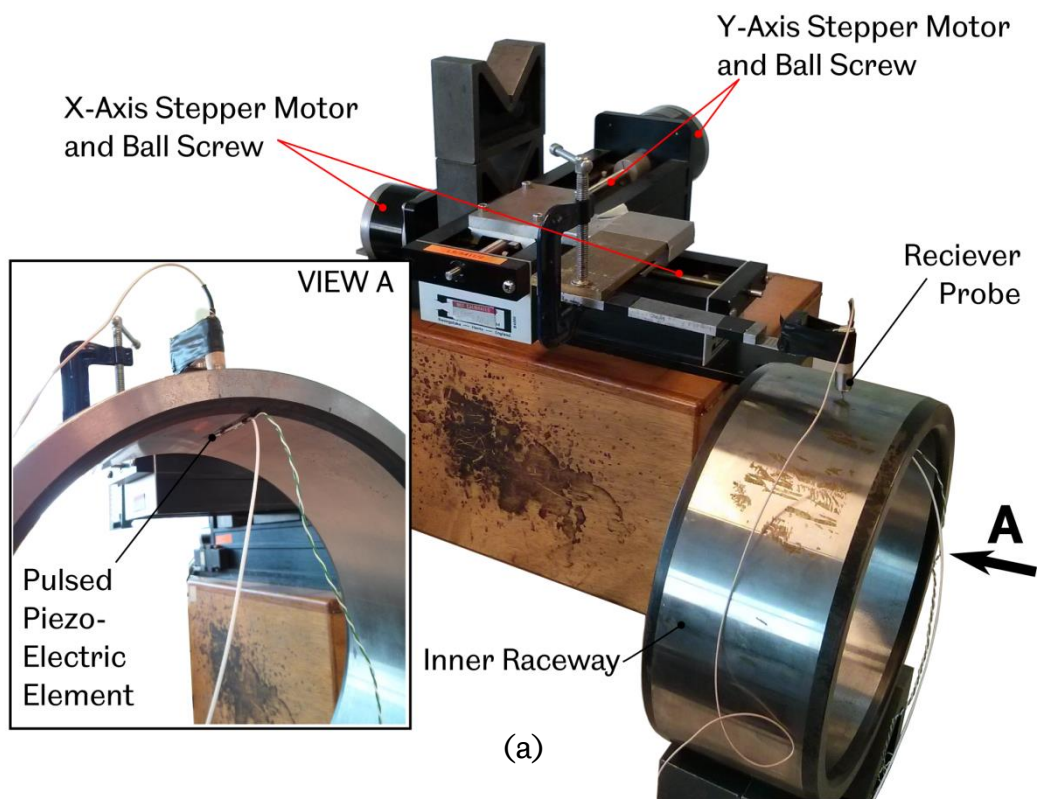
Figure 4.13: (a) Experimental setup for mapping of the ultrasonic wave-front through a test bearing inner raceway, (b) Schematic of the measurement technique

Ultrasonic waveforms obtained across the entire scan area were assessed to find the single waveform with maximum peak-to-peak amplitude. This is shown in Figure 4.14, and corresponded to the energy present within the centre axis of the emitted ultrasonic beam. Peak reference amplitude, $A_0$, was defined from this as 1296, occurring at time $t_0 = 0.53\mu s$ from the start of the data capture window (the same for all captured waveforms).
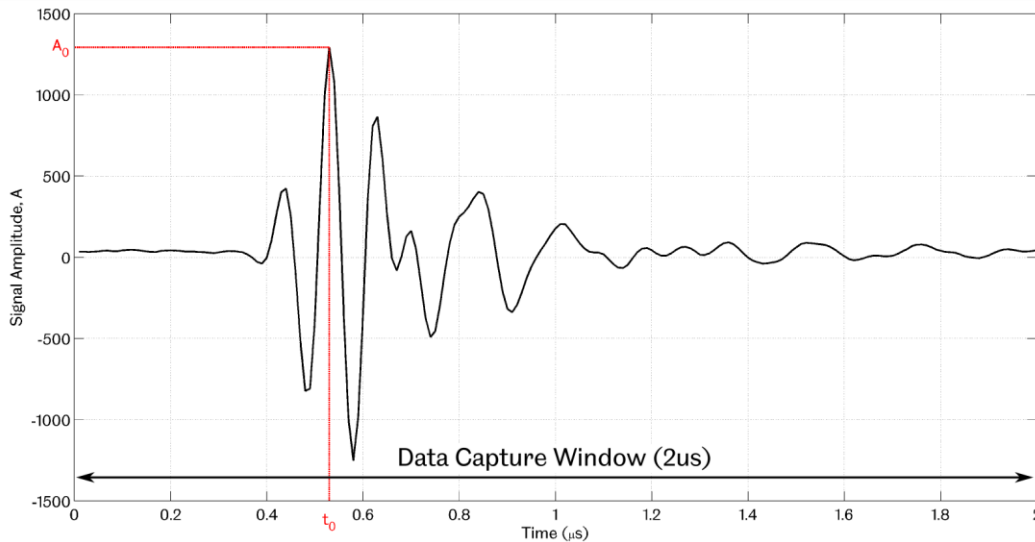


Figure 4.14: Ultrasonic reference waveform obtained from the scan of the rolling surface representing the energy within the centre axis of the ultrasonic beam

Figure 4.15 compares the signal amplitude of all ultrasonic waveforms obtained across the scan area at specific instances in time relative to $t_0$. Here, data maps have been centralised to the ultrasound beam centre axis and clipped to an area of $10 \times 10mm$. Outer dimensions of the pulsed ultrasonic source element are represented by a black bounding box overlaid on each scan. Clearly, beam spread was significant, and the amount of energy present at the rolling surface at any given time was variable.

In order to define beam width appropriately, the case of peak ultrasonic energy at the rolling surface ($t_0$, Figure 4.15(c)) was analysed further. Figure 4.16 illustrates the relative signal amplitude for $x = 0$ (i.e., moving circumferentially across the rolling surface in a plane through the beam centre axis). Typically, beam spread is defined as the point at which sound pressure has reduced by half that of the peak pressure at the centre axis [90], corresponding to a $-6dB$ amplitude reduction. This threshold is indicated on the figure, and leads to the definition of $3.8mm$ as the outer boundary of the ultrasonic beam at the rolling surface, circumferential to the inner raceway. Geometry of both the roller and raceway dictate that a

line contact will form between the two, and thus contact length will be much larger than width, (the contact patch will span almost the entire length of the roller). As such, beam divergence was not considered in the axial plane, as the beam would be significantly smaller than the contact length for all load cases.
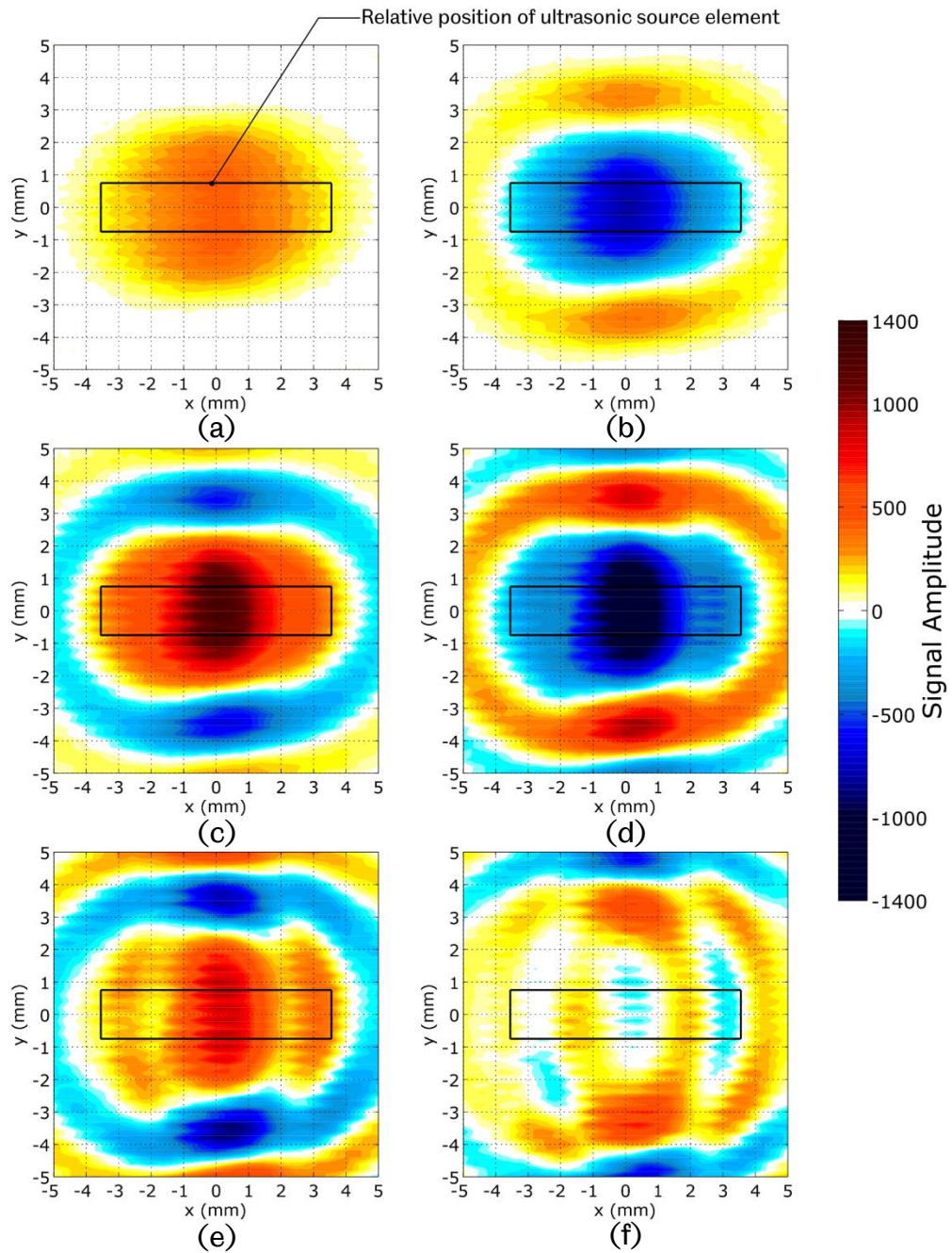


Figure 4.15: Maps of ultrasonic energy present at the rolling surface at (a) $-0.09\mu s$, (b) $-0.05\mu s$, (c) $0\mu s$, (d) $0.05\mu s$, (e) $0.10\mu s$ and (f) $0.14\mu s$ from the instance of peak ultrasonic energy in the centre axis of the beam, $t_0$
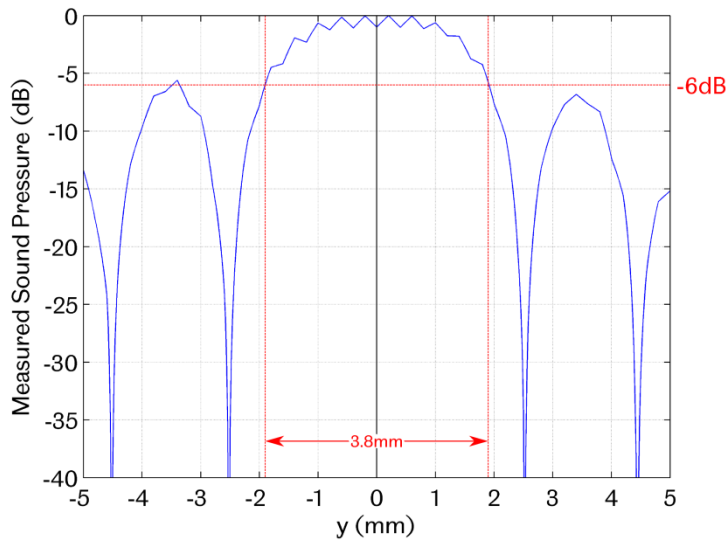
Figure 4.16: Relative sound pressure measured at the rolling surface at $t_0$ in the rolling direction ($x = 0$)

Figure 4.17 illustrates the assumed beam path through the bearing inner raceway based upon its measured outer width of $3.8mm$ at the rolling surface. Considering the geometry of the raceway and the width of the ultrasonic source element, divergence angle was consequently calculated as $6.93°$. It was then assumed that the angle of trajectory of ultrasonic energy emitted from a discrete point across the width of the source element would decrease linearly from maximum divergence at the outer edge of the piezo-element ($y = \pm0.75mm$), to 0° at the beam centre axis ($y = 0$). The angle at which the return width of the beam was equal to that of the ultrasonic source element was then established, and the
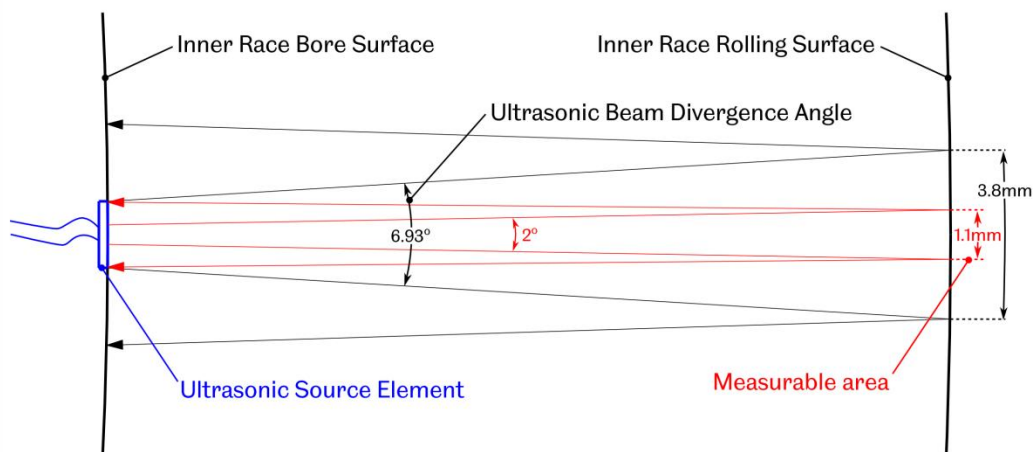


Figure 4.17: Illustration of inferred ultrasonic beam divergence angle based on measured width of the ultrasonic wave-front at the rolling surface, and the assumed region of the rolling surface that is "visible" to the source element

corresponding width at the rolling surface calculated. As shown on the figure, the limiting angle for usable energy within the incident ultrasonic beam was found to be approximately 2° which would consequently represent a 1.1$mm$ arc of the rolling surface.

Based on the above, it could be concluded that the installed ultrasonic element would be capable of resolving the oil film thickness within contacts of 1.1$mm$ width (2$b$) or greater. Given the proposed test load of 1140$kN$, as defined in §4.2, and the fact that the sensor would be located at bearing BDC and would therefore observe maximum internal bearing load, $Q_{max} = 380kN$ (according to Equation (2.1)), the expected Hertzian contact width at this location would therefore equate to 2.14$mm$ (2$b$, according to Equation (2.8)) and would therefore be well above this "minimum observable width."

The consequence should contact widths become smaller than the effective observation width of the ultrasonic beam however was suggested in [72], where it was stated that an over prediction in ultrasonic reflection may occur. This was explained, as illustrated in Figure 4.18, through consideration of the ultrasonic wave as a series of discrete parallel beams, which consequently experience varying degrees of reflection at the measurement interface due to the varying thickness of the lubricant film outside the contact region. The discrete wave segments are then received back across the face of the single source element, whose overall response will then be a summation of the effects induced by each individual component of the reflected beam. The recorded response is then
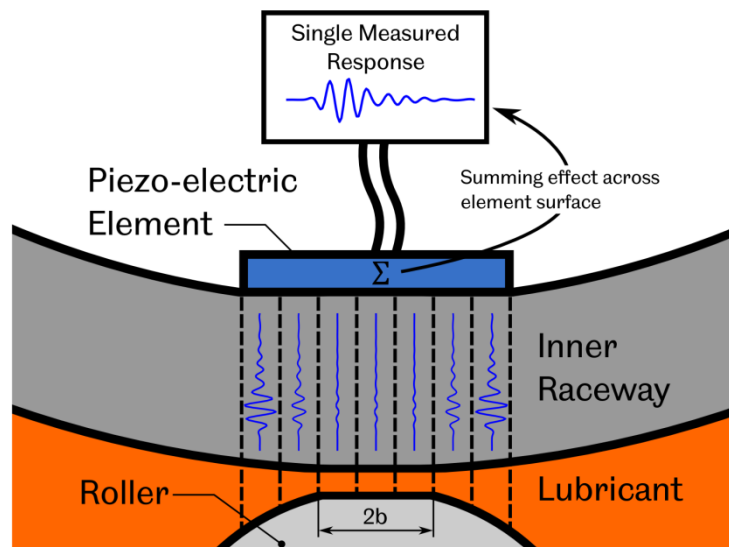


Figure 4.18: Consequence of non-parallelism across the ultrasonic wave-front leading to high observed reflection from the interface thus resulting in over-prediction of central film thickness (adapted from [72])

effectively an average across the whole wave-front, thus resulting in a larger reflection coefficient value as regions outside the contact (higher $R$) influence the response of the lower reflections within the contact region.

## 4.4  Conclusions

- A bespoke test rig has been developed to test cylindrical roller bearings representative of those used in the planetary stage of a $2 - 3MW$ WTGB. The key aim of the test rig was to induce failures in bearings such that the failure mode would be representative of those observed in the field.
- Test conditions were defined based on ISO281 and ISO16281 life rating calculations that would theoretically permit a full bearing life cycle test to be performed in around 2 weeks.
- A method to simulate the impact events known to occur in-service was also defined through the use of geometric stress raising features in test bearing inner raceway rolling surfaces.
- An appropriate means to instrument test bearings with ultrasonic sensors has also been established, utilising piezoelectric elements permanently bonded to test bearing inner raceway bore surfaces at the point of maximum bearing load.
- The ultrasonic energy emitted from a raceway bonded sensor was measured to determine the divergence of ultrasonic energy through the raceway material, and to subsequently determine the effective measurable width of the incident ultrasonic beam.
- A measurable width of $1.1mm$ was established, which was established as being lower than the expected contact width under proposed test loads, the sensor was therefore considered to be well suited to the application.

# Chapter 5

# INDEXING STRATEGY ANALYTICAL MODEL

As a precursor to the testing phase of this study, based around the test rig detailed in the previous chapter, a theoretical approach was employed to understand the potential advantages of raceway indexation in terms of bearing life enhancement. This chapter describes an analytical model that was developed to predict bearing life for an indexing bearing and identify regimes of raceway indexation at which bearing life may be optimised.

## 5.1 The Basis for an Indexed Bearing Life Model

As discussed in §2.2, bearing failure via rolling contact fatigue is driven by sub-surface material yield, which develops cyclically as a result of sub-surface shear stresses induced during over-rolling, and is thus a progressive failure mechanism. The failure of planetary stage WTGB bearings however is understood to be initiated by impact events, which effectively circumvent the incremental sub-surface damage mechanism, and may effectively induce sub-surface yield instantaneously, thus vastly enhancing the onset of spalling failure.

The overall rolling contact fatigue failure mechanism is considered the main factor pertaining to bearing life, and thus forms the basis of the bearing life rating principles introduced in §2.2. Consequently, it was considered appropriate for the analysis of an "Indexed Bearing Life Model" to also draw upon the standard bearing life rating principles and thus

provide the basis for the analysis. In the development of the basic model, only failure via classical rolling contact fatigue was initially considered, and the initiation (impact) events were dealt with subsequently, as will be discussed later in the chapter.

## 5.2 Elemental Modelling of Bearing Life

Standard life rating methods consider the fatigue life of a bearing as a whole, which presented two initial problems. Firstly, since the MultiLife principle relates specifically to bearings with non-rotating inner raceways, and failures are known to propagate from this location as a consequence, it was necessary to focus attention on the fatigue damage induced within the inner raceway material only. Furthermore, in order to analyse the effects of raceway indexation, the inner race needed to be discretised into a series of circumferential elements such that life ratings could be applied individually to each. In this way, fatigue accrual could then be appropriately calculated across only the segment of bearing material under load, and consideration could be made to the fact that an indexed inner raceway would develop varying degrees of fatigue damage throughout its life based on its position relative to the bearing loaded zone.

The basic life rating of a bearing, as defined by Equation (2.5), considers its dynamic load rating $C$ and the equivalent radial load $P$ (both of which are expressed in Newtons). This is reproduced in Equation (5.1), where $P$ is replaced with $F_r$ for this application, since applied loads here were taken to be purely radial. Furthermore, since operation on the test platform would be at constant rotational speed $N_r$ (rpm), it was possible to further define the life rating in terms of total operational hours, $L_{10(h)}$ through addition of a speed factor $a_r$.

$$L_{10(h)} = a_r \left(\frac{C}{F_r}\right)^{\frac{10}{3}}$$

(5.1)

Where,

$$a_r = \frac{10^6}{60N_r}$$

(5.2)

The bearings under consideration were known to fail at their inner raceways, primarily due to the fact that this is the stationary component in planetary bearing applications, as discussed in §1.2.2. In fact, more advanced bearing life rating methods are available [91], permitting adjustment of the bearing life rating equations to account for such alternative modes of operation, however – for simplicity – such analyses were not considered during development of the model presented here. As such, it was postulated that the *standard* life rating of Equation (5.1) would provide a reasonable approximation to the fatigue life of the inner race material.

Considering bearing failures in general; failure will typically propagate from one discrete point within the race and, in a purely fatigue based failure, this will generally relate to the region in which load is applied. Since a bearing may be installed in any orientation about its axis of rotation, then the stationary raceway could be said to be just as likely to fail in any given orientation, and thus possess the same statistical life rating in any case. As such, if the inner race was split into a series of individual segments of equal size it was assumed that each element would possess the same *initial* life rating, that is; before the bearing has accrued any operational cycles.
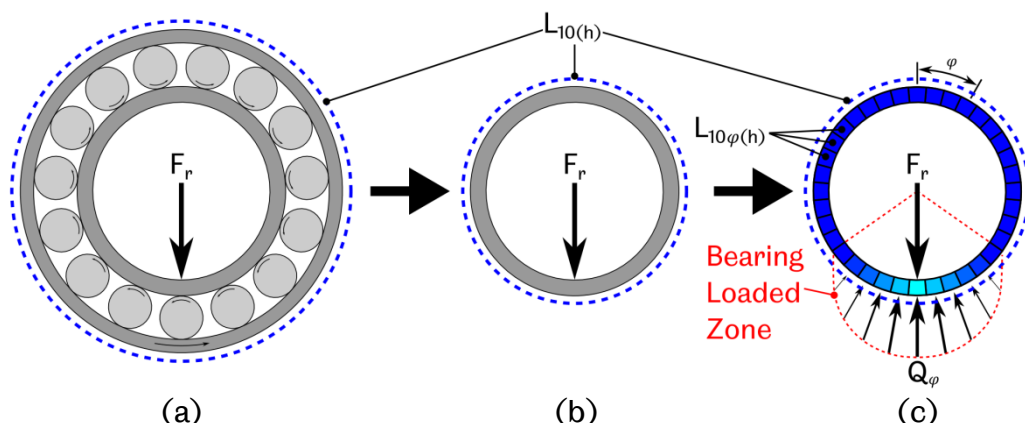
$$L_{10(h)} \equiv L_{10\varphi(h)} \tag{5.3}$$



Figure 5.1: (a) and (b) Development of the effective $L_{10(h)}$ life for a MultiLife Bearing and (c) subsequent representation at an elemental level

An elemental life rating was thus developed – through modification of Equation (5.1) – as a function of the load on the raceway element $Q_\varphi$, and

95

its elemental load rating $C_\varphi$, as shown in Equation (5.4), where $\varphi$ represents the angular position of the raceway element from bearing top-dead-centre (TDC).

$$L_{10\varphi(h)} = a_r \left( \frac{C_\varphi}{Q_\varphi} \right)^{\frac{10}{3}} \tag{5.4}$$

As illustrated in Figure 5.1(c), elemental life could then be said to be consumed due to local loading of individual raceway elements. For a non-indexing bearing, elemental load outside the bearing loaded zone would never be consumed and would effectively be infinite. Elemental life within the loaded zone however would be consumed at a rate proportional to the internal bearing load acting upon the given raceway element; where the amount of life available to each element is dictated by the initial global $L_{10(h)}$ life, and the ultimate failure of the bearing would arise at the first instance where the life of any individual elemental was completely consumed.

The normal load on each raceway element, $Q_\varphi$, was established based on Equation (5.5) [21].

$$Q_\psi = Q_{max} \left[ 1 - \frac{1}{2\epsilon} (1 - \cos \psi) \right]^{1.11} \tag{5.5}$$

Where $Q_\psi$ is the normal load at angular position $\psi$ away from the location of $Q_{max}$ (bearing BDC in this case). $\epsilon$ is defined according to Equation (5.6), and is a function of the bearing diametral clearance $c_d$, and the "ring radial shift" $\delta_r$.

$$\epsilon = \frac{1}{2} \left( 1 - \frac{c_d}{2\delta_r} \right) \tag{5.6}$$

Physically, $\delta_r$ represents the radial deflection of the bearing rings at $\psi = 0$ under the applied bearing load, which – in reality – is difficult to define. However, a procedure is described in [21] to define it through the iterative adjustment of $\delta_r$ until the vertical component of all internal bearing loads, $\sum Q_\psi \cos \psi$ match the applied radial bearing load $F_r$.

As discussed above, the bearing life rating principles are formed on the basis that fatigue failure in bearings is driven by sub-surface shear stresses that are induced within the raceway material during rolling contact. As such, Equation (5.4) was further manipulated to the form of Equation (5.7), where elemental life was expressed as a function of the maximum principal shear stress induced within the raceway element, $\tau_{1\varphi}$, calculated according to Equations (2.7) and (2.10), based on the normal contact load $Q_\varphi$ for the given raceway element, as defined in Equation (5.5) above. Maximum principal shear stress, $\tau_1$, was chosen for simplicity over that of orthogonal shear stress, $\tau_0$, given the fact that the former is known to act directly below the contact and could therefore be directly related to the load applied to the given raceway element.

$$L_{10\varphi(h)} = a_r \left( \frac{C_{\sigma\varphi}}{\tau_{1\varphi}} \right)^{\frac{10}{3}} \tag{5.7}$$

Here, a new parameter $C_{\sigma\varphi}$ is introduced, termed the *Dynamic Stress Rating*, which was used to represent the cumulative cyclic stress capacity of each raceway element.

## 5.3  Indexing and Indexing Strategy

An "Indexing Strategy" for a MultiLife bearing essentially involved definition of two key parameters; the "Index Period" and the "Index Angle". The former defining the period of time, or number of cycles accrued by the bearing before its stationary raceway was indexed, and the latter describing the magnitude by which the raceway would be rotated at each instance of indexation. Clearly, variation of these two parameters alone would provide a vast range of potential indexing strategies for a MultiLife bearing, and as such, a method was required to assess such a broad range conveniently and effectively. To limit the scope of analysis, a constraint was imposed that both index angle and period would remain constant throughout a given indexing strategy.

In establishing the initial dynamic stress rating of all the bearing elements; $C_{\sigma\varphi}^{[0]}$, it was assumed, based on earlier discussion, that any part of the bearing inner raceway material would possess the same statistical likelihood of failure, and thus the initial stress carrying capacity of any

given element would be the same for all such elements across the bearing circumference, such that:

$$C_{\sigma\varphi}^{[0]} = C_{\sigma}^{[0]}$$

(5.8)

Furthermore, prior to any loading or operational cycles being performed, this stress carrying capacity would be purely a function of the bearing material and geometry, and would thus be linked to the global $L_{10(h)}$ bearing life as per the standard life rating procedures [12]. Accordingly, Equation (5.7) could be rearranged to the form of Equation (5.9), and considered initially for the case where $\varphi = 180°$; relating to the point of maximum internal bearing load, thus representing worst-case loading.

$$C_{\sigma}^{[0]} = \tau_{1\{\varphi=180\}} \left(\frac{L_{10(h)}}{a_r}\right)^{\frac{3}{10}}$$

(5.9)

Global $L_{10(h)}$ bearing life was therefore established based on the standard procedures outlined in [12], where consideration could also be made to the effects of lubrication, contamination and the fatigue stress limit of the bearing, through inclusion of the relevant life modification factors.

An iterative process was then adopted, where the stress rating of each bearing element was assessed at every inner raceway "rest position" throughout a given indexing strategy (considering the time spent at each static position as dictated by the index period $x$, in hours). The reduction in elemental stress ratings could thus be tracked, and total "Indexed Bearing Life" established through identification of the first raceway element whose stress rating was reduced to zero. Bearing life could then be found through assessment of the number of index periods through which the raceway had survived prior to this. The iterative process by which a given indexing regime was assessed is detailed as follows.

Elemental life for the first index period, $L_{10(h)\varphi}^{[1]}$, was calculated for each element according to Equation (5.10), where for elements outside the loaded zone $\tau_{1\varphi} = 0$ and thus $L_{10(h)\varphi} = \infty$.

$$L_{10(h)\varphi}^{[1]} = a_r \left( \frac{C_{\sigma\varphi}^{[0]}}{\tau_{1\varphi}^{[1]}} \right)^{\frac{10}{3}}$$

(5.10)

The bearing would then operate normally for a given number of hours, $x$ (the *index period*), during which elemental life would consequently be consumed. The life remaining, $L_{10(h)\varphi r}^{[1]}$, within each raceway element after this operational period could thus be determined by simply subtracting the bearing life at the start of the index period by the index period duration, $x$.

$$L_{10(h)\varphi r}^{[1]} = L_{10(h)\varphi}^{[1]} - x$$

(5.11)

Consequently, stress carrying capacity of those elements placed under load would be reduced, and a new elemental dynamic stress rating could then be determined through rearrangement of Equation (5.10):

$$C_{\sigma\varphi}^{[1]} = \tau_{1\varphi}^{[1]} \left( \frac{L_{10(h)\varphi r}^{[1]}}{a_r} \right)^{\frac{3}{10}}$$

(5.12)

The raceway would then be indexed by $\theta$ degrees (the *index angle*), and each raceway element would subsequently be placed under a different load, $Q_\varphi^{[2]}$, based on its new position relative to the bearing loaded zone ($\varphi^{[2]} = \varphi^{[1]} + \theta$). As such, each element would then experience a different maximum principal shear stress, $\tau_{1\varphi}^{[2]}$. Elemental life at the new location could again be established from Equation (5.10) which, for the second index period, would become:

$$L_{10(h)\varphi}^{[2]} = a_r \left( \frac{C_{\sigma\varphi}^{[1]}}{\tau_{1\varphi}^{[2]}} \right)^{\frac{10}{3}}$$

Similarly, the life remaining after the second index period, again of duration $x$ hours, could then be calculated:

99

$$L^{[2]}_{10(h)\varphi r} = L^{[2]}_{10(h)\varphi} - x$$

And the further de-rating of elemental dynamic stress carrying capacity could be observed according to:

$$C^{[2]}_{\sigma\varphi} = \tau^{[2]}_{1\varphi}\left(\frac{L^{[2]}_{10(h)\varphi r}}{a_r}\right)^{\frac{3}{10}}$$

This procedure was repeated until the condition defined in Equation (5.13) was satisfied at any point around the bearing, thus indicating that a raceway element would fail during the $n^{th}$ index period.

$$L^{[n]}_{10(h)\varphi} \le x \tag{5.13}$$

As such the total life to failure of the given raceway element could then be established from Equation (5.14).

$$L_{10(h)\varphi T} = x(n-1) + L^{[n]}_{10(h)\varphi} \tag{5.14}$$

This was performed for all raceway elements at index period $n$ that satisfied condition (5.13). Total bearing life for the given indexing strategy was thus identified from the minimum instance of $L_{10(h)\varphi T}$.

$$L_{10(h)T} = \min_{0<\varphi\le 360}\left\{L_{10(h)\varphi T}\right\} \tag{5.15}$$

### 5.3.1 Application of the Initial Model

The process presented above is illustrated in Figure 5.2, for a MultiLife Bearing whose inner raceway is indexed by 30° ($\theta = 30$) every 29 hours ($x = 29$). Calculations were performed considering the bearing test conditions proposed in §4.2 for accelerated life; where outer race rotational speed would equal $100rpm$, a constant radial load of $1140kN$

would be applied and operating temperature was estimated to equal around $60°C$. As discussed previously, $L_{10(h)}$ life under the proposed conditions was found to equal 69.4 hours.

Figure 5.2 illustrates the incremental reduction in the dynamic stress rating, $C_\sigma$, of all the inner raceway elements throughout the indexing regime. The initial stress rating, $C_\sigma^{[0]}$, of each of the elements was equal to $2.9GPa$, as shown by the outer black curve in each of the plots. The extent of the bearing loaded zone is illustrated by red lines in each of the plots, where normal bearing clearance and the applied load in the current case resulted in load distribution across an arc spanning 111.6° to 248.4°.

Within the first 29 hour index period (Figure 5.2(a)), elemental stress rating reduction only occurs for elements residing within the loaded zone,
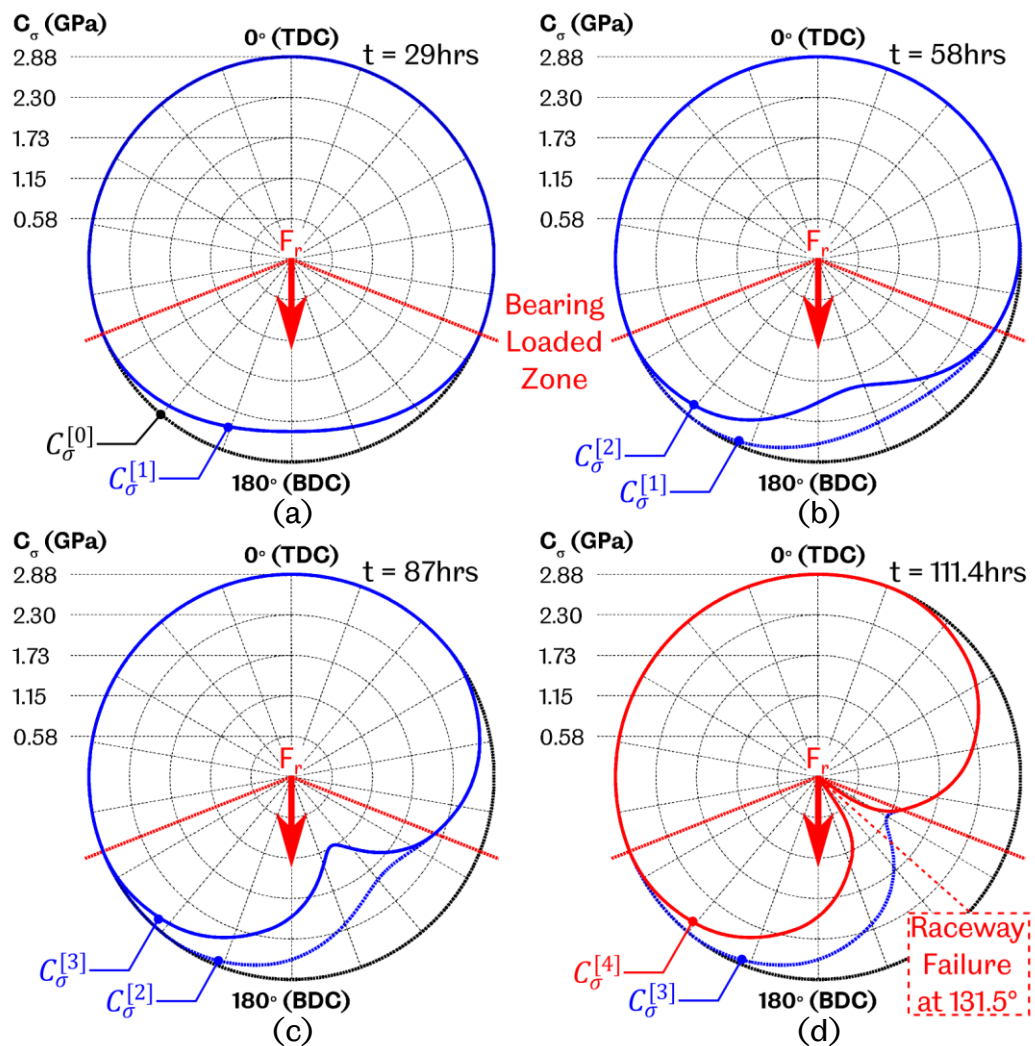


Figure 5.2: Incremental reduction in the dynamic stress rating and subsequent failure of a MultiLife bearing over four indexing periods (a)-(d), with a $1140kN$ applied radial load ($F_r$), and a regime of 30° indexation every 29 hours ($\theta = 30$, $x = 29$)

and the magnitude by which each element is de-rated occurs proportionally to the magnitude of the peak principal shear stress induced within each raceway element. As such, stress rating is minimised at bearing BDC to a magnitude of $2.4GPa$. The raceway is then indexed by 30° anti-clockwise and thus some of the previously loaded elements are be moved out of the bearing loaded zone to the right, as shown by the broken curve in Figure 5.2(b), and fresh, un-loaded elements are subsequently introduced at the opposite side. Over the subsequent index period (up to $t = 2x = 58$ hours) further fatigue damage is induced, resulting in the $C_\sigma^{[2]}$ curve as shown, which is then indexed again prior to the third index period as shown in Figure 5.2(c). Ultimate bearing failure occurs in the fourth index period (Figure 5.2(d)), where $C_\sigma$ is reduced to zero at 131.5° as illustrated. Applying Equation (5.14) here gives a life rating of 111.4 hours for the failed element, thus indicating the total theoretical bearing life achievable from this particular indexing scenario at a 10% rate of failure ($L_{10}$). This analysis therefore indicates that a potential 60% enhancement to bearing life may be achievable through utilisation of the given strategy of indexation, when compared to the $L_{10(h)}$ life under standard bearing operation as discussed above.

What is also evident throughout the process of raceway indexation is that the point of failure would no longer necessarily occur at the point of maximum bearing load. In fact, as observed in Figure 5.2, it appears that if highly stressed elements are not moved sufficiently far away from peak radial load during indexation, the rate of fatigue may be slowed, but ultimate failure of such critical elements could still be triggered at any point within the loaded zone.

## 5.3.2 Investigating a Range of Indexing Regimes

Upon confirmation of the suitability of the model, a wider range of indexing strategies were assessed to determine the optimum regime for raceway indexation. Figure 5.3 presents a matrix of bearing life results for all strategies of indexation within the ranges $0.1 \leq x \leq 48$ hours and $0 \leq \theta \leq 90$ degrees, in the form of an intensity plot. The plot has been normalised to illustrate the relative enhancement to $L_{10(h)}$ bearing life under no indexing regime and at the same operational conditions as those discussed in §5.3.1, as per the definition in §4.2. Clearly, application of an index period of 0 hours is meaningless, hence the limit applied to the y-axis here. An index angle of 0 degrees essentially represented normal bearing operation, and bearing life enhancement was thus equal to 1

(indexed life is equal to the standard $L_{10(h)}$ life) for all such cases across the plot.
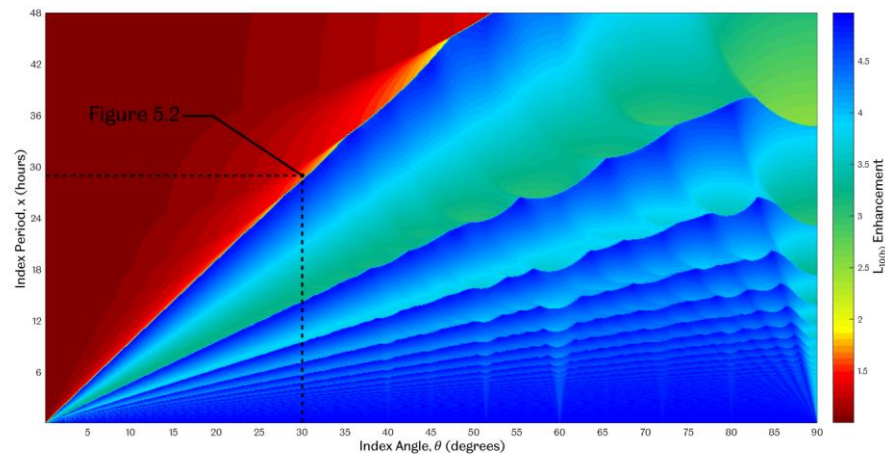


Figure 5.3: Intensity plot to illustrate the enhancement to the standard $L_{10(h)}$ life for a MultiLife Bearing run at $100rpm$ and $1140kN$ under a range of indexing strategies

What is immediately apparent from the figure is that theoretical life enhancements of up to $5$ times the standard $L_{10(h)}$ bearing life have been identified. Shorter index periods and larger index angles generally appear to be beneficial, and indeed for very small index periods ($x \leq 1$) bearing life is consistently high for $\theta \geq 2$. Additional optimum zones exist however, illustrated by the dark blue bands that run diagonally across the plot with a positive gradient. These striations move closer together and reduce in slope, gradually becoming horizontal as they merge with the blue region at the bottom of the plot. In contrast, the red region in the top left of the plot signifies the area of lowest relative bearing life enhancement; the example in Figure 5.2 illustrates an indexing scenario within this region.

The general patterns observed in Figure 5.3 are logical, given that the longer the inner raceway remains stationary (large $x$); the more fatigue damage is accrued within the localised portion of raceway material under load. Consequently, a higher index angle must be employed to move such critical elements away from peak load in order to avoid further damage and maintain a high bearing life. Conversely, if indexing occurs almost continuously (very low $x$), raceway elements do not reside within the loaded zone for any significant duration, and thus localised damage takes much longer to reach a significant level, resulting in the high bearing lives observed for $\theta \geq 2$ and $x \leq 1$ here.

Indeed, the red region of the plot, where bearing life is significantly lower than all other areas, occurs due to a combination of long periods of operation with relatively small angles of indexation. As illustrated in Figure 5.2, a significant amount of damage is accrued by raceway elements residing within the loaded zone per index period, and such elements are not moved sufficiently far away from the loaded zone before their ultimate failure takes place. As such, bearing life does not extend beyond the first few indexing cycles in these cases.
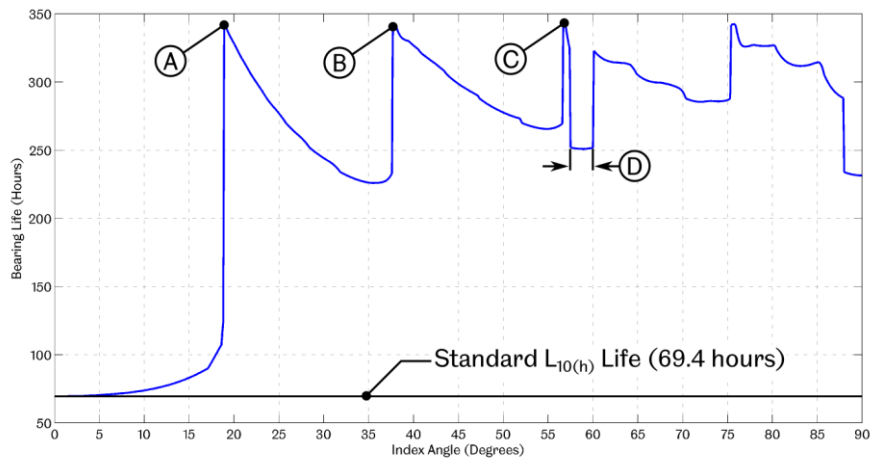


Figure 5.4: Bearing life for all indexing strategies where $x = 18$ hours and $0 \leq \theta \leq 90$ degrees

Figure 5.4 shows the relationship between theoretical bearing life and index angle when a constant index period of 18 hours is utilised. The region of low relative life enhancement, as discussed above, occurs from 0 to 17 degrees for the given index period. Beyond this, bearing life rises sharply and peaks briefly at approximately 18.9°, this however is not maintained as index angle is increased further, where a gradual decay is subsequently observed. This pattern of sharp rise in bearing life followed by gradual decay repeats with increasing index angle.

Figure 5.5 shows the dynamic stress rating reduction curves for the indexing scenario represented by **Point A** on Figure 5.4, where the inner raceway is rotated by 18.9 degrees at 18 hour intervals. Similarly to the case presented in Figure 5.2, elemental damage builds up significantly over the first few indexing cycles (7 in this case), yet the index angle here is such that critically damaged elements are just moved out of the loaded zone before an elemental failure can occur. Consequently, bearing life is dramatically increased because a stabilised phase of operation is reached where, during each subsequent index period, a fatigue peak of the same

magnitude $(0.53GPa)$ is induced at 123.3°, which is then moved just outside of the loaded zone and thus does not progress any further. The inner raceway can subsequently undergo a complete revolution of indexes during this stabilised running phase and failure will only then occur once such near-critical elements move back into the loaded zone.
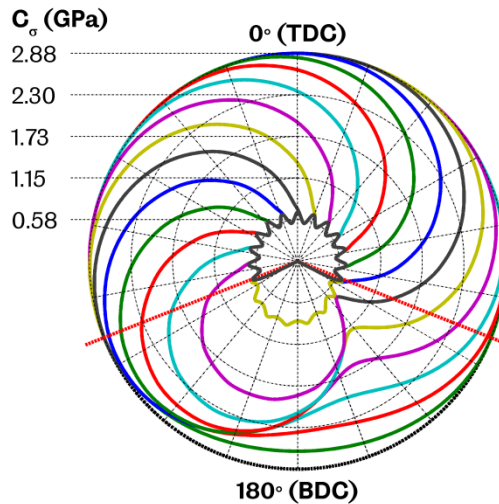


Figure 5.5: Reduction in $C_\sigma$ per index period for $x = 18$ hours and $\theta = 18.9°$

Bearing life incrementally reduces beyond **Point A** in Figure 5.4, as index angle is increased within the range $19 \le \theta \le 36$. Here, stress rating reduction per index period occurs similarly to Figure 5.5, however, larger angles of indexation result in critically stressed elements moving back into the loaded zone sooner, thus reducing bearing life. In addition to this, the radius of the stabilised stress rating band that forms – allowing the bearing to complete a full revolution of indexes – becomes increasingly large as index angle increases (i.e. elemental damage becomes less critical, moving away from $C_\sigma = 0$). This is illustrated for $\theta = 20, 25$ and 30° in
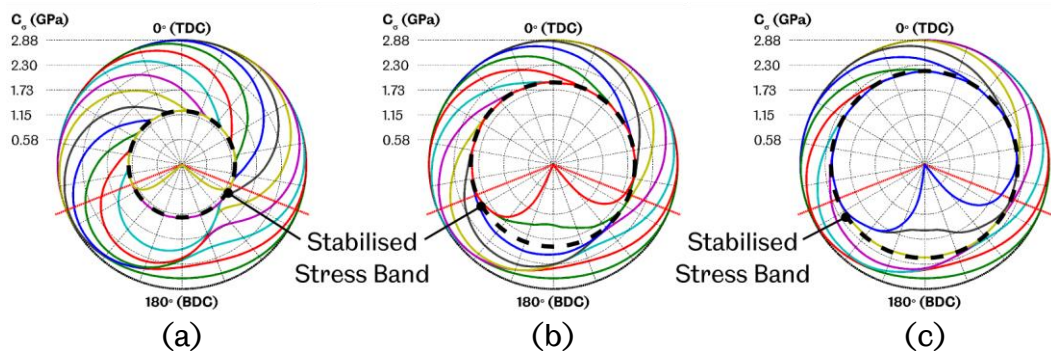


Figure 5.6: Incremental widening of the stabilised stress rating band for index angles of (a) 20, (b) 25 and (c) 30° at $x = 18$ hours

Figure 5.6(a), (b) and (c) respectively, illustrating the process by which raceway indexation becomes increasingly less efficient.

This subsequently results in a further bearing life peak at **Point B** ($\theta = 37.8°$), where stress build up occurs at a particularly low rate, thus permitting two stabilised stress rating bands to form as the bearing is now able to complete two full revolutions of indexation prior to ultimate elemental failure.
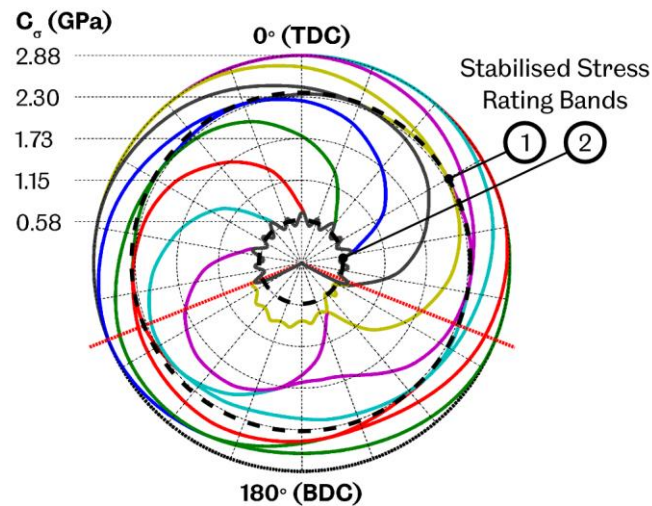


Figure 5.7: Reduction in $C_\sigma$ per index period for $x = 18$ hours and $\theta = 37.8°$

This pattern is found to repeat for the next peak in the bearing life curve of Figure 5.4 and indeed for every other peak in bearing life that can be observed. This explains the formation of the blue striations in Figure 5.3, where each consecutive band – moving from left to right across the figure, is formed as a result of the bearing undergoing an additional full rotation of indexes prior to failure. As such, **Point C** of Figure 5.4 occurs as a result of three full bearing rotations becoming possible due to the much wider angle of indexation.

A break in the general trend is however observed in the subsequent decay following **Point C**, as index angle is further increased. Across **Point D** (within the region $57.5 \leq \theta \leq 60°$), bearing life is relatively low, rising again beyond this and following a more steady rate of decay, as observed previously, as index angle is further increased. Further analysis of $C_\sigma$ curves prior to, during, and subsequent to **Point D**, as shown in Figure 5.8, indicates this to be a result of temporary breakdown in the third stabilised stress band as shown in Figure 5.8(b).

106

Figure 5.8: $C_\sigma$ curves for (a) $\theta = 56.8$, (b) $\theta = 58.6$ and (c) $\theta = 60.1°$ index angles at $x = 18$ hours

Clearly, a more uneven stress band develops across the raceway elements by the third full raceway rotation as a consequence of the much larger index angles now being employed. As such, under certain angles of indexation the raceway regions where peak stresses have been induced will be indexed in-phase with the location of peak internal bearing load as these elements move back into the loaded zone on 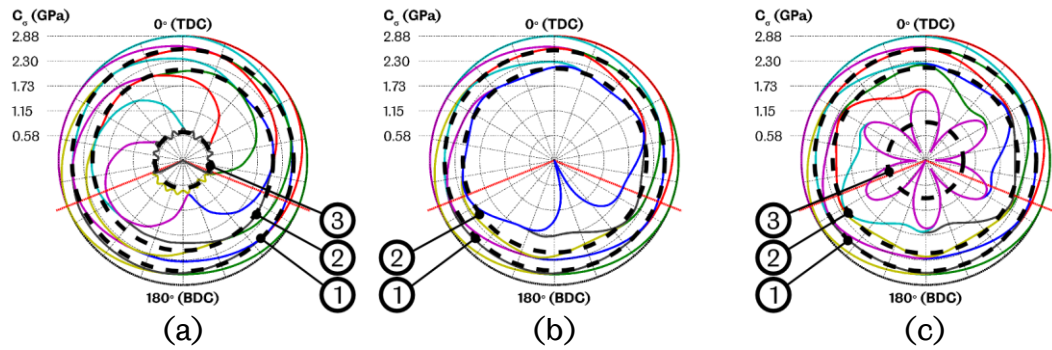their second full raceway rotation. As a result, such critically stressed raceway segments will sustain the largest degree of further damage as they sweep through the loaded region, causing those elements to fail before they can move through and out of the other side. Therefore, when index angle is such that this can occur, indexation to three full raceway rotations is not possible and bearing life is compromised as a result. Once the phasing of peak raceway stresses does not coincide with peak loading again (by 60.1° in the case of an 18 hour index period as shown in Figure 5.4) elemental damage no longer results in complete failure, a third stress rating band is again formed, and an additional complete raceway revolution is once again attainable.

This effect is again observed at $\theta = 90°$, resulting from the same phasing effects inducing significantly lower bearing lives. Indeed, when considering the full bearing life matrix of Figure 5.3, it seems that the region around 90° is generally detrimental to overall bearing life enhancement for all values of $x$. This low life region expands with increasing index period, beginning from around 75° by $x = 48$ hours. This is a result of higher loading durations inducing larger fatigue peaks within raceway elements, thus causing greater variability in stress bands and therefore increasing the general sensitivity to premature elemental failure across a wider range of index angles.

For this reason, bearing life enhancements become much more variable as both index angle and period are increased. This can be observed in Figure 5.4 as an increase in the waviness of the general sawtooth curve profile for $\theta \geq 60°$. This is also evident upon the diagonal striations in the full bearing life matrix of Figure 5.3 as waviness can also be seen to become more intense towards the top right of the plot.

## 5.3.3 Optimum Strategies of Indexation

Development of the basic indexing model and its application to a variety of different operational regimes has provided a better understanding of the impact raceway indexation can have upon bearing life. The patterns observed in Figure 5.3 have been explained through the above analysis, leading to a better understanding of how and why bearing indexation can be optimised.

In the general application of an indexing strategy for a MultiLife WTGB planetary stage bearing it seems that an optimum strategy is one which ensures at least one complete raceway rotation is achieved. Furthermore, in order to utilise as much of the load carrying capacity of the inner raceway as possible, the material should be loaded such that induced stresses take the raceway as close to failure as possible, whilst still ensuring a full rotation of indexes are achieved. The scenario illustrated in Figure 5.5 is an example of this, as well as any such regime running along **Line A** in Figure 5.9. However, from a practical standpoint, operation within this zone puts bearing life on a knife edge, and the implications of, for example; an indexing instance being delayed, or the actuation mechanism failing to achieve the prescribed indexing angle, could result in significant reduction to life as operation moves into the red region of the plot. Furthermore, the natural variation in bearing material quality and the variable occurrence of material inclusions means that, in fact, some bearings may be weaker than others, and thus may be more predisposed to failure.

Given such variability, maintaining both a small angle of indexation and a low index period may indeed be the most advantageous approach; giving greater confidence that bearing life will remain high despite fluctuations in the application of raceway actuation or the variation in bearing strength. Here, raceway actuation would take place almost continuously, resulting in several full rotations throughout the life of the bearing and meaning that only small amounts of damage would ever be accrued in any one portion of the material. The region denoted **B** in Figure 5.9 illustrates strategies

lying with this operation range. The practicality of achieving such regular actuation may not be feasible in reality however, due to the demand placed upon the mechanical system.
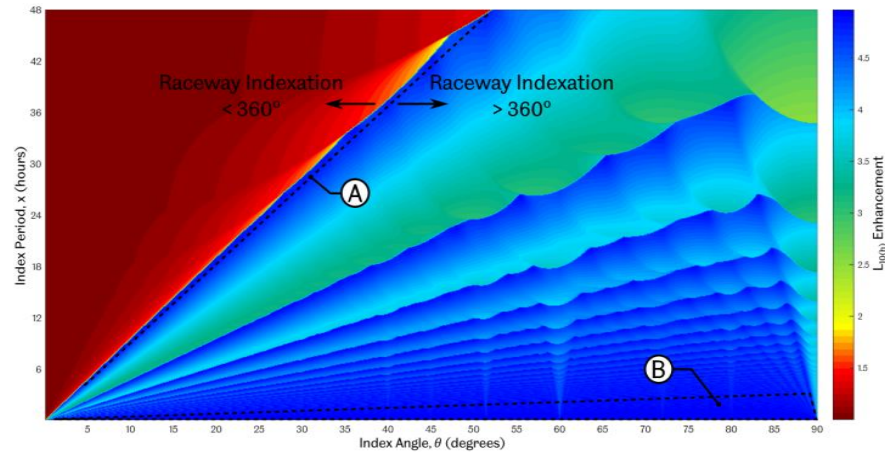


Figure 5.9: Illustration of the optimum strategies of raceway indexation

## 5.4 The Stress Concentrating Effect of a Defect

The analytical model presented thus far permits a statistical indexed bearing life to be evaluated based on classical rolling contact fatigue failure. As discussed in §5.1 however, fatigue lives of planetary WTGB bearings will likely be considerably shorter due to stress induced sub-surface material yield that will arise in the event of an impact, thus advancing rolling contact fatigue considerably. As discussed in §4.2.1, in order to simulate such events and appropriately accelerate life on the test platform, seeded defects were to be introduced in the form of geometric changes to the rolling surface. This would generate stress concentrations and thus advance fatigue failure in a similar way to that of the real scenario. The above model would thus require further development in order to assess the impact such surface changes would have to overall bearing life enhancement of a MultiLife Bearing under a given indexing strategy.

A Wire-Cut EDM process was to be utilised to generate a fine groove or "notch" across the entire rolling surface, normal to the direction of rolling. Such defects would be introduced at a width of approximately $0.2mm$, and

$0.1mm$ depth. Due to the geometry of the wire electrode, defects would effectively be of semi-circular geometry, as shown in Figure 5.10.
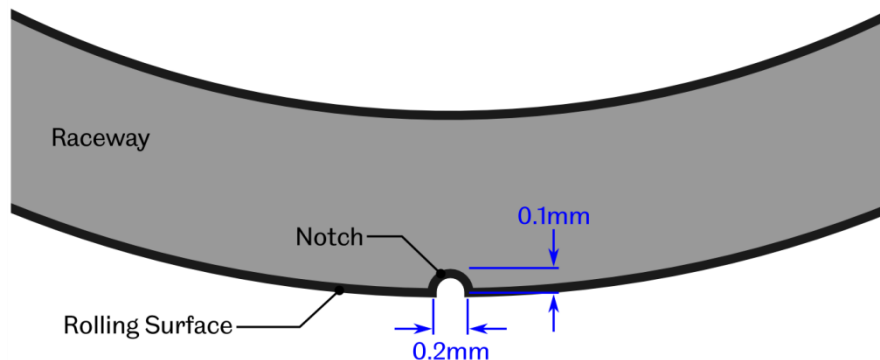


Figure 5.10: Exaggerated illustration of the notched inner raceway rolling surface

Features such as grooves, shoulders and holes in structural members generate localised regions of high stress within the body of the structure. These localised stresses are quantified by the degree to which peak stresses are increased in the presence of such features, as compared with the normal stress distribution of the un-deformed structure under the same loading conditions. Such multipliers are termed *Stress Concentration Factors* (SCFs), typically denoted by $k$ and evaluated as follows [92].

$$k = \frac{\sigma_{max}}{\sigma_{nom}}$$

$$k_s = \frac{\tau_{max}}{\tau_{nom}}$$

(5.16)

As shown in Equations (5.16), determination of $k$ requires knowledge of a reference stress value, $\sigma_{nom}$ or $\tau_{nom}$, that would typically represent either the stress at – or near to – the stress raising feature, or peak global stress within the structure before the feature is introduced. $\sigma_{max}$ and $\tau_{max}$ then represent the same stress parameters when the stress raising feature under analysis is present. Typical theoretical values can be found for standard structures and cross-sections that have been modified with various geometrical irregularities. SCFs can be calculated in terms of normal or shear stresses within the material, where a subscript $s$ is usually employed to indicate representation of the latter.

Values are generally determined through experimental techniques and measured through methods such as photoelasticity or the use of strain gauges, but finite element methods have also previously been used. Published values typically refer specifically to a *theoretical* stress concentration factor, to indicate that modifications are a function of the geometrical discontinuities alone, and are independent of material type [93].

Clearly, introduction of the proposed features to test bearing inner raceway rolling surfaces will generate such stress concentrations, thus affecting the maximum principal sub-surface shear stress, $\tau_1$, induced within raceway elements residing near to the geometric fault. This will act to enhance elemental life reduction when such features are placed under load, and will thus similarly influence overall bearing life. Similarly, higher stresses will be induced as a result of higher loading, and thus the probability of sub-surface yield and the subsequent onset of fatigue failure is increased accordingly in such cases as well. The remainder of this chapter investigates the stress concentrating effect of the defect proposed in Figure 5.10, in terms of the amplification to sub-surface shear stress that will be induced under the range of expected elemental loads under an applied bearing load of $1140kN$. Quantification of the magnitude of $k_s$ and its angular distribution away from the feature will be explored such that an elemental stress concentration function can be defined and included within the basic indexing model proposed earlier in this chapter, such that the effects of fault seeding can be included and assessed for any given strategy of raceway indexation.

### 5.4.1 Determining the SCF through Finite Element Modelling

Given the somewhat specific load case and geometries relevant for this application, the likelihood of finding appropriate stress concentration factors in literature was considered to be low. Furthermore, as discussed above, it was desirable to understand how much of the raceway material would be affected by such local surface discontinuities, such that stress concentration factors could be applied appropriately to individual raceway elements. Finally, in order to model such effects more accurately it was considered prudent to also take into account specific material properties of the bearing steel. As such, a finite element analysis (FEA) approach was deemed the most appropriate method to establish the stress concentrating effects of the seeded defect. Models were constructed in

the *HyperMesh* Finite Element Pre-Processor for analysis in *Abaqus Standard 2D*, the details of which are discussed presently.

### 5.4.1.1 Model Geometry

In order to represent the contact appropriately, both the rolling element and raceway were modelled. Since loading and geometry of the contacting bodies would be symmetric through the centre of the contact (in the plane normal to the rolling direction) it was deemed sufficient to model only half of the rolling element and inner raceway. In addition, since the stresses induced from the loading of the two bodies was not expected to extend particularly far from the contact, only a 20° segment of the inner raceway was modelled; considered to be more than sufficient to capture the full extent of the stress field and reduce overall computational time. Finally, since geometry of the two bodies would lead to the formation of a line contact, loading could be assumed to be uniform across the contact length, thus a 2 dimensional analysis was considered sufficient and the two bodies were modelled at a thickness of $0.01mm$.

### 5.4.1.2 Material Properties

NU2244 test bearing raceways are composed of standard 100Cr6 (AISI 52100/EN31) Chrome Bearing Steel, whilst rollers utilise a slightly harder 100CrMn6 Chrome Steel variant. Aside from hardness, both are known to have similar mechanical properties, and were therefore modelled identically in the analysis. Mechanical properties attained from [94] for 100Cr6 bearing steel were utilised, as detailed in Table 5.1.

| | |
|---|---|
| Modulus of Elasticity, $E$ | $210GPa$ |
| Poisson's Ratio, $v$ | 0.3 |
| Density, $\rho$ | $7810kg/m^3$ |

Table 5.1: Mechanical properties of 100Cr6 Chrome Bearing Steel used to represent the roller and raceway bodies in the FEA

It was assumed that loading of the two bodies would be entirely elastic so as to maintain agreement with Hertzian theory.

### 5.4.1.3 Boundary Conditions

Modelled geometry and boundary conditions are illustrated in Figure 5.11. Load was applied at Point B, up to a peak magnitude chosen to be

representative of the maximum roller/raceway contact load $Q_{max}$ under an applied bearing load of $1140kN$. Maximum contact load was found to equal $380kN$, and would subsequently be distributed across the entire length of the roller ($L = 82mm$) such that load per unit length would be $4.634kN/mm$. Considering the $0.01mm$ width of the modelled geometries, and the fact that only half of the contact was to be analysed, a peak load of $23.15N$ was subsequently applied at Point B. Simulation results were obtained at load increments of 10% up to this equivalent peak bearing load.

The inner raceway was modelled using a cylindrical coordinate system; $r$ and $\theta$ ($x = 0$), whilst the rolling element was modelled using standard Cartesian coordinates; $y$ and $z$ (again with $x = 0$ ), all of which are illustrated on Figure 5.11. As also illustrated in the figure, the inner raceway internal bore surface (surface 1) was constrained in both $r$ and $\theta$, since this surface would be supported by the stationary internal bearing shaft in reality, which was assumed to be rigid. The raceway line of symmetry (surface 3) was constrained in $\theta$ since displacements either side of this plane would be equal and opposite, and therefore the $\theta$ displacements at surface 3 could be assumed to be equal to zero. Deflection would however take place in the $r$ direction as a result of the
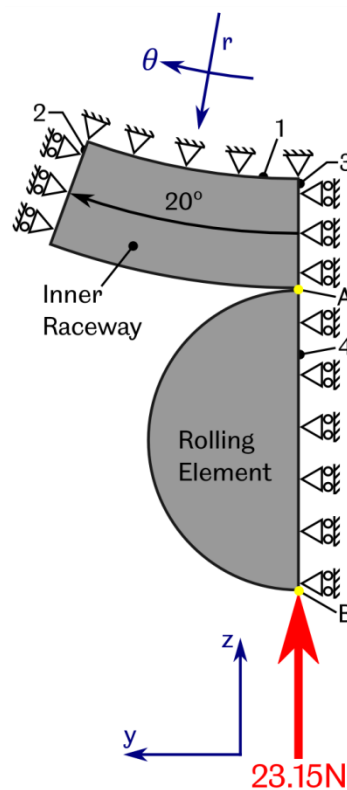


Figure 5.11: Illustration of the finite element model constructed to represent the roller-raceway contact

applied load, and the surface was therefore permitted to move freely in this axis to allow the raceway material to be compressed accordingly. Surface 2 was also constrained in the same way, however this was considered to be somewhat trivial, since stresses were expected to be negligible here due to the distance of this surface away from the loading point. Finally, the rolling element was also only constrained along its line of symmetry (surface 4) in the $y$ axis, so as to permit transferral of the load applied at Point B onto the raceway surface in a purely radial orientation. Given the two-dimensional nature of the model, analysis was performed for the condition of plane strain; that is $\varepsilon_x = 0$.

Finally, the contact interface was modelled as being non-frictional so as to be further representative of Hertzian theory.


### 5.4.1.4   Meshing of Components

To facilitate the analysis, the bodies were discretised with a combination of 3 and 4 node, plane strain elements (CPE3 & CPE4R).  It is well documented that such discretisation of geometries for the purposes of finite element analysis will inevitably result in errors.  An exact solution can only be obtained through application of an infinitely fine mesh, which is clearly unattainable in practice [95].  Ideally then, the smallest possible mesh elements should be applied to reduce error in results, however a trade-off must be made since utilisation of such small elements will greatly reduce computational efficiency.  It is therefore necessary to perform a mesh refinement study when developing any finite element model in order to understand the minimum number of mesh elements required to converge upon a result, such that a mesh independent solution is obtained and the greatest degree of accuracy is achieved.

This was carried out through iterative analysis of the un-notched raceway and rolling element, both modelled with an increasing number of mesh elements.  Mesh refinement was performed to both the inner raceway and rolling element simultaneously.  Components were meshed so as to create as even a distribution of elements as possible throughout each body.  As a starting point, both components were meshed such that the average element area was $4mm^2$ and subsequent refinement was then performed by quartering each individual element.

Since maximum principal shear stress was of key relevance to the indexing model, this parameter was used as the metric in the mesh refinement study.   As such, for each mesh refinement case, the "Max Shear"

parameter at each mesh element was obtained; relating to the maximum shear stress in any orientation within the 2D plane. The single maximum instance of this across the entire stress field was then extracted for each mesh refinement case. Results were only considered for stresses within the raceway structure, since only stresses arising within this component were of interest in the subsequent analysis. Figure 5.12 shows the obtained results, where shear stress is quoted against the number of elements comprising the raceway structure only, again due to this being the component of key interest.
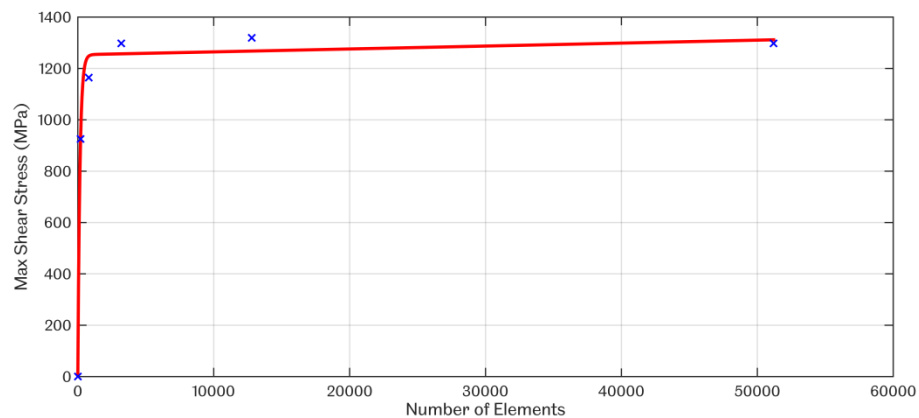


Figure 5.12: FEA results for maximum shear stress vs number of mesh elements for the modelled inner raceway segment

As can be seen from the figure, convergence to $\sim 1300 MPa$ is achieved through utilisation of 3200 elements or more. The finest mesh investigated comprised of 51200 raceway elements, corresponding to an average elemental area of $0.0156mm^2$. This therefore provided indication of an appropriate elemental size, and also an expectation of the maximum extent of the stress field within the raceway structure. Based upon this, the final mesh employed utilised only 8780 inner raceway elements, but was generated with biases such that the density of the mesh elements was gradually increased towards Surface 3 and Point A in Figure 5.11. As such, even higher resolution was achieved at the peak contact point, such that elemental area within this region was around $4.8 \times 10^{-5}mm^2$, ensuring more accurate representation of the stress field within this region, whilst maintaining computational efficiency. Similar biases were also applied to the rolling element, and the final meshed bodies are illustrated in Figure 5.13.
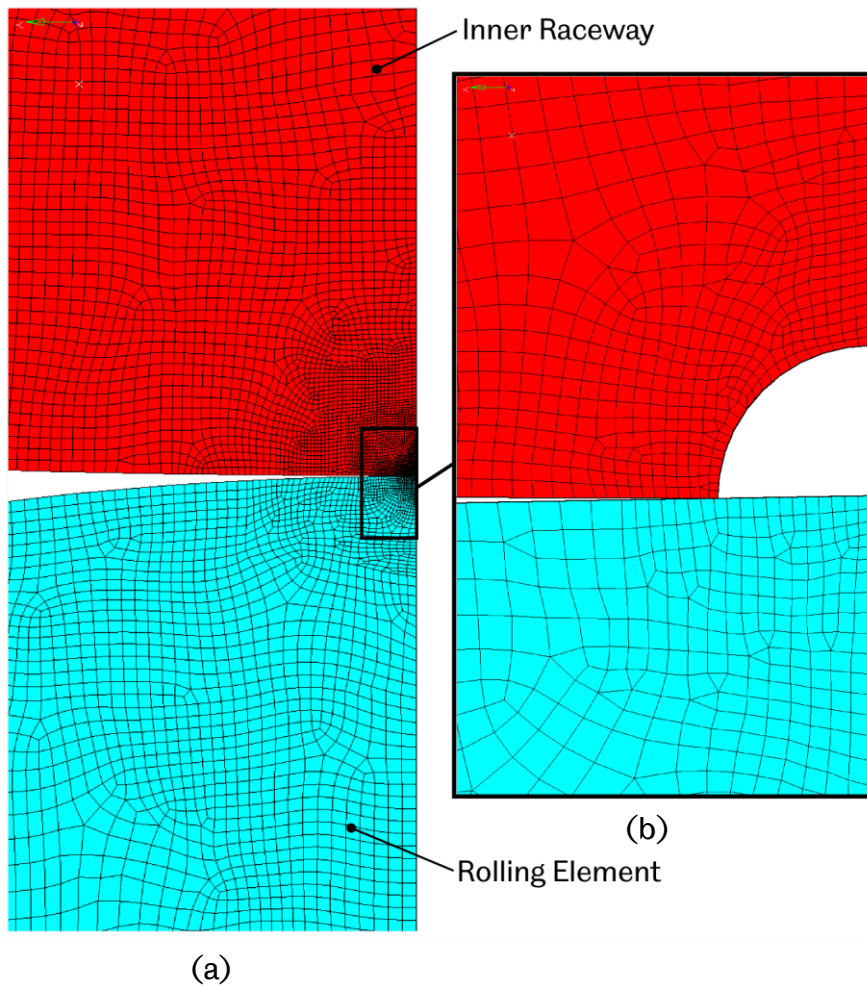
Figure 5.13: (a) Illustration of meshes applied to the rolling element and inner raceway bodies near to the contact centre point (b) close up of the mesh at the contact centre point for the fault seeded inner raceway

### 5.4.1.5  Suitability of the Model

To assess the accuracy of the model, simulation results were compared with Hertzian theoretical calculations for the same contact and loading conditions, as shown in Table 5.2.

|  | Hertzian | FEA | % Variation |
|---|---|---|---|
| **Peak contact pressure, $p_{max}$ (GPa)** | 2.76 | 2.83 | +2.47 |
| **Max principal shear stress, $\tau_{1max}$ (GPa)** | 0.83 | 0.89 | +6.74 |
| **Contact half width, $b$ (mm)** | 1.07 | 1.08 | +0.93 |

Table 5.2: Comparison between FEA model results and Hertzian theory

It is important to note the key assumptions of Hertzian theory when seeking validation for a contact model in this way so as to understand the applicability of the analytical methods to the problem in question. Indeed, when applying Hertzian analysis it is assumed that both the contact length is large by comparison to the reduced contact radius $R^*$ (as defined in Equation (2.9)), and that the contact area is comparably small [96]. In this case however, due to component dimensions and applied loads being relatively large, the contact length $L$ is in fact only around 3.7 times the reduced contact radius $R^*$ and furthermore; contact deformation and the resultant contact area will in fact be reasonably large such that $R^*$ will only be around 20 times greater than the theoretical contact half width $b$. Subsequently, validity of the theoretical model itself may be somewhat questionable, and the differences observed in Table 5.2 may in fact be due to errors in the analytical results.

Despite the above, it is clear from Table 5.2 that both model results and Hertzian predictions were found to agree well, where maximum variation was less than 7%. It is however apparent that model results provided higher stress and contact width values than those predicted theoretically. This may be due to the points highlighted above, but could also be attributed to the fact that the inner raceway here essentially represents a hollow cylinder, where bodies considered in a truly Hertzian case would in fact be solid. This would logically account for the slightly increased levels of stress calculated numerically, since the FEA model represents a thinner structure which would be more prone to deflection.

Figure 5.14(a) illustrates the principal shear stress field within the un-notched raceway body in the vicinity of Point A (Figure 5.11), as obtained from the FEA results and expressed as contours of $\tau_1/p_{max}$. Theoretical contours for an elastic line contact (as presented in [29]) are provided alongside the model results in Figure 5.14(b) for comparison. From this, the obtained FEA results can be seen to again agree reasonably well with the theoretical model. As shown in Table 5.2, the peak principal shear stress obtained from the FE model was found to be slightly higher than the $0.3p_{max}$ predicted by Hertzian theory, and as a result the $0.3p_{max}$ contour in Figure 5.14(a) bounds a larger area than that of the same contour in the adjacent theoretical plot. Similarly then, all additional contours are also slightly larger than their equivalent theoretical contours due to shear stress results being slightly higher throughout the entire stress field for the FEA model. Nonetheless, the location of the peak principal shear stress in Figure 5.14(a) coincides with that which is expected for a

Hertzian line contact; occurring at around $0.8b$, in-line with that presented in the adjacent plot as well as Figure 2.5.



(a)                                                                 (b)
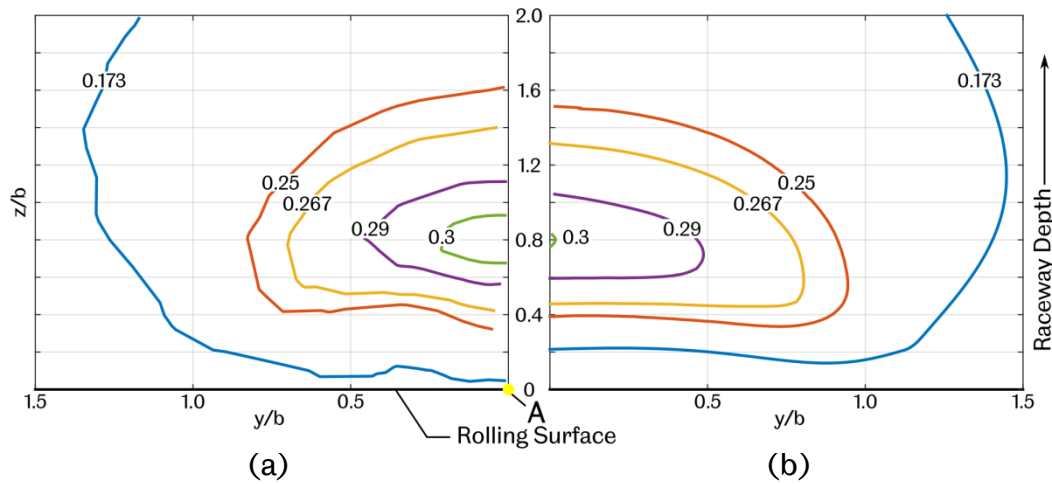
Figure 5.14: (a) Contours of constant principal shear stress normalised to maximum contact pressure ($\tau_1/p_{max}$) obtained from the FEA model for the un-notched raceway. (b) Comparison to the same contours obtained theoretically for the elastic contact of cylinders (adapted from [29]). Point A in Plot (a) is equivalent to that shown in Figure 5.11.

The above assessments were therefore considered sufficient to confirm validity of the FEA model, and this was subsequently utilised further to identify the stress concentrating effects of the defect under consideration.

### 5.4.1.6 Model Results

In determining the stress concentration factor from the FEA results the raceway stress field was investigated in angular strips as illustrated in Figure 5.15. Shear stress values through the entire raceway thickness at a given angular position were assessed and averaged to obtain a single shear stress for each corresponding raceway strip per angular position $\theta$. This was done similarly for both the notched and un-notched raceway results and a stress concentration factor was then obtained according to Equation (5.17) for each such angular position away from the contact centre point (Point A, Figure 5.11).

$$k_s = \frac{\overline{\tau_{1\theta u}}}{\overline{\tau_{1\theta}}} \tag{5.17}$$

Where $\overline{\tau_{1\theta u}}$ denotes the average principal shear stress obtained from the notched raceway at angular distance $\theta$ from the defect.



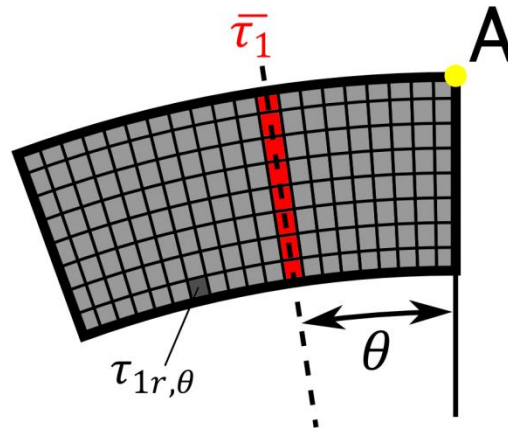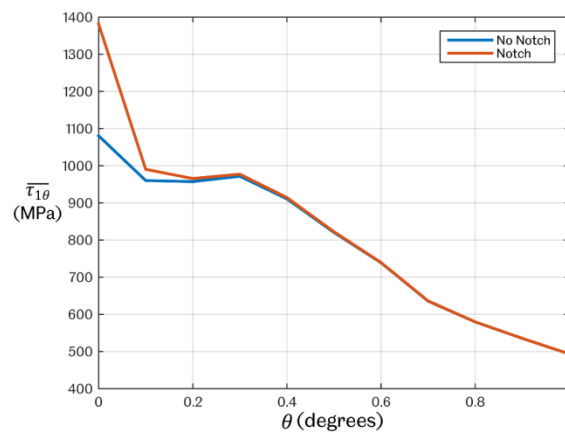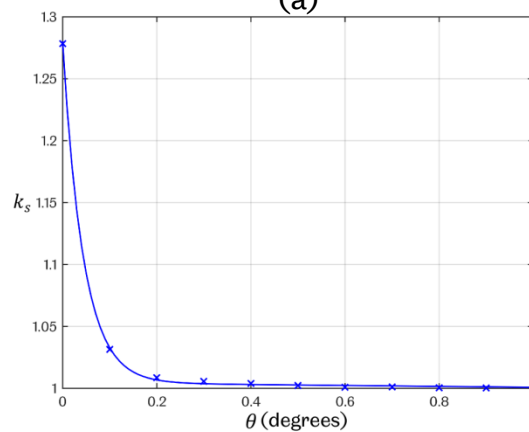Figure 5.15: Illustration of the averaging procedure to determine a single shear stress per angular position $\theta$



(a)



(b)

Figure 5.16: (a) Comparison of shear stress results obtained under an equivalent bearing load of $1140 kN$, with and without a surface defect present in the raceway, (b) Resultant stress concentration factor through the inner raceway vs. angular distance away from the defect

Figure 5.16(a) shows the principal shear stress results obtained from the model under equivalent maximum bearing load. Figure 5.16(b) then illustrates the resultant stress concentration factor, $k_s$, calculated from the two curves of Figure 5.16(a). As can be seen in the figure, $k_s$ peaks at 1.28 directly underneath the seeded defect, and then decays exponentially with increased angular distance away from it, settling quickly at 1. Indeed, it can be observed that the stress concentrating effect of the defect is extremely local, and does not extend through the raceway structure beyond roughly 0.6° from bearing BDC.

As discussed in §5.4.1.3, analysis was performed at load increments of 10% up to the peak internal bearing load $Q_{max} = 380kN$, arising from maximum applied radial load $F_r = 1140kN$. The peak stress concentration factor was found to increase with increased load applied to the contact interface. This is due to the increased amounts of deflection that will inevitably occur under higher loading, thus generating higher stresses around the geometric discontinuity. Such variation in both magnitude and distribution of $k_s$ with respect to load is shown in Figure 5.17, where curves are plotted from results obtained for equivalent contact loads at 10%, 50% and 100% of $Q_{max}$.

Subsequently, the magnitude and distribution of the stress concentration factor through the raceway structure was found to be characterised by Equation (5.18) with regards to the angular distance $\theta$ away from the
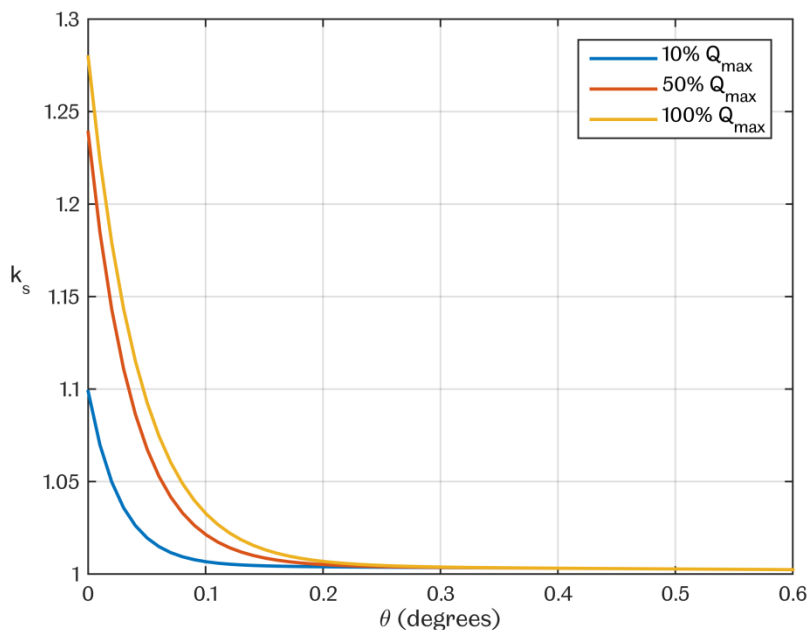


Figure 5.17: Stress concentration factor vs angular distance away from the raceway notch for normal load to the raceway-roller contact equivalent to 10%, 50% and 100% of $Q_{max}$

defect.

$$k_s(\theta, Q_u) = \beta e^{\gamma\theta} + 1.0046 e^{-3.7 \times 10^{-3}\theta}$$

<div align="right">(5.18)</div>

Where, the coefficients $\beta$ and $\gamma$ were functions of the normal load applied at the defect location $Q_u$, as defined in Equations (5.19).

$$\beta = 0.2238 e^{2.091 \times 10^{-7} Q_u} - 0.2354 e^{-5.601 \times 10^{-6} Q_u}$$

$$\gamma = -396.2 Q_u^{-0.2072}$$

<div align="right">(5.19)</div>

## 5.4.2 Integrating Stress Concentrations into the Indexing Bearing Life Model

Definition of Equations (5.18) and (5.19) allowed the effects of the proposed surface defect to be integrated within the indexing model. The defect location, relative to the bearing loaded zone, would of course move during any indexing strategy and would subsequently be subjected to varying degrees of normal load, thus altering the induced stress concentrations, as illustrated in Figure 5.17. Thus, by tracking the location of a defect throughout a given indexing strategy and considering the internal bearing load, $Q_\varphi$, at each static position, the stress concentration curve could be characterised at each such location, and applied to the indexing calculation procedure as detailed in Equations (5.10)-(5.15) as appropriate.

The calculation procedure would remain largely the same, except that in order to accommodate the stress concentrating effect of the defect it was necessary to modify some of the calculation steps, such that Equation (5.10) would become:

$$L_{10(h)\varphi}^{[n]} = a_r \left( \frac{C_{\sigma\varphi}^{[n-1]}}{k_{s\varphi} \tau_{1\varphi}^{[n]}} \right)^{\frac{10}{3}}$$

<div align="right">(5.20)</div>

And Equation (5.12) would also be modified to the form:

$$C_{\sigma\varphi}^{[n]} = k_{s\varphi} \tau_{1\varphi}^{[n]} \left( \frac{L_{10(h)\varphi r}^{[n]}}{a_r} \right)^{\frac{3}{10}} \qquad (5.21)$$

Where $k_{s\varphi}$ is the stress concentration factor of the raceway element residing at angular position $\varphi$ from bearing top-dead-centre as a result of the geometric change to the rolling surface. For most of the raceway elements around the circumference of the inner race $k_{s\varphi}$ would be equal to 1, except in the vicinity of the geometric discontinuity, where SCF's could be determined based on Equations (5.18) and (5.19).

To illustrate the effects of the above enhancements to the model, an indexing strategy is presented in Figure 5.18 which is comparable to that of Figure 5.2; that is for a strategy of 30 degree raceway indexation every 29 hours. Here however a defect of the proposed geometry is seeded from the outset of bearing operation at the point of maximum bearing load ($\varphi = 180°$).
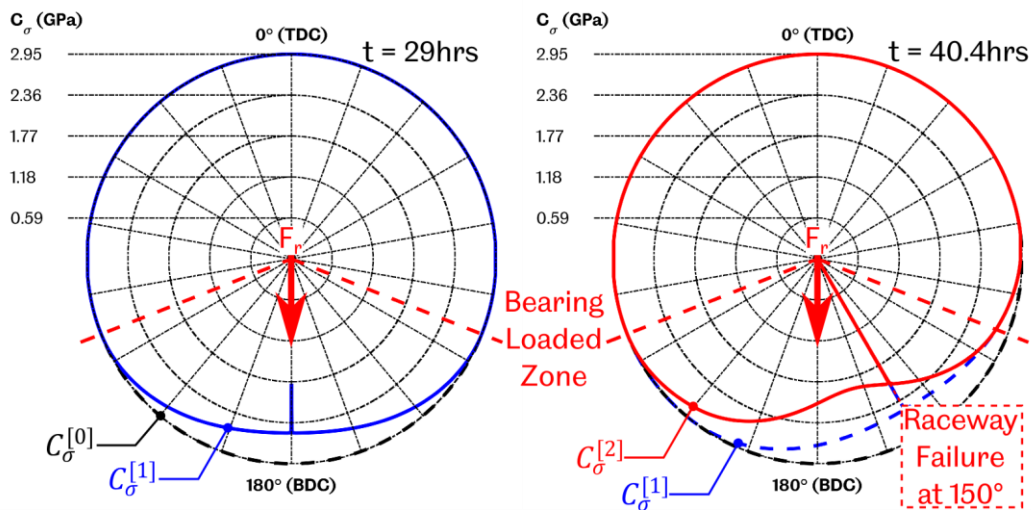


Figure 5.18: Incremental reduction in dynamic stress rating for a MultiLife bearing indexed by 30 degrees every 29 hours with a surface defect introduced at bearing bottom-dead-centre

As expected, bearing life is reduced through introduction of the surface defect, and bearing failure occurs as early as the second index period such that total $L_{10(h)}$ life here is 40.4 hours, representing 36% of the life attained with an undamaged bearing. As illustrated in the figure, the effects of the defect are considerable, resulting in rapid reduction of elemental stress ratings in the extremely localised area around the defect, and failure ultimately occurs at this location.
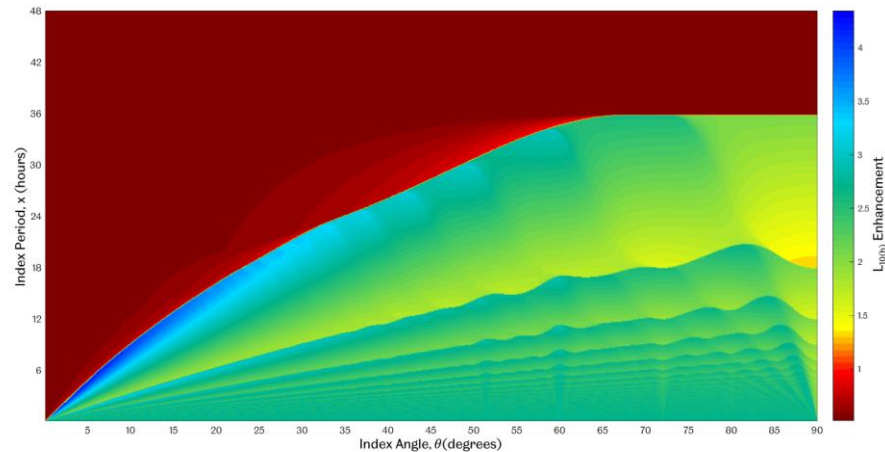
Figure 5.19: Life enhancement matrix for a bearing with a seeded defect at bottom-dead-centre, with a total radial bearing load of $1140kN$ applied

Figure 5.19 is comparable to Figure 5.3, where life enhancement is illustrated for the same range of indexing strategies but with the surface defect present at bearing BDC. What is immediately apparent through comparison of the two is that maximum bearing life enhancement is reduced in the presence of the defect to 4.3 times the standard $L_{10(h)}$ life. Furthermore, there appears to exist only a small range of indexing strategies that are capable of achieving such life enhancement, illustrated by the narrow dark blue band towards the bottom left hand side of the figure. Indeed, the region of higher bearing life is predominantly green in Figure 5.19, indicating average life enhancement of around 2-2.5 times $L_{10(h)}$. In Figure 5.3, for the case of the undamaged raceway, the plot was predominantly blue, indicating many operational areas at which a life multiplier of 4.5-5 was achievable.

The region of low bearing life is also expanded in the presence of a surface defect, where the red region extends across the entire width of the plot for $x \geq 36$ hours. This is logical, since high stresses are induced as a result of extended load durations, which subsequently increases the likelihood that material fatigue will initiate, as observed in Figure 5.18, and thus the utilisation of larger index angles between such static loading periods will not alleviate the onset of damage. It should also be noted that the red region in the presence of a raceway defect actually represents life factors below 1, illustrating the importance of moving a damaged region of the raceway away from the loaded zone as soon as possible in practice.

Figure 5.20 illustrates the progression of elemental stress distribution for four indexing cases where index period is maintained at 8 hours for each

case while index angle is increased. As observed in the figure, multiple stabilised stress bands can still be formed resulting in multiple full raceway revolutions (Plots (b)-(d)). However, achievement of this in the presence of a defect is not as effective in the extension of bearing life as was the case for an un-damaged bearing (as shown in Figure 5.4) due to the fact that higher angles of indexation ultimately result in high stress concentration regions occupying the bearing loaded zone more frequently, and thus increasing the likelihood that failure will be induced at such locations.

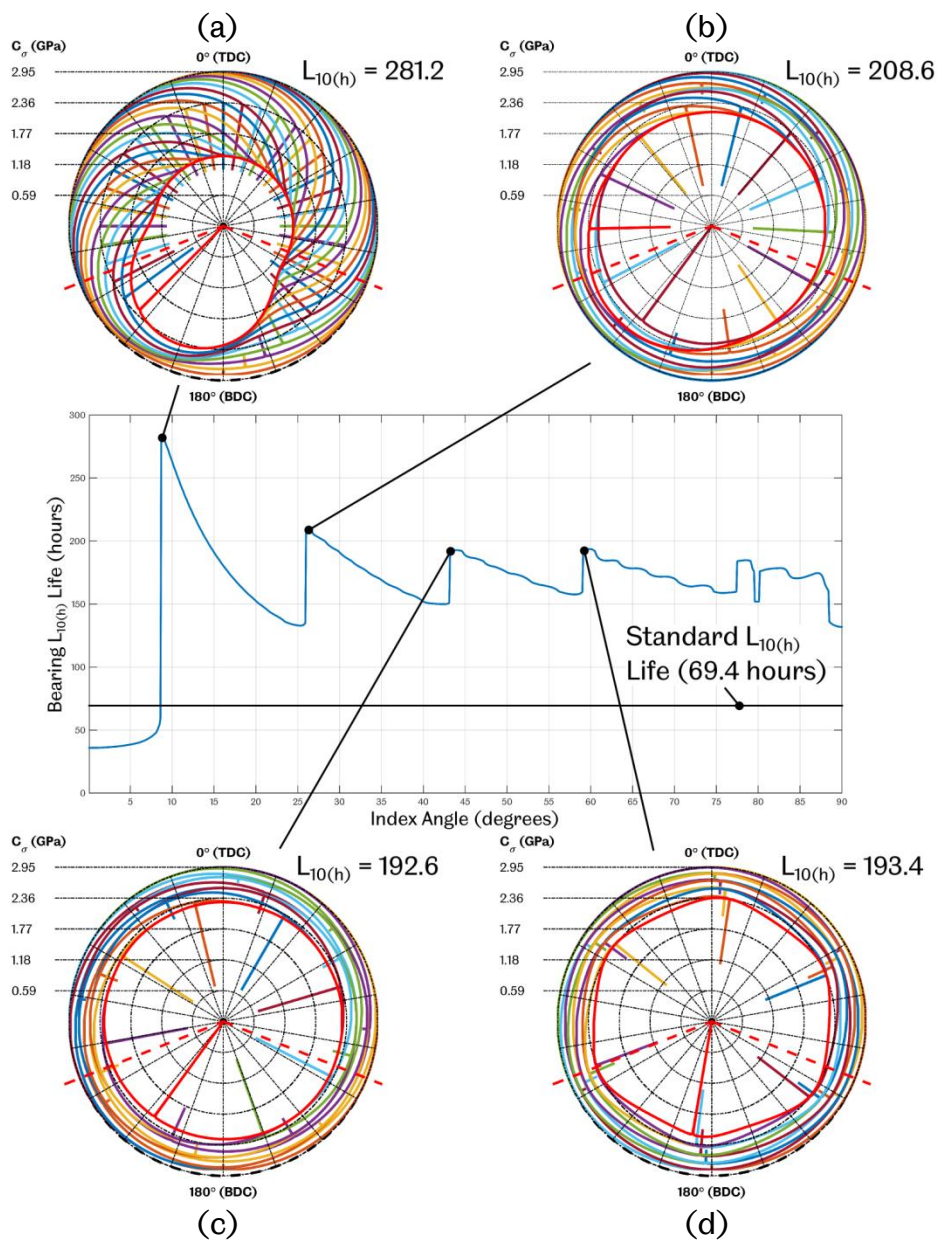

Figure 5.20: Stress rating plots for indexing strategies at a 8 hour index period and (a) $\theta = 9°$, (b) $\theta = 26.3°$, (c) $\theta = 43.5°$ and (d) $\theta = 59.6°$ with a surface defect introduced initially at bearing bottom-dead-centre

### 5.4.3 Applicability to WTGB Planetary Stage Bearing Life

In reality, the faults induced in WTGB planetary stage bearings during impact events will be *sub-surface* initiated, and the resultant stress concentrations will likely be somewhat less than those approximated in the above analysis, given the more extreme *geometric* changes that have been modelled here. It is therefore anticipated that bearing life would generally be higher for a single impact event occurring in the field than that illustrated in Figure 5.19 for a given indexing strategy – assuming that only one impact is observed throughout the bearing life.

Nonetheless, it could be concluded from Figure 5.19 that maintaining a low index period would generally appear to be advantageous for generic implementation of a MultiLife bearing system. As discussed in §1.2.2 however, the occurrence of fault seeding impact events may be numerous in reality, taking place randomly throughout the operational life of a bearing. Considering this, it would be reasonable to conclude that moderate angles of indexation employed relatively frequently may in fact be the best operational approach to ensure that optimum life is achieved. Ultimately however, it is clear that robust condition monitoring techniques would enhance MultiLife operation considerably. Indeed, if bearing faults could be both detected and located within a planetary bearing, such information could be used to determine the timing and degree by which raceway indexation should be performed in real-time so as to maximise the remaining useful life of the bearing material.

## 5.5 Considerations for Bearing Life Testing

In gaining further knowledge of the effects that surface defects may have on bearing life through the above analysis, it was subsequently possible to fully define the full scale bearing life tests, as introduced in Chapter 4. Overall aims of the testing were to provide proof-of-concept for the MultiLife principle, whilst also determining the effectiveness of ultrasound in detecting bearing faults. To simplify instrumentation (as discussed in §4.3.3) ultrasound sensors would be permanently bonded to bearing raceways at one single location relating to the point of maximum bearing load. Therefore, it was logical to seed a surface defect at this location so that failures would be more likely to propagate from this point (given the evidence indicated in Figure 5.17), and thus the likelihood of detection by the ultrasonic element would be high.

As also discussed in the previous chapter, the main drawback of utilising a permanently bonded ultrasonic element is that this would move with the raceway as it is indexed, and in the case of the test platform developed for this study, only a single instrumentation channel exists (as illustrated in Figure 4.6), thus the element would be removed in the event of such rotation. However, the MultiLife principle could still be tested without the need for raceway indexation, through seeding of surface faults in multiple locations. Based on the above analysis, those which reside within the loaded zone would be expected to propagate at a rate proportional to the normal internal bearing load at the defect location and, by contrast, those which reside outside the bearing loaded zone should experience no propagation whatsoever. This therefore still satisfied the key aims of the testing program since firstly, the ultrasonic source element could still be utilised to observe the progression of damage from a surface fault residing at the point of maximum bearing load. Secondly, the MultiLife principle could be tested and validated through indication that defects further away from the maximum internal bearing load would observe a lower rate of propagation, and indeed those outside the loaded zone should display no signs of propagation whatsoever.

It was therefore considered that three defects would be sufficient to achieve this, and it was thus proposed that such defects should be positioned on the rolling surface at the locations $\varphi = 148, 180$ and $356°$. These angles were selected specifically so that – based on the distribution of rolling elements within the bearing – no two surface defects would be impinge upon simultaneously by rolling elements during rotation. While this was largely trivial with respect to the acquisition of ultrasonic data at the maximum loaded point, such impingements could potentially be detected simultaneously by acoustic emissions sensors. Thus it was beneficial to phase such acoustic sources appropriately so as to avoid the potential overlapping of acquired bursts, and thus make it easier to link specific data features with the associated surface defects.

Figure 5.21 illustrates the stress distribution for a test bearing inner raceway run under the proposed operating conditions (as used consistently throughout this chapter) and with the three surface defects introduced into the rolling surface. From this, the 10% failure rate for the test bearings was estimated at 35.9 hours, resulting from a raceway failure in the region of **Defect 2**; located initially at bearing bottom-dead-centre, as expected. A stress concentration is also observed around **Defect 1**, which is much smaller in magnitude due to the fact that the normal load in

this region is equivalent to $0.74Q_{max}$, and thus elemental fatigue here will propagate at a much slower rate as expected.
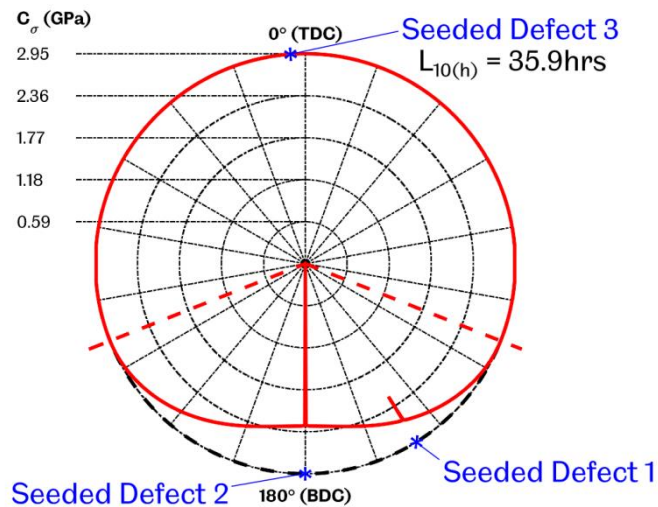


Figure 5.21: Raceway damage induced with three surface defects leading to bearing failure after 35.9 hours

## 5.6 Conclusions

- An elemental bearing life model has been developed to establish the bearing life that may be attainable when a MultiLife Indexing Mechanism is utilised.
- The basic model was based upon the standard life rating principles and considers the progression of failure via rolling contact fatigue.
- The model was applied to a range of indexing strategies for an NU2244 test bearing run on the experimental test rig under the proposed test conditions as defined in Chapter 4. Up to a five-fold increase in the standard $L_{10}$ bearing life was identified through use of the MultiLife mechanism and, generally, low angles of indexation applied at relatively short intervals appeared to be optimum.
- The model was further enhanced to consider the effects of transient loading events that arise in WTGB planetary bearings acting to advance rolling contact fatigue failure in-service, thus reducing bearing life. This was considered in terms of the stress concentrations that are induced from such events.
- Specifically, the stress concentrating effects of a geometric surface defect were considered, given that this would be the method by which stress raisers would be induced within test bearings for the

experiments within this study. Such effects were investigated through the development of a finite element model and comparisons of associated inner raceway stress fields with and without a surface defect present.

- Bearing life was found to be significantly reduced in the presence of a surface defect, and the ability of the MultiLife Mechanism to enhance bearing life was found to be limited to a much smaller range of indexing strategies.
- Based on the model results, the bearing tests established in Chapter 4 were further defined in terms of the specific seeded defects that would be investigated, and a prediction was made to the expected bearing life in such cases.

# Chapter 6

# CHARACTERISING BEARING OPERATION FROM ULTRASONIC MEASUREMENTS

This chapter provides an initial discussion of ultrasonic measurements taken on the test platform relating to more fundamental characteristics of the obtained signals, rather than exploring the general trends throughout a bearing life test, as will be discussed in Chapter 7. Various analyses are performed in an attempt to provide further explanation of certain features that are often observed when applying the ultrasound technique to bearing applications, but which are not well understood. Following this, and based upon some of the earlier findings within the chapter, the features to be extracted during the full bearing life testing are defined, and a prediction is made to the expected change in such features as bearing degradation is observed.

## 6.1  General Analysis of a Single Roller Pass

Bearing tests would typically be initiated through incremental application of applied load up to the required test set point, and it was therefore possible to obtain ultrasonic measurement across a range of load cases during such instances. This enabled various features of the ultrasonic measurements to be characterised as dependant on – or independent of – applied load.

Figure 6.1 presents reflection coefficient curves for a single roller pass at $100 rpm$ for bearing loads within the range $100 - 1000 kN$ at $100 kN$

increments. Each data set within the plot was recorded at a $10kHz$ pulse repetition rate, and reflection coefficient has been established at the $10MHz$ resonant frequency of the ultrasonic source element. The reference signal was obtained for each load case using the modal reference approach as described in §3.4.3.1.
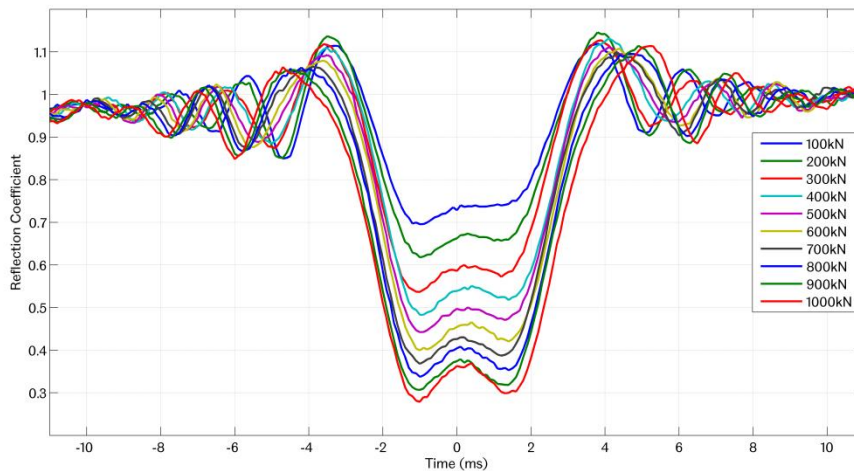


Figure 6.1: Reflection coefficient curves for a single roller pass at bearing loads between 100 and $1000kN$

In the figure, the contact between inner raceway and roller is clearly observed by way of a general reduction in reflection coefficient at $-2 \leq t \leq 2ms$. All curves have been centralised on the figure to the temporal mid-point of the roller passage; i.e. the centre of the contact patch, where the rolling element is directly aligned with the centre axis of the ultrasonic element. A clear relationship can be observed between reflection coefficient and load, where reflection coefficient is increasingly minimised across each roller passage as applied load is increased. A number of additional features are also observed from the data, as highlighted in Figure 6.2(a) and introduced within the following paragraphs.

Firstly, no single minimum can be observed in any of the reflection coefficient curves that could logically be associated with minimum film thickness within the contact. In fact, a "W-shaped" profile is apparent such that reflection coefficient tends to rise at the centre of the roller pass, where it would typically be expected that ultrasonic reflection would be minimised due to peak contact pressures and the subsequent constriction of the oil film. This profile appears to become more pronounced with increased load, and indeed by plotting the difference between the central peak and minimum trough for each curve (Feature C, Figure 6.2(a)) with respect to load, as shown in Figure 6.2(c), a generally

130

linear trend can be observed. This clearly suggests the observed feature to be a **load dependant** effect.

The width of the two sequential minima on each curve (Feature A, Figure 6.2(a)) appears to be similar for all load cases, occurring across a duration of around $2.4ms$, which can therefore be said to be **load independent**. Considering the total width of the W profile however (Feature B, Figure 6.2(a)), it is evident that this again becomes **load dependent**, as illustrated in Figure 6.2(b).

Finally, signal "flutter" can be observed either side of the roller passage in all load cases, where reflection coefficient tends to fluctuate around 1 prior and subsequent to a roller passage (Feature D, Figure 6.2(a)). This suggests reflected energy to be higher at brief instances than that of the steel-air reference condition, where all receivable energy emitted by the source element (i.e. that within the measurable area, as established in §4.3.3.1) should theoretically be received back. Generally, the same number of reverberations can be observed for all load cases. In addition to this, while the relative phase of such reverberations appears to change as a function of load due to the widening of the W profile, all appear to follow the same rate of attack and decay at contact entry and exit respectively, as shown in Figure 6.3. Clearly then, this is a **load independent** phenomenon.
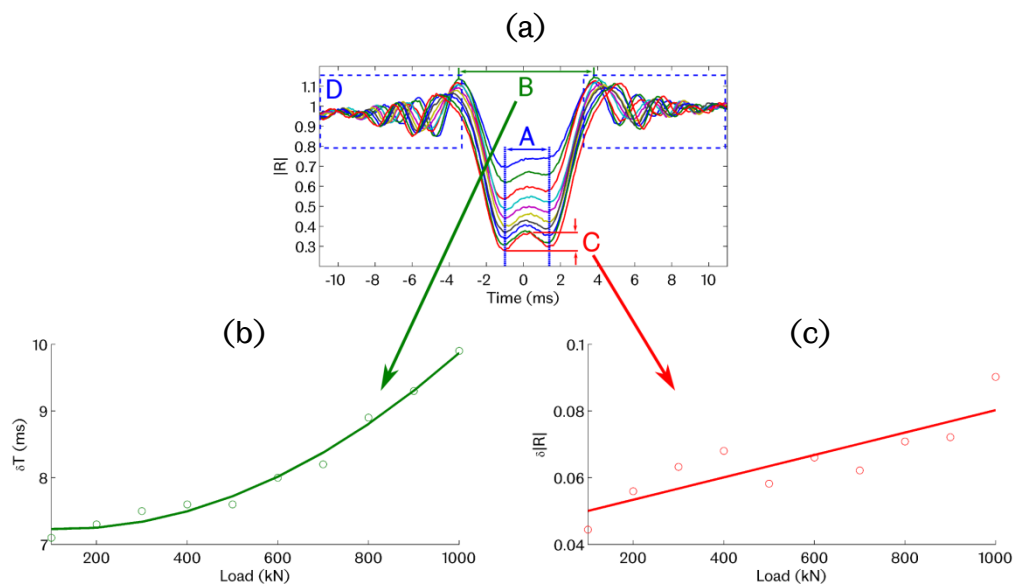


Figure 6.2: (a) Features of interest in Reflection Coefficient curves of Figure 6.1; (b) Correlation between width of contact region (Feature B) vs load; (c) Correlation between magnitude of central rise in reflection (Feature C) vs load
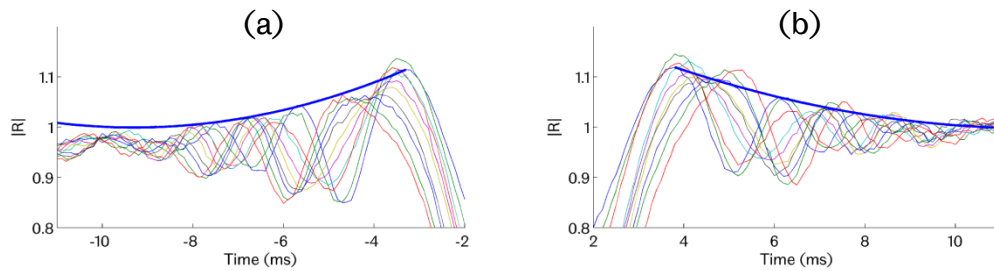
Figure 6.3: Illustration of the broadly similar rate of (a) attack and (b) decay in signal flutter across all load cases at the contact entry and exit regions (Feature D, Figure 6.2(a))

## 6.2 Fringe Effects at Contact Entry and Exit

The first point of discussion focusses on the flutter in reflection coefficient that is observed at the contact entry and exit, as highlighted in Figure 6.3. As discussed above, this is unusual not only due to the apparent instability in ultrasonic reflection around the contact, but also due to the fact that instances exist where reflection coefficient rises above 1; suggesting more energy to arrive back to the ultrasonic source element than that which should feasibly be available to it, i.e. that emitted within the measurable segment of the incident beam. In addition to this, such effects lead to difficulty in establishing the absolute boundaries of the contact under observation, which – if determined – could provide direct indication of contact width and subsequent determination of applied load.

For simplicity, the following analysis deals exclusively with the $1000kN$ load case from Figure 6.1. While reflection coefficient data presented in the figure was obtained through analysis of the first reflection from the inner raceway rolling surface, subsequent reverberations through the raceway were also recorded. In order to observe all such reverberation pulses from the roller-raceway interface simultaneously an intensity plot was constructed. To improve clarity, each individual waveform was clipped to reduce the effects of extreme maxima and minima, which would otherwise cause reduced detail of the mid-range amplitudes of key interest.

The negative component of signal amplitude was firstly removed from analysis, since any patterns emerging within reflected pulses were known to occur similarly in both the positive and negative wave directions. This

was done primarily to reduce the range of values that would need to be represented by the colour map, thus increasing resolution of the remaining data; hence the first expression in Equation (6.1) was derived. In a similar manner, the range of positive values to be represented was also reduced. Here, considering the magnitude of the fringe patterns highlighted in Figure 6.3, it was expected that such effects could only be induced by signals of comparatively low amplitude. Furthermore, large scale changes to peak signal amplitude were only expected to occur within the contact region, and thus would only act to widen the required range of values to be represented by the colour map, if included in the analysis of the contact entry and exit regions. In order to avoid this, whilst still maintaining detail within the mid-range amplitudes, each raceway reflection pulse was clipped to 1/3$^{rd}$ of its peak reference amplitude. To apply this, an exponential curve was utilised, as shown in Figure 6.4(a), obtained from the modal reference waveform of the data stream, which was found to be represented by expression $\boldsymbol{B}$ as defined in Equation (6.2).

$$A(t) \leq \boldsymbol{0} = \boldsymbol{0} \ \& \ A(t) \geq \boldsymbol{B}(t) = \boldsymbol{B}(t) \tag{6.1}$$

Where:

$$\boldsymbol{B}(t) = 509.7e^{-\frac{t}{1024.4}} \tag{6.2}$$

This threshold was applied across each recorded waveform to clip any values which exceeded it, and was then used again to normalise the data such that effects of signal attenuation were compensated for across the entire data capture window $(21\mu s)$.
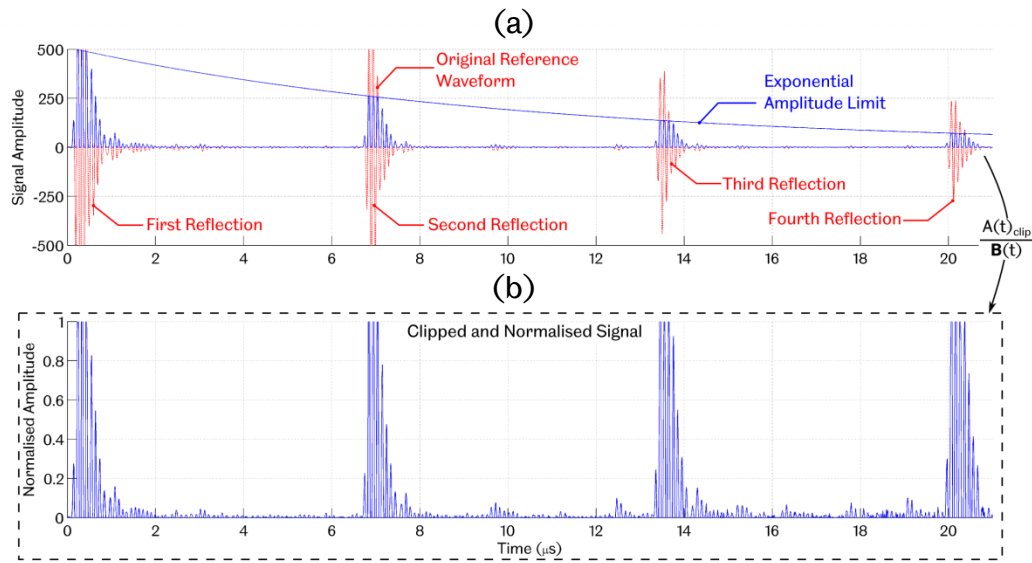
Figure 6.4: (a) Clipping of original ultrasonic reference signal for the $1000kN$ load case (b) Resultant waveform after normalisation
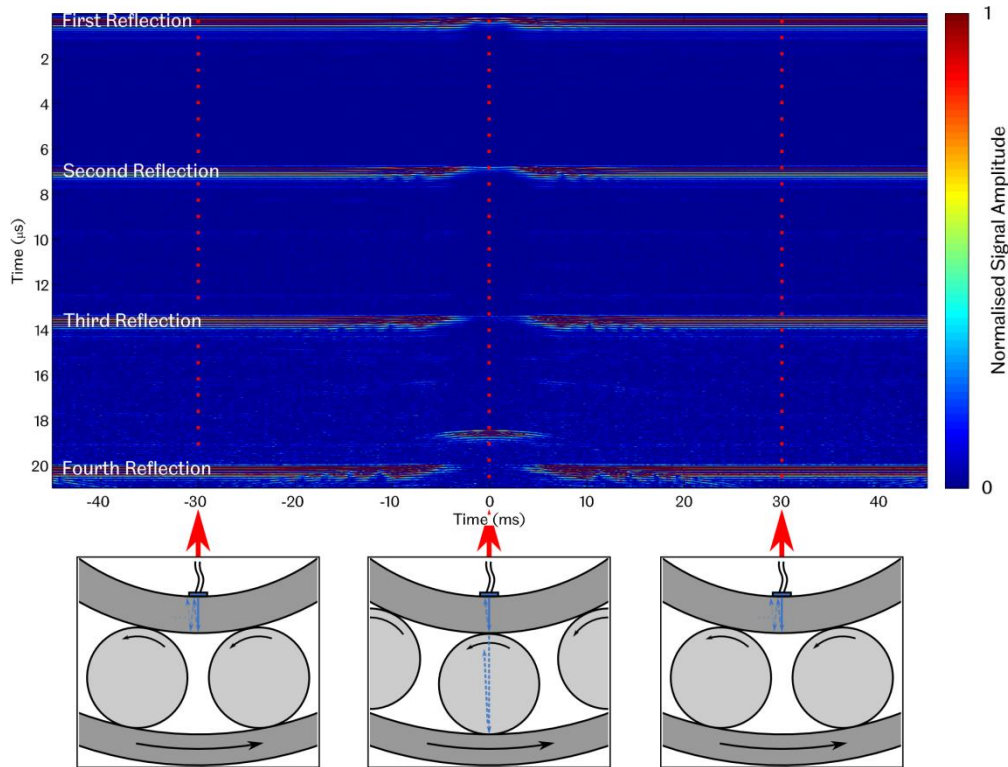


Figure 6.5: Colour plot to illustrate variations in signal amplitude throughout a single roller pass at $1000kN$

134

Figure 6.5 shows the resultant intensity plot, where the normalised waveforms, such as that shown in Figure 6.4(b), are orientated vertically and arranged side-by-side along the $x$ axis to illustrate variation throughout the roller passage. Here it can be observed that all four reflections are generally stable in the time between a roller pass, as observed by red horizontal banding between $-45 \geq t_x \geq -30$ and $30 \leq t_x \leq 45$, indicating maximum signal amplitude ($A = 1$). At $t_x \approx \pm 25ms$ a saw tooth pattern begins to emerge within the fourth reflection pulse, which then cascades through to the third, second and first reflections, becoming visible in the first reflection pulse at $t_x \approx \pm 10ms$. The roller-raceway interface becomes observable in reflected pulses between $-5 \leq t_x \leq 5$, illustrated by large scale amplitude drop ($A \leq 1$) as ultrasonic energy is transmitted through the interface, consequently resulting in reduced reflection back to the source element. In the third and fourth reflection pulses transmission is such that reflected signal amplitude drops below the level of background noise, causing red banding to become entirely blue ($A = 0$). Within this region, additional reflected energy becomes observable at $t_y \approx 18.5\mu s$, which is a result of energy being transmitted through the roller and reflected from its opposite face. The ability to measure this may be particularly useful, and will be discussed further in §6.4.2.

## 6.2.1 Higher Order Reverberations

Figure 6.6 provides further detail of the patterns in both the first and fourth reflection pulses of Figure 6.5. Also provided here are "Spectrograms" of reflection coefficient vs frequency for each of the time domain waveforms represented in the intensity plot below. Frequencies within the range of $6 - 14MHz$ are presented, which approximately represents a $-24dB$ bandwidth of the reference spectrum.

Fringe patterns manifest themselves differently in the time and frequency domain, yet both are clearly linked to the same physical effect since the duration over which they occur is the same in each case. In the frequency domain, fringe patterns are observed as sequential light and dark striations, oriented vertically either side of the roller passage, present across the full frequency range but diverging at lower frequency, illustrating a shift in phase away from the contact. In the time domain however, and when viewed in closer detail (as in Figure 6.6), such patterns occur as circular ripples leading in towards the centre of the contact, which seems to be a result of additional, low amplitude, ultrasonic waves moving in and overlapping the major reflected pulse.
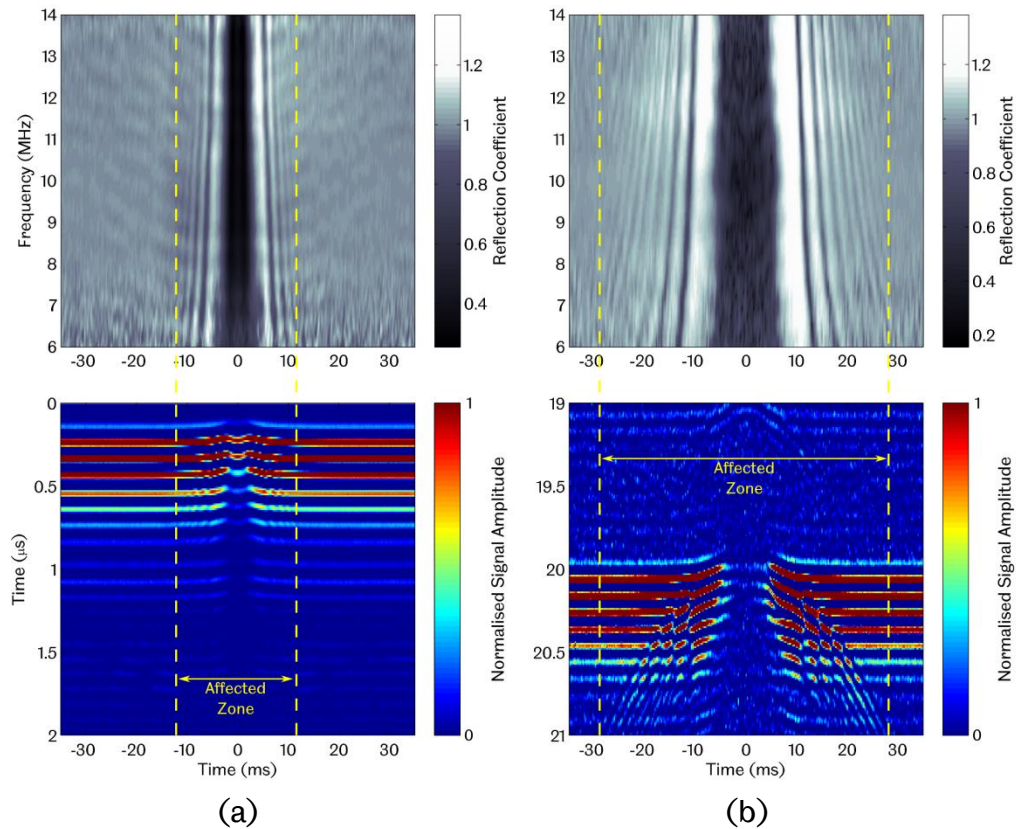
Figure 6.6: Closer inspection of (a) first reflection and (b) fourth reflection signals displayed as both frequency "Spectrograms" (top) and time domain intensity maps (bottom)

Fringe effects become much clearer in higher order reverberations through the inner race, where the patterns are not only more defined, but also occur over a much longer time duration. Indeed, by the fourth reflection, as observed in Figure 6.6(b), fringe effects become observable over a duration of approximately $25 - 30ms$ either side of the rolling contact, significantly longer than the first reflection (Figure 6.6(a)) where the affected zone is much narrower; bound within a $\pm 12ms$ time frame. If the frequency maps of Figure 6.6 are plotted at a single frequency this becomes even more apparent, as shown in Figure 6.7, where reflection coefficient is shown at the resonant frequency ($10MHz$) only. This further highlights the increased magnitude by which fringe effects occur upon the fourth reflection measurements by comparison to that of the first reflection case.
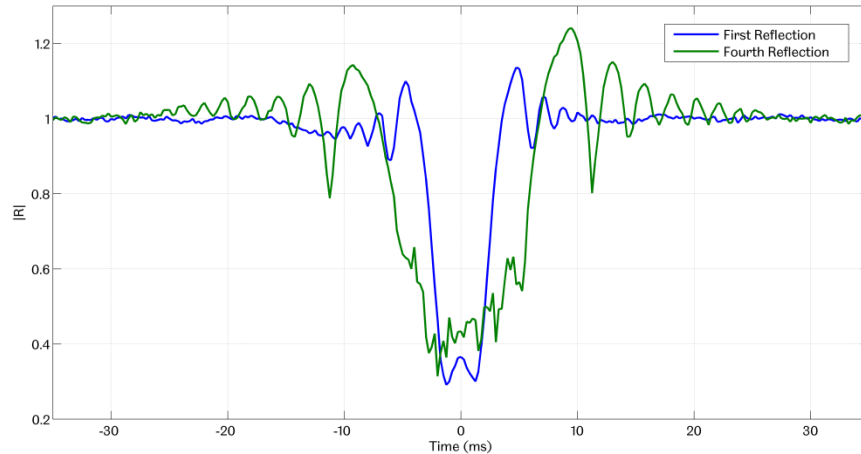
Figure 6.7: Comparison of reflection coefficient at $10MHz$ for the first and fourth reflection data of Figure 6.6

Clearly then, interference effects have a much greater influence when ultrasonic residence times within the raceway material are longer. This is logically explained by the divergence of ultrasonic energy known to occur through the material as established in §4.3.3.1, yet consideration in the aforementioned section was only made to the divergence of *incident* ultrasonic energy emitted from the source element. Considering, therefore, the implications of such divergence when the sound wave begins to reverberate through the raceway material, it is clear that ultrasonic energy will rapidly spread across a significant segment of the rolling surface as illustrated in Figure 6.8. Consequently, as a roller approaches the line-of-sight of the sensor, small amounts of ultrasonic energy within the periphery of the emitted beam would be reflected back from the roller-raceway interface and consequently induce responses at the source element. As such, effects are induced upon reflected signals before the contact is within the field of view of the ultrasonic element.

As discussed above, the intensity of fringe patterns increases over higher order reverberations. This is most clearly observed in the $10MHz$ reflection curves of Figure 6.7, where spectral reflection coefficient peaks at ~1.25, significantly higher than the first reflection curve which peaks at ~1.1. Evidently, divergence of energy will increase significantly with further reverberations of the incident beam and more energy consequently becomes available across the peripheral roller-raceway interface. As such, higher intensity waves will consequently be reflected back to the source element. This can also be partly attributed to the fact that signal amplitude of higher order reverberations is generally much lower than earlier reflections due to attenuation, and thus additional

energy from the peripheral interface has a greater influence upon the observed variation in such higher order reflected pulses.
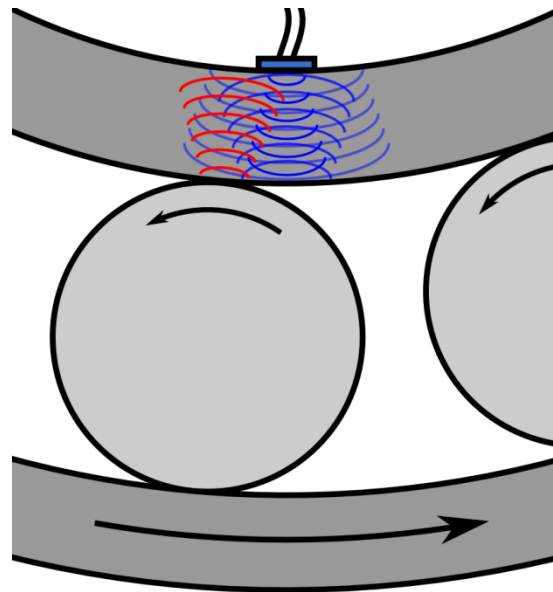


Figure 6.8: Illustrating the spread of acoustic energy across the raceway with multiple reverberations of an ultrasonic pulse, and the subsequent reflections from the nearby roller-raceway interface

Similarly to the emitted ultrasonic energy, reflected energy from the peripheral roller-raceway interface will also reverberate through the raceway material and this energy will consequently diverge as well. This may explain the general temporal widening of fringe patterns at increased residence times. As illustrated in Figure 6.9, the ultrasonic wave-front will broaden with subsequent arrivals at the inner raceway bore, and thus the leading region of the wave-front occupies a much wider area. This consequently results in detection by the ultrasound element further ahead of the contact when observing higher order reverberations.
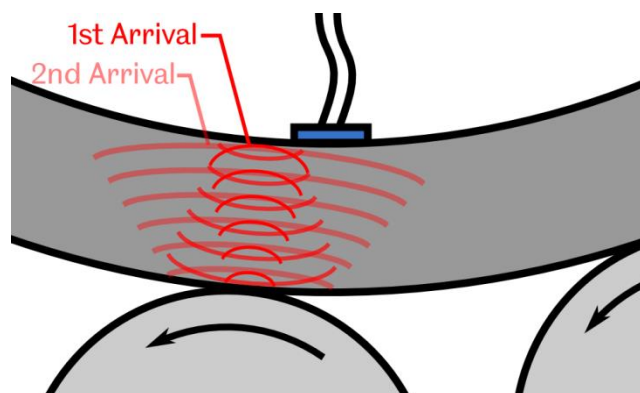


Figure 6.9: Illustration of the broadening of the ultrasonic wave-front from the peripheral interface with multiple reverberations and effects on the arrival at the source element

This therefore explains the occurrence of the fringe patterns highlighted in Figure 6.3 and the reasons for the earlier detection of such patterns when observing higher order raceway reflections, as identified in Figure 6.5.

## 6.3 Reflection Coefficient vs Load during Rolling Contact

Relationships between reflection coefficient and load would logically be determined by assessing the minimum point on the reflection coefficient curve for a given roller pass, since this should represent the central film thickness formed at the interface. For a line contact, film thickness will be generally uniform across the contact length and is understood to remain largely parallel across its width (in the direction of rolling), particularly under increased normal loading, as shown in Figure 6.10. A parallel film is maintained until the fluid exit point, where a constriction in film thickness must occur to enable pressure to drop back to that of the near ambient levels outside the contact [29]. In general, therefore, we should expect the reflection coefficient curve for a roller pass to broadly match the form of the theoretical film thickness profile at any given load.
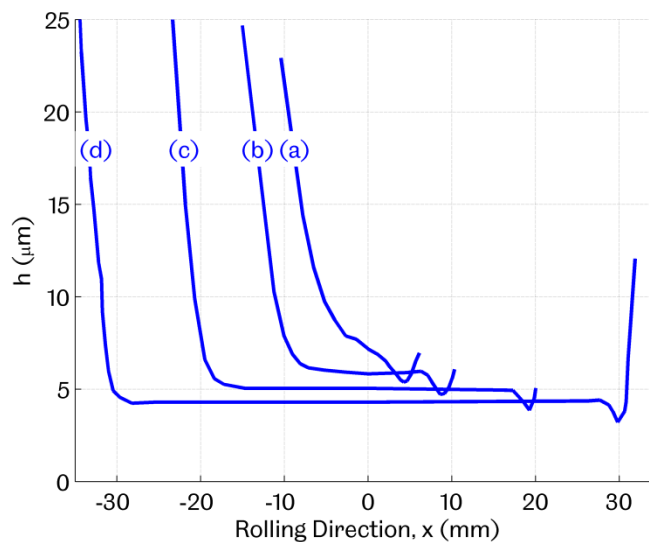


Figure 6.10: Film thickness profiles for a line contact with increasing contact pressure from (a) to (d) in the ratio 1:2:4:6 (adapted from [97])

Clearly however, as observed from Figure 6.1, this is not the case, and measurement curves present a different form entirely with two sequential minima, both generally similar in magnitude, but with the leftmost (fluid

inlet side) being slightly more prominent in each case. At all load cases both minima are found to have the same temporal separation, indicating that the observed width is not dictated by load. However, a rise in reflection coefficient is observed at the centre of the roller pass – in between the two minima – which can again be seen in all load cases, but becomes increasingly prominent at higher load. It appears therefore that load on the contact induces additional changes in ultrasonic reflection, likely a result of local mechanical changes in the bearing material adjacent to the contact. Consequently, the known changes in ultrasonic reflection pertaining to lubricant film formation and induced contact pressure – as dictated by the spring model predictions – become masked, leading to misleading results when applying such models directly to the obtained results.

## 6.3.1 Comparisons between Obtained Results and Spring-Model Predictions

The central film thickness, $\bar{h}$, for an EHL line contact can be established from Equation (6.3) [29].

$$\frac{\bar{h}}{R} = 3.11\{2\alpha E^*\}^{0.56} \left\{ \frac{\bar{U}\eta_0}{2E^*R^*} \right\}^{0.69} \left\{ \frac{Q}{2LE^*R^*} \right\}^{-0.1} \tag{6.3}$$

This was established for the test bearing lubricant (Castrol Hyspin VG32), based on the estimated dynamic viscosity, $\eta_0$, at an operating temperature of $65°C$, as was the case for the tests in question. $\alpha$ is a pressure-viscosity coefficient, taken to be $2.076 \times 10^{-8} m^2/N$ for the test oil, and $\bar{U}$ is the entrainment velocity into the contact, and is thus based on the bearing rotational speed $N_r$. Given the sensor location at bearing BDC, maximum internal bearing load would be observed, hence – considering the range of bearing loads under consideration ($100 - 1000kN$) – normal contact load would range from $33kN$ to $333kN$ (according to Equation (2.1)).

Subsequently, theoretical film thickness was found to range from $153.9nm$ at $100kN$ to $122.3nm$ at $1000kN$. This, in turn, would be expected to yield reflection coefficient values within the range $0.37 - 0.34$ from ultrasonic measurement data, based on application of the spring model for a lubricated contact (Equation (3.16)). The above calculations indicate that despite the range of applied loads being significantly large, this would in fact only command relatively small changes to film thickness within the

contact. As such, reflection coefficient curves would be expected to deviate within the contact region by only 0.03 across the entire load range. Despite the absence of the expected minima in the measurement data shown in Figure 6.1, it is clear that there is a much greater step change between each load case than is predicted by the spring model.

Referring to the first of the spring model assumptions, as outlined in §3.3.3 however; it is required that the lubricant film acts to completely separate the two bearing surfaces. It is therefore relevant to consider the Lamda ratio for the contact scenario, as defined in Equation (3.17). Roughness measurements were performed on an NU2244 test bearing for both the inner raceway and rolling element surfaces and were found to be $0.09\mu m$ and $0.17\mu m$ for both components respectively, giving a total composite roughness for the interface of $0.19\mu m$. Consequently, Lamda ratios were found to range from $0.79 - 0.63$ across the $100 - 1000kN$ load range. As such, it is evident that at all load cases investigated, oil film formation was insufficient to separate the two bearing components, and thus the spring model may in fact not be applicable without also considering the asperity interactions that will likely dominate the contact.

Figure 6.11 plots theoretical reflection coefficient values across the range of test loads according to the spring model for a lubricated contact. In addition to this however, consideration has also been made to the variations of the model for both mixed and solid contact, as discussed in §3.3.4, to assess whether either of the latter represent the measurement
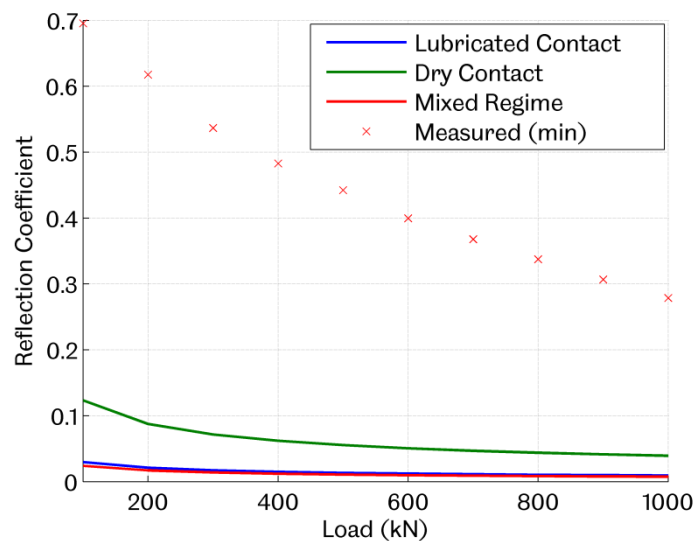


Figure 6.11: Comparison of spring model predictions of ultrasonic reflection (at $10MHz$) for a dry, mixed and lubricated interface with test measurements overlaid

data more appropriately than that of the oil film model. To facilitate the comparison, the minimum reflection coefficient recorded at each test load taken from the measurement data presented in Figure 6.1 is also overlaid on the plot.

It should be noted that, in order to apply the dry contact model, knowledge of the relationship between interfacial stiffness and contact pressure for the material couple was required. Since no calibration experiment had previously been performed on the bearing steel in question a calibration curve for a typical steel-on-steel contact was applied, taken from [98], where the following relationship was derived:

$$K = \frac{p_{nom}}{53.5}$$

(6.4)

However, since this relates to a mild-steel material couple, $K$ would in fact be significantly larger for the bearing steel in this case. This is not only a result of the increased hardness of the bearing steel, but also due to the lower initial surface roughness of the surfaces, thus resulting in higher real area of contact due to increased asperity interactions, consequently inducing higher effective stiffness at the interface. It should therefore be considered that relevant curves presented in Figure 6.11 will in fact be an over prediction of the real case, but should be acceptable for a general comparison.

In application of the lubricated and mixed regime models, bulk modulus for the lubricant was required, which can easily be determined from the speed of sound, $c$, and fluid density, $\rho$, under ambient conditions according to Equation (3.12). In EHL scenarios however, contact pressures are extremely high, and bulk modulus can change significantly. In an attempt to compensate for such changes, a bulk modulus vs pressure relationship was used, as described in [99], to allow the fluid film to be modelled more appropriately within the relevant spring model predictions.

$$B = \left\{ 1 - \frac{1}{1 + B_0'} \ln \left[ 1 + \frac{p}{B_0}(1 + B_0') \right] \right\} [B_0 + p(1 + B_0')]$$

(6.5)

Here, $B_0$ is the bulk modulus of the fluid at ambient pressure and $B_0'$ is the pressure rate of change of $B$, which is stated to be approximately equal to

11. $p$ is the fluid pressure, which was taken to be equal to the average Hertzian contact pressure for the interface at the given bearing load.

Ambient bulk modulus, $B_0$, was determined through use of Equation (3.12) for the test bearing lubricant, where speed of sound through the fluid was determined experimentally for temperatures within the range $25 - 80°C$. The curve of Figure 6.12 was obtained, leading to definition of a temperature vs speed of sound relationship for the test lubricant from the least squares best fit of the data.
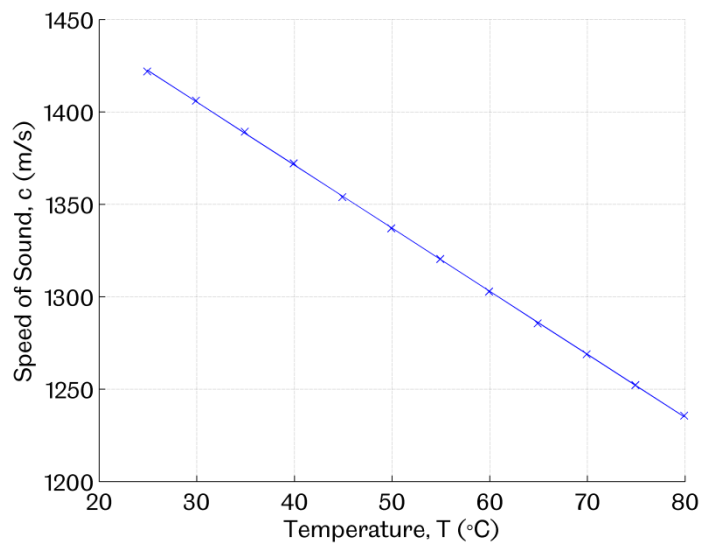


Figure 6.12: Temperature vs Speed of Sound relationship for the Castrol Hyspin VG32 test bearing lubricant

$$c = 1507.65 - 3.412T$$

(6.6)

Thus, speed of sound through the lubricant at the test temperature of $65°C$ could be approximated as $1285.9 m/s$.

Oil density at the same test temperature was calculated, according to Equation (6.7)[23], to give $B_0 = 1.39 GPa$.

$$\rho = \rho_0 \big(1 - 0.00063(T - T_0)\big)$$

(6.7)

Where $\rho_0$ is the reference density of the lubricant, typically defined in manufacturer data sheets at a given reference temperature $T_0$, stated to be equal to $870 kg/m^3$ at $15°C$ for the VG32 test oil.

Referring back to Figure 6.11, it is clear that measured reflection coefficients do not match any theoretical prediction, where experimental results are significantly higher across the entire load range and experience a much greater rate of change with load. As can be seen, both the lubricated and mixed regime models are almost identical and relatively flat across the entire load range. The dry contact model predicts slightly higher values of reflection than the other two and has a slightly steeper gradient across the load range. Given the use of a contact stiffness relationship for a mild-steel material couple (as discussed above), true values would likely be lower than presented, thus providing further contrast to the measured case. All model predictions exhibit an increased rate of change at low load; below around $200 kN$, and this is particularly pronounced in the dry contact model.

As discussed in §3.3.4, it was reported in [80] that the actual contribution of asperity contact to total composite stiffness $K_T$ was low. This is supported by the results in Figure 6.11, where the lubricated and mixed regime curves are comparable, suggesting minimal contribution of the solid stiffness component to the overall reflection coefficient results.

Applicability of the spring model to the measurement data was explored further through direct calculation of film thickness via Equation (3.16). Film thickness calculations were performed on ultrasonic reflection waveforms obtained at $t = 0$ in Figure 6.1 across the full range of test load cases. Figure 6.13(a) shows the frequency spectrum of the modal reference waveform for each load case in terms of its relative decibel amplitude, and illustrates the $-6 dB$ bandwidth of the acquired signal. Figure 6.13(b) shows the reflection coefficient vs frequency relationship for each of the measurement pulses at $t = 0$ (Figure 6.1), and Figure 6.13(c) shows the corresponding film thickness results obtained from each of the reflection coefficient curves.
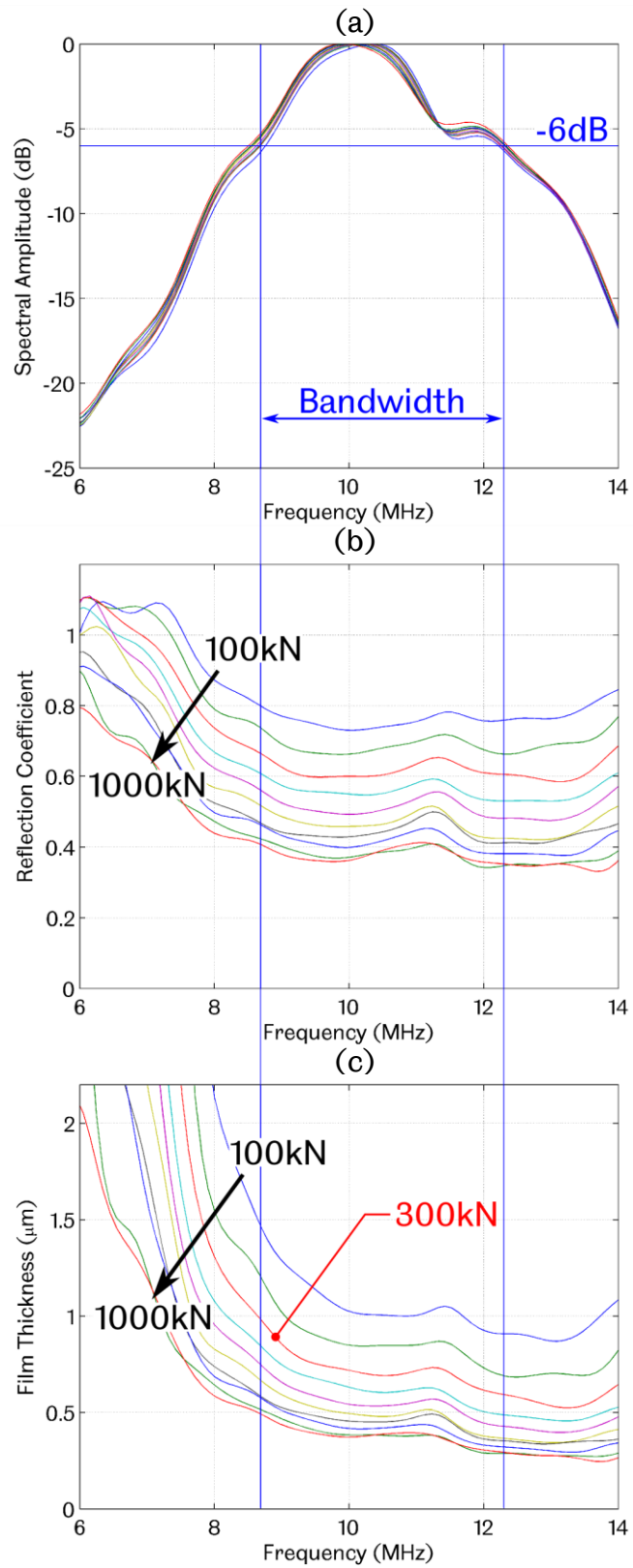
Figure 6.13: (a) Ultrasonic reference spectra for all load cases presented in Figure 6.1; (b) corresponding reflection coefficient and (c) film thickness vs frequency at $t = 0$

While we would expect to observe frequency variation across the reflection coefficient curves, as dictated by Equation (3.11), film thickness clearly cannot exhibit such variation and must be constant for all frequencies within the bandwidth of the signal. Figure 6.13(c) shows that, in fact, film thickness is highly variable across the frequency range at low load, but profiles tend to level as load is increased, suggesting that the model becomes more applicable at higher load.

A potential explanation for this may be sought through consideration of the second of the spring model assumptions as detailed in §3.3.3; requiring that the measurement surfaces be parallel across the ultrasonic wavefront. Whilst the bearing surfaces themselves will be inherently non-parallel, Hertzian theory dictates that – when in contact – surface deformation will occur such that the surfaces become parallel across the contact width, thus satisfying the spring model assumption. Clearly though, contact width is highly dependent on the normal load applied to the contact.

Film thickness profiles similar to that of Figure 6.13(c) were obtained in [80], where commercial ultrasonic focussing transducers were utilised to obtain contact measurements in a ball-on-flat scenario. The use of commercial transducers in this study meant that ultrasonic frequency could be linked directly to the focused spot size of the ultrasound beam based on the relationship:

$$d_f = 1.025 \frac{Fc}{fD}$$

<div align="right">(6.8)</div>

Where $d_f$ is the focussed diameter of the ultrasound beam, $F$ is the focal length, $D$ is the transducer diameter, and $f$ and $c$ are the wave frequency and speed of sound through the transmission medium. The film thickness profile was found to become highly sloped (with a consistently negative gradient, such as that observed in Figure 6.13(c)) at low frequencies, where the focal spot size was larger than the contact width.

Comparing the ultrasonic beam width, as determined in §4.3.3.1, with the Hertzian contact width approximation for the test bearing geometry at the range of loads investigated, it is observed that beam width in this case exceeds the theoretical contact width at loads below $300kN$, as shown in Figure 6.14. This, alongside the comments made in [80], supports the observations in Figure 6.13(c) where film thickness profiles become much flatter for loads exceeding this $300kN$ contact width threshold.
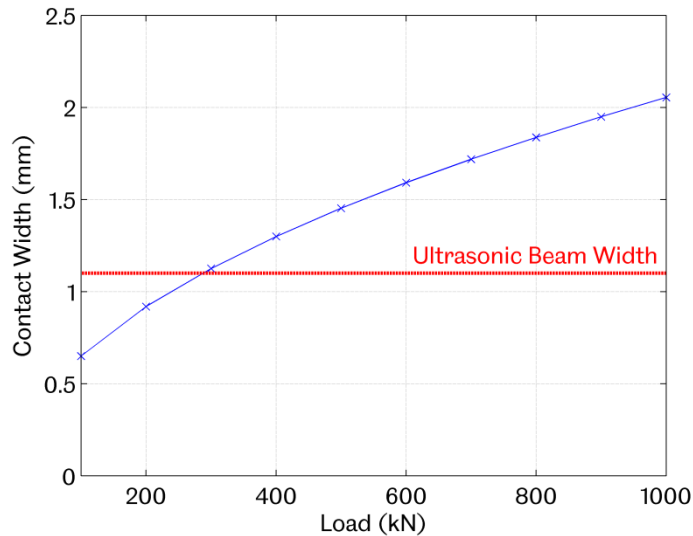
Figure 6.14: Theoretical contact width at each test load and comparison to the effective ultrasonic measurement width as determined in §4.3.3.1

## 6.3.2 Surface Deformation

As discussed earlier, the shape of the measured reflection coefficient profiles across the roller passage are not as would be expected when compared to theoretical film thickness profiles illustrated in Figure 6.10. The W-shaped profile observed during the passage of the rolling contact is clearly a result of more complex interactions of the acoustic wave with the interface, making it difficult to apply the spring model to ultrasonic measurement data with reasonable confidence. Indeed, this feature may well be masking the true response related to interfacial stiffness alone, resulting in the discrepancies between obtained results and model predictions as observed in Figure 6.11.

Given the clear relationship between the reflection coefficient profile and applied load, as observed in Figure 6.2(c) (Feature C), the effects may be partially attributable to the deflections that will inevitably take place as the rolling surface flattens under load during the rolling contact. This is considered since the degree of raceway deflection will increase under higher load as the contact width becomes larger.

Exploring this further, an initial issue immediately becomes evident as a result of such deflections in that the validity of the reference used to establish $|R|$ in the presence of the contact (according to Equation (3.20)) becomes questionable when surface deformation takes place. The

reference waveform essentially represents the reflection from the inner race rolling surface when no roller is present in the line-of-sight of the sensor, and thus represents the reflection from the un-deformed surface. Indeed, a true reference would need to be obtained from the deformed surface whilst a steel-air interface was visible, which is clearly unattainable in reality. As illustrated in Figure 6.15, flattening of the surface would effectively increase the angle of reflection from the measurement surface and thus divert energy away from the ultrasonic source element. This means that a "true" reference would have lower amplitude than the modal reference waveform obtained from the full data stream, owing to such alteration of the beam path and the subsequent reduction in reflected energy received across the source element. However, this effectively means that in determining $|R|$ (Equation (3.20)), $A(f)_{ref}$ is essentially an overestimate of the true reference amplitude, which would subsequently act to provide an underestimate of $|R|$, which is therefore contradictory to what is actually observed in the measurement data.
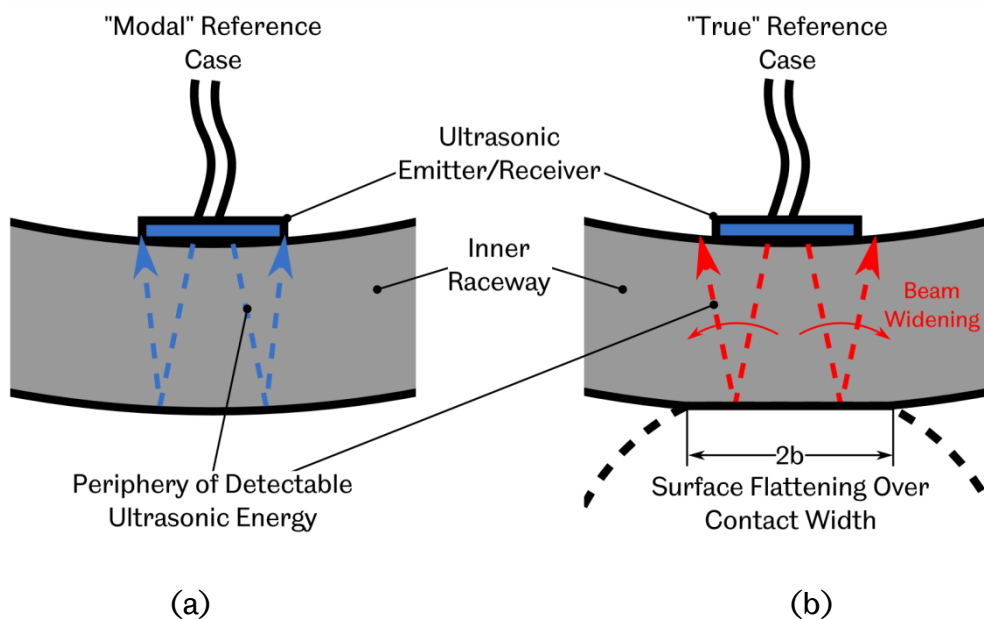


Figure 6.15: Illustrating the assumed widening of the ultrasonic beam width as a result of surface deformation under load (b) as compared with the normal reflection of energy under no surface deformation (a)

An alternative explanation therefore may relate to the change in the elastic response of the bearing material under stress. This is known to induce changes in the velocity of an acoustic wave in a process known as the "Acoustoelastic Effect" [100]. This is evident in the time domain wave maps of Figure 6.6, where the ultrasonic reflection bands shift upwards within the contact zone $(-10 \leq t \leq 10)$ as a result of increased sound velocity

through the bearing material when compressed. Simultaneously to velocity changes, ultrasonic attenuation levels have also been found to change with applied stress, in [101] for example it was found that attenuation levels increased with tensile stress, and therefore it would be logical to assume that the opposite may be true for material in compression. This would indeed provide a potential explanation for reflection profiles observed here, since the amplitude of measurement signals, $A(f)_{meas}$, would increase as a result of reduced attenuation, thus resulting in an observed increase to $|R|$ in the centre of the contact. Subsequently, such changes would be superimposed upon the ultrasonic reflection response as a function of the oil layer stiffness alone.

Overall, the discrepancies between measured reflection and spring model predictions, as discussed in this section, ultimately demonstrate that use of the spring model to provide accurate film thickness information for the proposed testing would be misleading. This is a result of high bearing load and low viscosity oil resulting in a boundary lubricated contact at bearing BDC. Analysis of Figure 6.13(c) and Figure 6.14 however, has led to an understanding that the sensor itself should indeed be appropriate for indicative assessment of the contact interface at test loads of $> 300kN$, since contact width will exceed the width of the ultrasound beam at the rolling surface.

# 6.4 Further Analysis of Higher Order Reverberations

Analysis of reflection coefficient and subsequent determination of oil film thickness is generally performed on first reflection pulses from the rolling surface, since this will typically provide the greatest clarity in terms of signal-to-noise ratios (SNR). Analysis could similarly be performed on any further reverberations through the bearing raceway, provided that pulses with reasonable SNR's are available, and bearing in mind that observed reflection coefficients become amplified by the power of the number of consecutive reverberations seen by the incident sound wave.

$$|R| = \sqrt[r]{R_{meas}}$$

(6.9)

Where $r$ is the reverberation number.
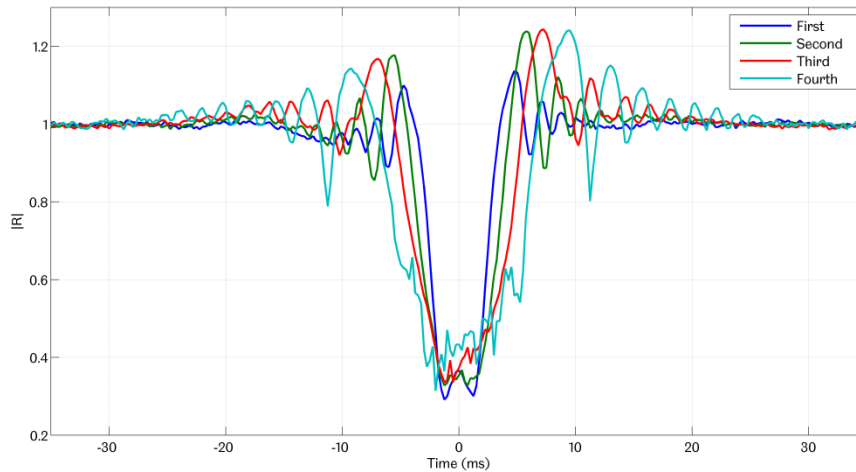
Figure 6.16: Reflection coefficient (at $10MHz$) established from analysis of the first four reverberations from the roller-raceway interface during a $1000kN$ roller pass

Figure 6.16, similarly to Figure 6.7, presents the $10MHz$ component of reflection coefficient for the $1000kN$ roller pass, here however, the first four reflections from the rolling surface are now illustrated. First and second reflections can be seen to yield similar reflection coefficient minima, yet third and fourth reflections yield slightly higher values. This is attributed to the fact that, at peak contact load, ultrasonic transmission through the interface is such that reflected energy of third and fourth reverberation pulses drop below the level of background noise, as observed in Figure 6.5, and thus all further information is lost. It is clear that reflection curves for second, third and fourth reverberation pulses are generally much noisier than the first reflection case, highlighting the earlier point that it is advantageous to utilise the first reflection for greater measurement clarity.

## 6.4.1 Amplification of Time Domain Shifts

Despite the above, the use of higher order reverberations may still be advantageous for alternative analyses. One such observation which has not yet been discussed is the apparent amplification of time domain shifts that can be observed by comparing the first and fourth reflection maps of Figure 6.6 (reproduced as Figure 6.17 at higher magnification).
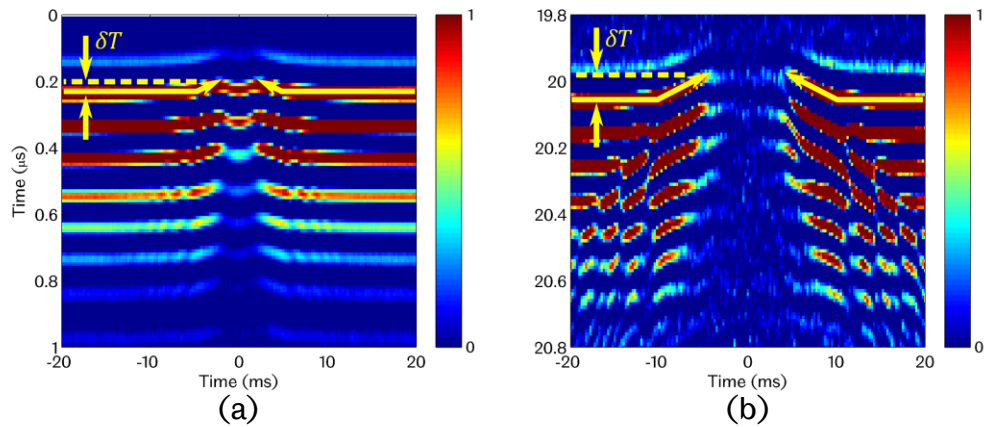
Figure 6.17: Time domain maps of Figure 6.6 in higher magnification to illustrate differences in time shift between (a) first and (b) fourth reflection pulses

As discussed briefly in §6.3.2, this observed shift is a result of the acoustoelastic effect. What is interesting however, is that the observed time shifts ($\delta T$) are much greater in the fourth reflection pulses (Figure 6.17(b)) than in first reflections (Figure 6.17(a)). This occurs since the sound wave, by the fourth reflection, has travelled through the same material four times, and therefore observes the same surface deflection and sub-surface stresses the same multiple of times as well. As is clear from Figure 6.17(a), time shifts are generally extremely small even under high contact loading, and therefore signals must be sampled at high digitisation rates in order to be able to resolve such changes accurately. Where digitisation rates are limited, such as with the FMS systems used extensively within this work ($100MHz$), the amplification of such effects in higher order reverberations may provide a means to still perform measurements at relatively low contact loads despite sampling frequency limitations.

## 6.4.2 Observation of Through-Roller Reflections

A further feature of interest was identified in Figure 6.5 relating to the measurement of a through roller reflection, illustrated in closer detail in Figure 6.18.
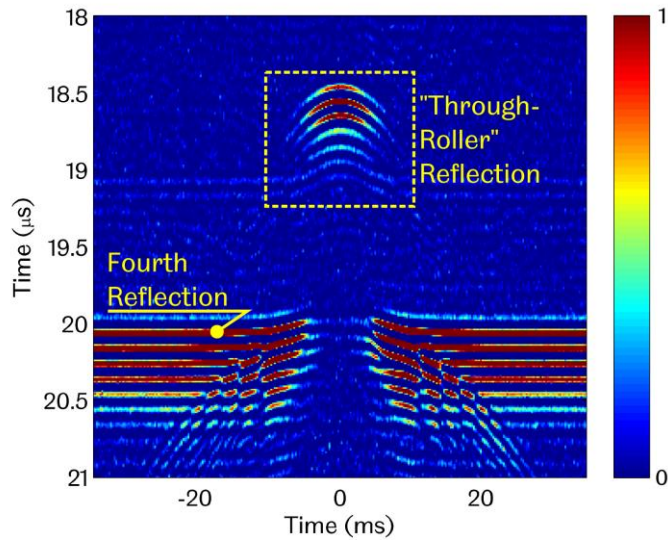
Figure 6.18: Closer observation of the colour map of Figure 6.5 illustrating the fourth reflection and the emergence of the through-roller reflection during the roller pass

Figure 6.19 presents the raw ultrasonic measurement data used to generate the plot of Figure 6.5 at two specific instances during the roller pass; (a) at $t_x \approx -30$ when no roller is present in the line of sight of the sensor (all four reflections from the rolling surface are visible), and (b) at $t_x = 0$ when the passing roller aligns directly with the ultrasonic element. As discussed previously, an additional pulse emerges with the roller present, resulting from transmission of ultrasonic energy through the roller-raceway interface, through the roller itself, and reflecting back from the opposing surface.
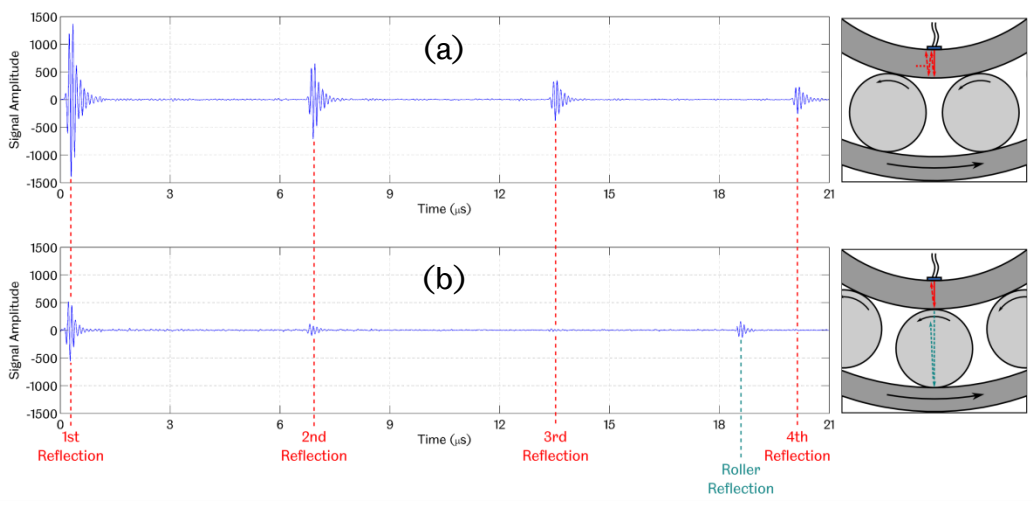


Figure 6.19: First four reflections (a) with no roller in the sensor line-of-sight and (b) when a roller is directly aligned with the sensor

Observation of this pulse may well be useful, due to its potential to indicate the boundaries of the contact patch, since contact between the inner raceway and rolling element must occur for transmission into the roller to take place, and the subsequent back face reflection to be received. The main advantage here is that reflected energy is not subject to the contact fringe effects as is the case for the reflections from the inner race rolling surface (discussed in §6.2). As such, the contact start and end points can be more easily observed and defined.
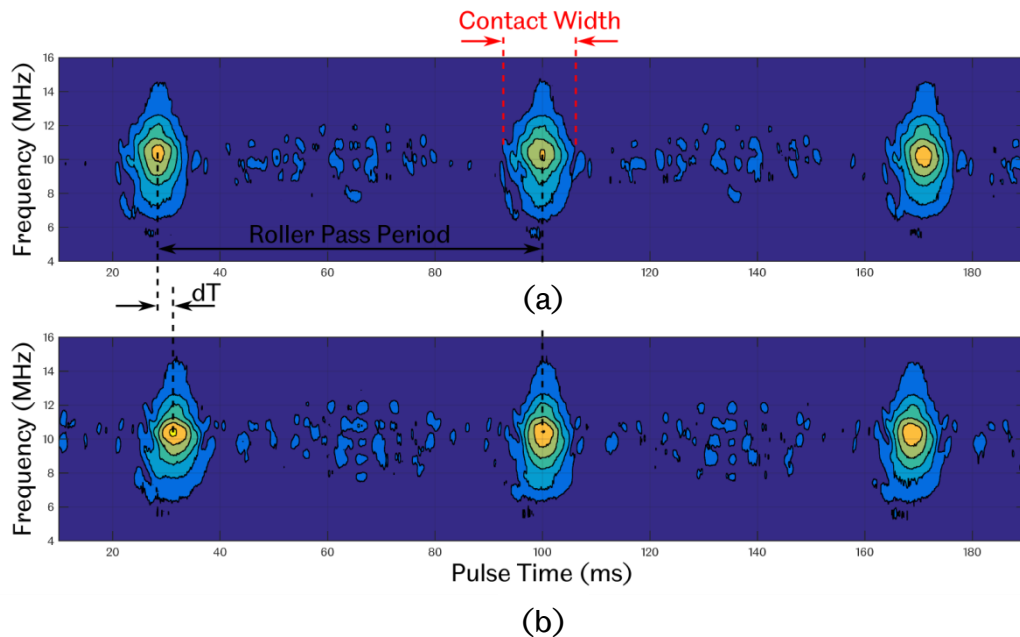


Figure 6.20: Spectrograms illustrating the emergence of the "through- roller" reflection within ultrasonic data obtained during three sequential roller passages at bearing loads of (a) $800kN$ and (b) $1000kN$

Figure 6.20 presents three sequential roller passages under bearing loads of (a) $800kN$ and (b) $1000kN$. These are presented as spectrograms of the recorded "through-roller" reflection as observed at $t \approx 18.5ms$ in Figure 6.19(b). The plots here are similar to those presented in Figure 6.6, however, where in the previous case plots represented the spectral reflection coefficient, $|R|(f)$, across the frequency range, here the plots represent the absolute spectral amplitude of each of the reflected pulses across a $4 - 16MHz$ frequency range. Actual values represented by the plots are not particularly important, and it is only relevant to note that yellow regions represent peak spectral amplitude, and areas of dark blue represent minimum spectral amplitude, which is essentially background noise. The test data used for these plots was obtained at a pulse repetition rate $(PRR)$ of $8kHz$, and thus a $0.125ms$ resolution is achieved in the $x$ axis.

As can be observed, spectral amplitude is of low magnitude until the through-roller reflection becomes visible during a roller passage, at which point amplitude rises above that of the background noise, and thus the boundaries of the rolling contact are visible. The contact width is thus defined, as shown in the figure, as the region over which spectral amplitude exceeds the background noise level.

It is however evident that the measurement data here appears to indicate a similar contact width for both load cases. This is a result of fluctuations in bearing rotational speed, as can be seen if the roller pass period is compared between the two load cases. As shown, the time between roller passages is longer for the $800kN$ load case, evidently a result of a brief drop in bearing speed during the data capture window in this case. As a result, the contact width appears larger, due to it taking longer for the roller to move past the ultrasonic element in this case. This therefore highlights the importance of an accurate knowledge of the bearing cage speed in order to permit assessment of contact width with greater certainty.

Ultimately, determination of contact width could then be used to provide direct indication of load applied to the rolling contact through application of Equation (2.8).

$$Q = \frac{\pi b^2 L E^*}{4R^*}$$

(6.10)

Clearly, this would provide significant insight into bearing operation. Further limitations to note are imposed by the fundamental principles of the ultrasonic technique. Clearly, in order to be able to accurately determine the boundaries of the rolling contact, it is desirable to obtain ultrasonic reflections at as high a rate as feasibly possible. Ultrasonic apparatus, as discussed in §3.4.2, permits pulse rates of up to $80kHz$ to be applied for single channel acquisition, meaning that the plots of Figure 6.20 could theoretically be resolved to the nearest, $12.5\mu s$ in the $x$ axis. According to time of flight through the inner raceway and roller, the back face reflection from the roller should be expected to take roughly $25\mu s$ to be received back at the source element, thus resulting in its location in between the third and fourth reflection pulses in Figure 6.19. However, due to attenuation through the bearing material, pulses must be generated at significantly high voltage in order for the amplitude of the roller back face reflection to exceed the levels of background noise. This

could be achieved by increasing the piezoelectric excitation voltage, but this would also however result in amplification of the inner raceway reverberations, and it is clear from the figure that even after the fourth reverberation through the raceway material, signal amplitude of such pulses are still significantly high. Rate of decay can in fact be estimated from the four reverberations that are observed in the figure, to indicate that in fact reverberations would only decay below background noise ($SNR \leq 1$) after approximately $50 \mu s$. This consequently dictates the maximum pulse rate that can be applied in order to avoid reverberations from a given pulse becoming observable in the subsequent pulse window and thus causing interference. Indeed, the rate of decay in this case dictates a maximum achievable pulse rate of around $20 kHz$, thus limiting the resolution at which the "through-roller" reflection curves can be obtained.

# 6.5 Identification of Lubricant Starvation

An interesting observation was brought to light as a consequence of a rig failure during an overnight test shortly after rig commissioning. A leak in the hydraulic system supplying oil to the test bearing meant that, over a period of several hours, the oil reservoir became depleted entirely, causing the test bearing to become starved of lubricant. It should be noted that the oil system at this point was less sophisticated than that described in §4.1.1.1, and no scavenge system was utilised at this stage. As such, oil was fed to the bearing at a fixed rate and return to the reservoir relied upon gravitational effects alone. A limitation to this was that rates of return were found to be much slower during the early stages of a test when temperatures were low and oil viscosity was consequently much higher. As such, the test bearing chamber was known to be essentially flooded until steady state operational conditions were achieved (oil temperatures of around $60°C$) and oil could circulate at an appropriate rate. This therefore provided observation of two extremes of bearing lubrication from which significant variations were found to exist within ultrasonic measurement data.

Figure 6.21 shows the raw ultrasonic data stream at two different stages of the test; (a) soon after initiation in the early evening, and (b) prior to rig shutdown when starvation had occurred. What is immediately apparent when comparing both plots is the shape of the region between roller passes, i.e. the reference portion of each data stream. Figure 6.21(a) – taken when the bearing is known to be flooded with lubricant – can be

observed to have a much wavier profile than that of Figure 6.21(b), where the bearing is known to have become starved.
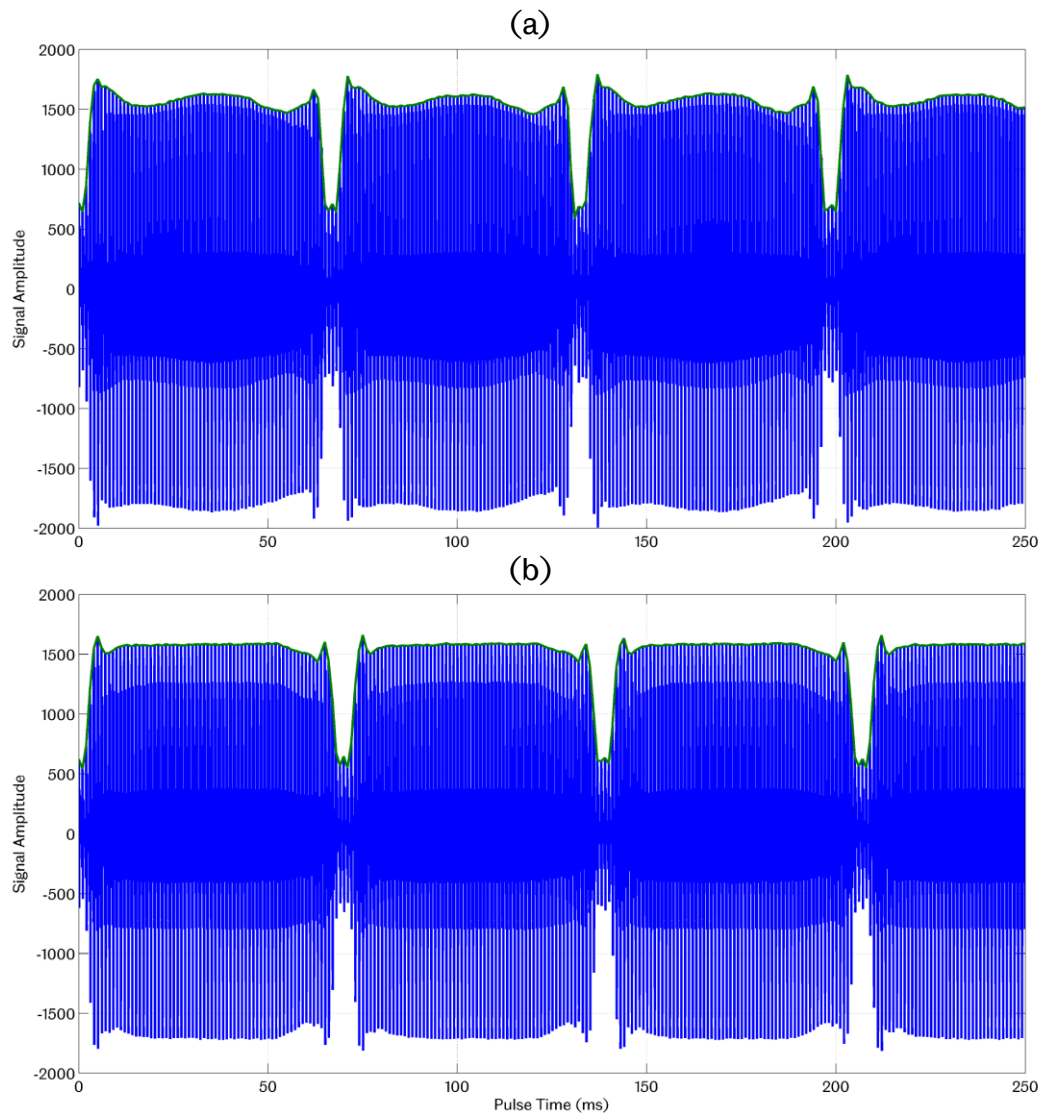


Figure 6.21: Raw ultrasonic waveform data captured at the beginning and end ((a) and (b) respectively) of an overnight test where an oil system failure was known to have occurred, consequently leading to lubricant starvation of the test bearing (bearing load and speed were maintained at $1000kN$ and $100rpm$ throughout)

To ease comparison of the data presented above, Figure 6.22 presents both datasets in terms of reflection coefficient at the resonant frequency ($8.6MHz$ for this particular data set) over a single roller pass in each case. Variations between the two are explained in the following paragraphs with reference to the annotations on the figure.

At **(1)** the presence of oil across the measurement region of the rolling surface will cause reflection coefficient to drop as ultrasonic energy is transmitted into the bulk oil and is effectively lost. Conversely, for the dry
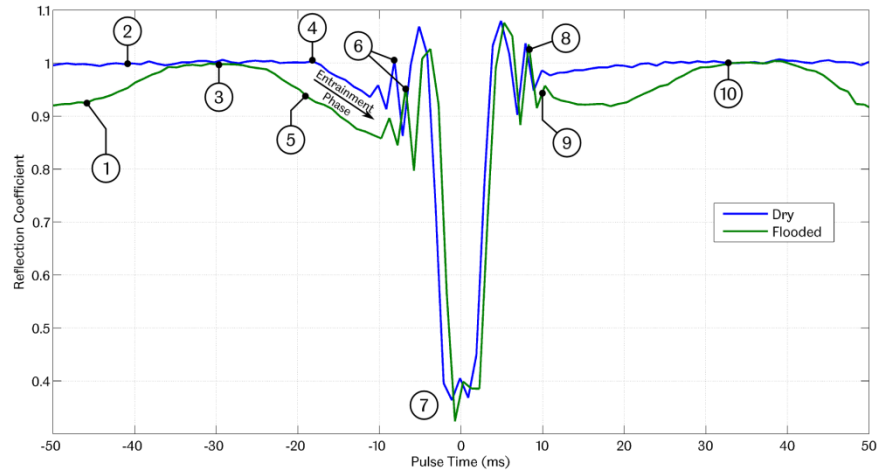
Figure 6.22: Reflection coefficient curves of the data presented in Figure 6.21 to illustrate a single roller pass under both dry and flooded lubrication conditions

case at **(2)**, reflection coefficient is generally stable at $R = 1$, as would be expected for a steel-air interface.

Reflection in the flooded case is generally unstable and only reaches unity for a brief period at **(3)**, which represents the temporal mid-point between roller passes. This suggests that at brief instances between roller passes the rolling surface is cleared of oil, despite the bearing being flooded. For the most part however, oil is present in and around the measurement region of the rolling surface, causing reduction in reflection coefficient. A similar trend in reflection coefficient reduction is still observed in the dry case, but to a much lesser extent and much closer to the contact region. This arises due to residual oil still being present within the bearing despite no new oil being supplied and thus entrainment of such oil still takes place; in this case at $-18ms$ prior to the contact centre point **(4)**. Rate of reduction in reflection coefficient during the entrainment phase **(5)** is generally linear and gradient is similar for both the dry and flooded cases. A step difference of around 0.07 is observed between the two curves, potentially a result of fluid pressure or velocity differences.

Fringe patterns at **(6)**, as discussed in §6.2, are similar in shape for both curves at the fluid inlet side of the contact, yet the same step change in magnitude between the two curves, as seen in the entrainment phase, is still observable. This again may be linked to the inevitable variations in fluid pressure or velocity that will exist within the two scenarios as a result of varying oil volumes.

Reflection coefficient profiles across the central contact region **(7)** are comparable, suggesting similar oil film formation in both cases despite the vastly differing bulk oil volumes.

Immediately after the passage of the roller **(8)**, fringe effects are again observable, but in contrast to the fluid inlet region, the curves directly overlap here. This suggests that no variation in oil conditions exist between the two cases at the fluid exit region, which should indeed be true if film formation and contact pressure within the contact were similar, as suggested above.

Curves again begin to deviate at **(9)** ($t \approx 10ms$). Reflection coefficient for the dry case returns to 1 rapidly, suggesting most of the residual oil to have been carried away in the wake of the rolling contact. Reflection for the flooded case returns to 1 gradually over the following $20 - 25ms$, finally reaching unity at **(10)**. This suggests that oil is still drawn away in the wake of the roller, but this takes much longer to clear from the measurement surface due to the much greater volumes of oil present.

The variations in ultrasonic reflection that have been observed from the test bearing under both dry and flooded lubrication regimes are particularly clear, and require very little processing of raw waveform data in order to be observable. Indeed, a lot of the patterns discussed within this section are immediately apparent from observation of the raw data streams of Figure 6.21. As such, the use of such observations as a means to assess the status of a bearing lubricant supply are easily applicable and could be feasibly applied in real-time to a wide range of bearing monitoring applications.

# 6.6  Reflection Coefficient vs Bearing Life

The primary aim of ultrasonic monitoring during full bearing life cycle testing is to identify trends and changes in data pertaining to bearing degradation. As discussed in §6.3, since oil film formation under test conditions is insufficient to fully separate the bearing surfaces, determination of real oil film thickness via application of the spring model is not possible, and would not provide meaningful insight. It was identified however, through assessment of beam width and understanding of the theoretical width of the contact that will form at the proposed test load of $1140kN$, that the sensor is still appropriate to observe the roller-inner raceway contact. Observations will therefore be made to general changes in reflection coefficient over the test duration through extraction of the

minimum reflection recorded per roller pass. As such, observation of changes to the measured minimum reflection per roller passage throughout the bearing life are thought to be sufficient for general trending.

Through full bearing life cycle testing therefore, it is hoped that the normal ultrasonic response at the roller-raceway contact (measured at the point of maximum load) can be characterised under normal operation. Once seeded defects are introduced into test bearings failures are then expected to propagate from the maximum loaded point, and thus it is anticipated that reflection coefficient changes pertaining to surface degradation will become observable as data begins to deviate from the "normal" baseline condition.

Data capture will be performed at regular intervals, defined so as to manage data volumes appropriately. Given that tests could potentially exceed two week run times, continuous streaming is deemed unfeasible due to the potential to generate unmanageable amounts of data. As such, it was defined that ultrasonic data capture should be carried out at maximum intervals of **10 minutes**, and for a duration of at least **3 seconds** in each case, so as to sufficiently observe the passage of the entire complement of rollers (according to Equation (2.18) the inner race roller pass frequency is $14.7Hz$ at $100rpm$, thus around 1.5 seconds would be required to observe the passage of all 15 rollers).

The average minimum reflection per roller pass across each three second data stream will be calculated, and from this a "Reflection Coefficient vs Bearing Life" curve will be generated to observe progression across the full life of the bearing.

Figure 6.23 illustrates the anticipated progression of such a curve from the outset of a test through to complete bearing failure. This can be broadly separated into a number of discrete phases, summarised as follows.

1. Given that test loads will typically remain constant at $1140kN$, reflection coefficient is expected to remain stable at around $0.2 - 0.3$ under normal operation (as inferred from the progression of $|R|$ vs load in Figure 6.1).
2. As discussed in §4.2.1, impact events are to be simulated after sufficient standard operational cycles have been accrued. This will be achieved via the introduction of geometric changes to the rolling surface rather than impact induced sub-surface stress concentrations. Geometric seeded defects are to be introduced

roughly at the stage indicated in the figure, after which no significant changes to reflection coefficient are expected immediately. However, it is anticipated that scatter in data will likely increase due to impingement of the rolling elements in passing.

3. With further impingement, surface damage is expected to develop at the point shown, and with this it is anticipated that average reflection coefficient will begin to drop. Since lubrication is already of a mixed regime it is expected that the relative ratio of oil separation to metal-on-metal contact will begin to reduce as any previous oil film will begin to breakdown leading to more instances of asperity contact. With this, the acoustic mismatch at the interface will begin to reduce as the individual instances of steel-on-steel asperity contact become more prevalent and dominate the contact area, leading to more transmission through the interface and a subsequent reduction in $R$.

4. Surfaces will become more heavily deformed as catastrophic failure ensues, and thus the rolling surface becomes severely non-parallel. This will likely cause scatter of the ultrasonic beam at the measurement interface, leading to less reflection back to the source element, thus resulting in further reduction to reflection coefficient.
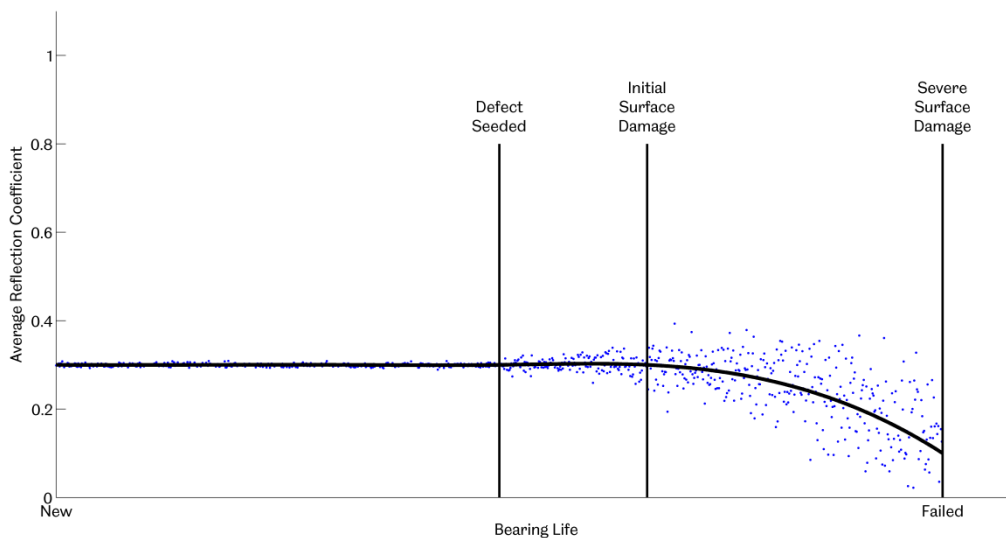


Figure 6.23: Anticipated progression of reflection coefficient through a full bearing life cycle

## 6.7  Conclusions

- Observations have been made to the ultrasonic measurements taken during a roller pass under a range of applied bearing loads. Various features have been highlighted that do not initially agree with the theoretical understanding of the rolling contact under observation.

- Divergence of ultrasonic energy through the bearing raceway has been found to induce interference or "fringe" effects in ultrasonic measurements either side of the observed rolling contact.

- It has been hypothesised that an anomalous rise in reflection coefficient at the centre of the rolling contact could be attributed to a reduction in attenuation as a result of the acousto-elastic effect.

- Due to high applied loads and the use of low viscosity lubricants it has been found that test bearings will always operate in a mixed lubrication regime and the spring model will therefore not provide meaningful insight into the thickness of the oil films formed within the rolling contacts in this case.

- The installed ultrasonic element is thought to still be suitable to observe the contact interface and provide insight into the changes occurring at the rolling surface as a bearing failure is induced.

- A "through roller" reflection has been found to be observable, which may be useful in providing a clearer indication of the width of the rolling contact, and thus permit the actual load on the contact to be inferred.

- A change in the shape of the ultrasonic data stream has been found to relate to the level of lubricant available within the bearing, and thus provides an indicator of poor lubricant supply.

- In the analysis of data during full bearing life cycle tests the minimum reflection coefficient per roller pass will be extracted and monitored. This is expected to generally remain constant until a bearing fault is induced, at which point it is anticipated that the average reflection coefficient will reduce with increased bearing degradation as a result of increased metal-to-metal contact.

# Chapter 7

# BEARING LIFE TESTING

In this chapter two bearing life tests are discussed which were based on the test conditions established in Chapter 4 to achieve accelerated life. An early test is firstly presented, which was performed prior to any ultrasonic instrumentation having been implemented, but which provides early validation of the MultiLife concept. The second test then includes a presentation of ultrasonic trends throughout the test, leading to conclusions regarding the effectiveness of the instrumentation in detecting the onset of bearing failure.

## 7.1 Un-instrumented Bearing Test

An early bearing life test was performed soon after the test rig had been fully commissioned, but prior to installation of any of the advanced instrumentation detailed in §4.3.3, in an attempt to provide early validation of the MultiLife hypothesis.

### 7.1.1 Test Conditions

A test was proposed, based upon the operating parameters identified in §4.2, which were expected to appropriately accelerate the onset of an inner raceway fatigue failure. A bearing speed of $100rpm$ was therefore utilised throughout, yet a more conservative approach was initially taken with regards to applied load, where it was considered prudent to scale back on the original definition somewhat for the initial test. Subsequently then, a slightly lower load of $1000kN$ was applied to the bearing for the

majority of the initial test. No active cooling of test bearing lubricant was performed so as to heat lubricating oils to the desired $60°C$ set point, however, due to lack of any sophisticated control, test temperatures were found to in fact stabilise at around $70°C$ for the majority of the test.

## 7.1.2 Results

Figure 7.1 summarises the test in terms of the operating conditions and associated outcomes, all according to the number of bearing cycles accrued. Discontinuities in the curves illustrate the points at which the test was temporarily stopped for short periods of time for bearing inspections or necessary rig downtime. As such, the observed discontinuities are due to test conditions being reapplied gradually up to the desired set points upon recommencement of the test. With the exception of such transients, bearing load was generally maintained at $1000kN$, as shown in plot (a), apart from a period of around 200,000 cycles where load was increased to $1200kN$.

Given the known variability in fatigue life that would exist within any bearing sample, it was desirable to firstly assess that the chosen bearing was not a particularly weak representation of the test set. The standard $L_{10(h)}$ life of the NU2244 bearing under a reduced load of $1000kN$ was found to be 120 hours, corresponding to around 720,000 cycles at a constant rotational speed of $100rpm$. Based on this, almost 500,000 test cycles were initially accrued to confirm that the test bearing did not display a particularly high propensity to failure (500,000 cycles would represent the $L_{6.5(h)}$ fatigue life based on Figure 2.6(b)).

As shown in Figure 7.1(c), at 480,000 cycles a defect was seeded in the inner raceway at bearing BDC, via the wire cut EDM process as discussed in §4.2.1. It was requested that the defect be machined to approximately $0.2mm$ width and $0.1mm$ depth, so as to be representative of the dimensions defined in Figure 5.10, however the exact dimensions of the feature were not measured at the time, and it was therefore assumed that the machining process had been carried out to specification. Bearing life was subsequently predicted through the methods discussed in Chapter 5, based on the assumed defect geometry and the test conditions illustrated in Figure 7.1(a) and (b). A 63.7 hour "post-defect" fatigue life was identified, equating to around 380,000 cycles and a reduction of almost half the standard fatigue life as defined above.
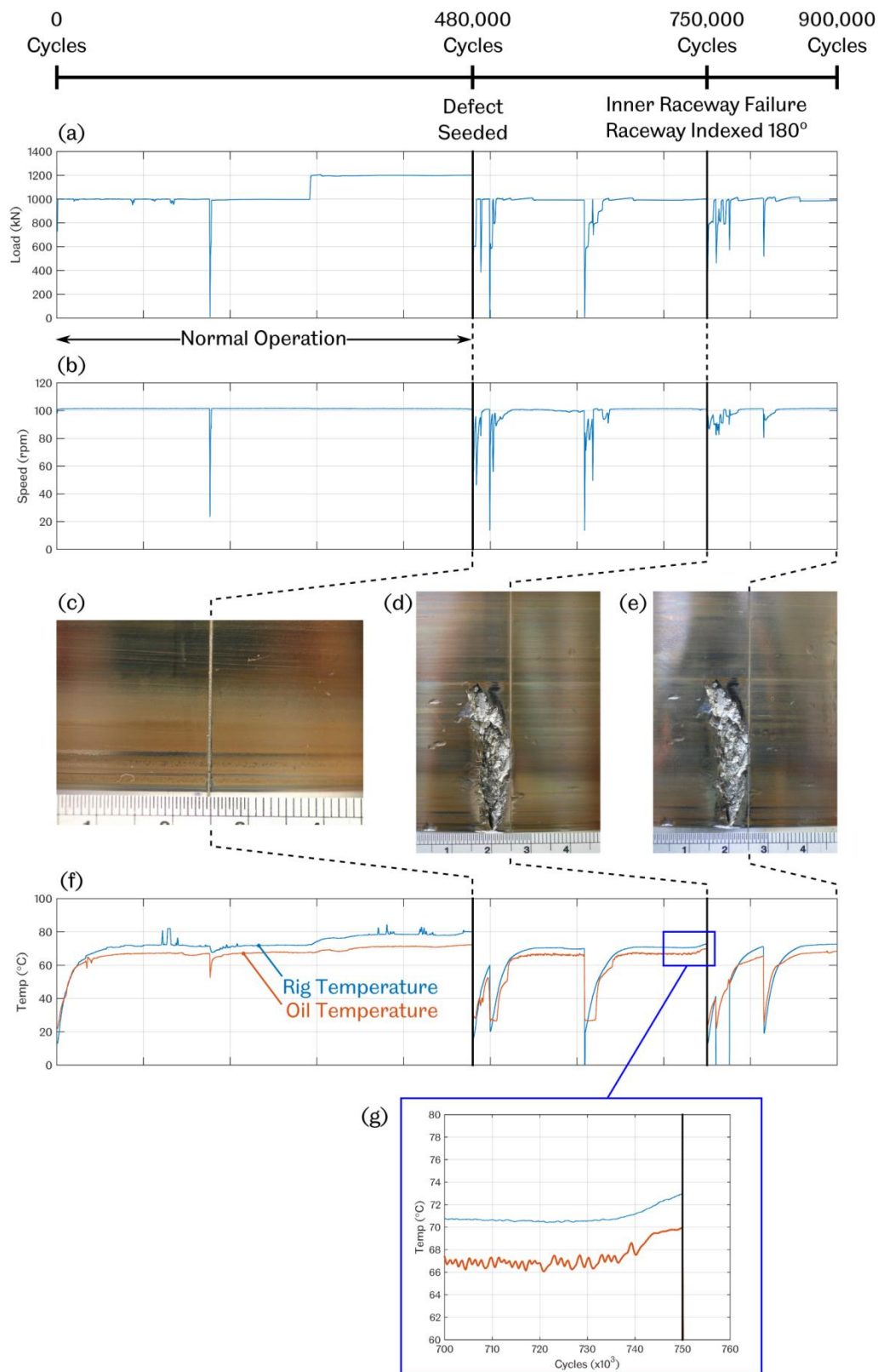
Figure 7.1: (a) and (b) Load and speed applied throughout the "Un-instrumented Bearing Test" (c) Illustration of the surface groove seeded at bearing BDC after 480,000 cycles (d) and (e) Illustration of the surface failure before and after indexing (f) and (g) Plots of lubricating oil and rig casing temperatures throughout the test

As shown in the figure, following the introduction of the seeded defect a further 270,000 cycles were performed before a failure was found to have occurred to the inner raceway. Clearly, this represented around 71% of the anticipated life as defined above, however the analysis did not account for the earlier 480,000 cycles of normal operation, and therefore a lower observed bearing life is unsurprising, given the additional fatigue cycles accrued. Nonetheless, it was clear that failure had certainly been induced due to the geometric stress concentration, at the point of maximum internal bearing load as expected.

As discussed in §2.3.1, bearing degradation can typically be identified through observation of a general temperature rise, and this was indeed evident in the temperature measurements taken during this particular test, as highlighted in plots (f) and (g). A rise is observed in both the lubricant and rig casing temperatures over a duration of around 10,000 cycles prior to rig shutdown, suggesting failure progression over a period of around 1.5 hours. It should be noted however, that measurements here were indirect, and a degree of "thermal lag" would be expected, meaning that the actual onset of failure to the inner raceway rolling surface may have occurred a considerable amount of time prior to the observed temperature rises.

### 7.1.2.1 Failure Analysis

Figure 7.2(a) shows the status of the test bearing inner raceway rolling surface across the full extent of the loaded zone, as recorded after 750,000 cycles. This highlighted an apparent misalignment of applied load resulting in edge loading of the rollers against the raceway, as evidenced by the fact that the failure has only occurred to one side of the rolling surface.

Failure had developed downstream of the surface groove (in relation to the rolling direction), which was expected since the groove would clearly cause disruption to the roller passage, therefore inducing effects subsequent to the incursion. From visual inspection, as evident in Figure 7.1(d) and (e), the wear scar appeared to be deepest at the very edge of the rolling surface, and it is therefore likely that the failure had initiated from this point. The gap observed between the seeded defect and this point suggests that in fact a degree of "lift-off" may have taken place, likely as the roller had struck the downstream edge of the groove as shown in Figure 7.2(b). This has then caused the roller to jump and strike the rolling surface further downstream, thus causing a small scale impact to

the rolling surface at this location, a process that will have then repeated following each roller passage. Given the apparent misalignment of load, this lift-off will have likely caused the roller to tilt, and strike one side of the rolling surface with greater force, as illustrated in Figure 7.2(c), causing impact at the location corresponding to the assumed initiation point.

An initial spall has then been generated as a result of a build-up of sub-surface stresses at this location eventually forming a crack which has propagated to the surface, as dictated by the classical rolling contact fatigue theory discussed in §2.1.1. Further damage has then developed from this location under continued bearing rotation, leading to the scar observed in the figure.
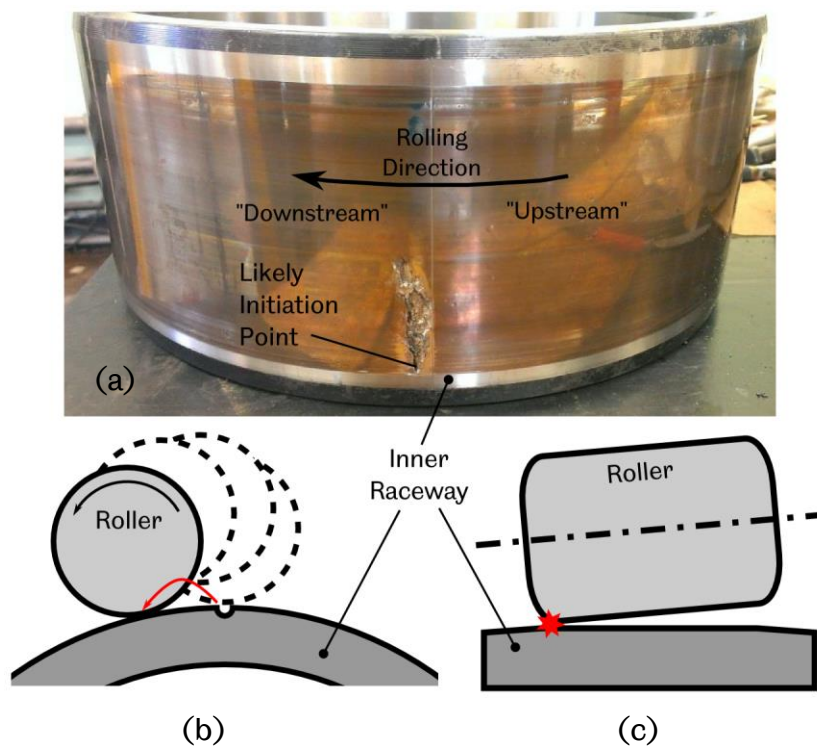


Figure 7.2: (a) Illustration of the inner raceway rolling surface across the loaded zone following inspection at 750,000 cycles (b) and (c) Exaggerated schematics of roller lift-off and edge impact assumed to have taken place as a result of the induced defect and misalignment of applied bearing load

### 7.1.2.2 Raceway Indexation

Following identification of the above surface failure the bearing inner raceway was indexed by 180°, such that the damaged region would be repositioned to bearing TDC so as to restrict any further load applied to it

and assess if all further wear progression could be halted. This was found to be successful, as illustrated in Figure 7.1(e), where – after a further 150,000 cycles – no significant change to the wear scar was observed.

## 7.2 Instrumented Bearing Test

A key conclusion of the above test was that the operating conditions defined for accelerated bearing life were indeed valid. Furthermore, a representative spalling failure could be simulated through seeding of a geometric stress concentration to the inner raceway rolling surface.

Above all however, the MultiLife concept had been proven to be effective in extending bearing life, despite significant surface failure. In reality the progression of failure to the levels observed here would be extremely undesirable due to the wear debris released and the likelihood of such material inducing further damage in other bearings and gearbox components. The need for robust condition monitoring techniques that can detect the onset of such failures were thus highlighted as being particularly important.

This is further enforced when considering the fact that only one single stress-raising "impact event" has been simulated here. In reality however (as discussed in §5.4.3), such events would occur numerous times throughout the life of a planetary bearing, and so knowledge of the occurrence and severity of such events would be highly desirable so as to enable raceway indexation only when required, thus maximising the enhancement to bearing life.

Accordingly then, the main bearing life test (as developed and discussed over the previous three chapters) was subsequently pursued and a test bearing was instrumented with both ultrasonic and temperature sensors, as described in §4.3.3. The key objectives of the test were therefore twofold; to provide further validation of the MultiLife concept, and to assess the capability of the ultrasonic technique in the detection of an inner raceway failure.

### 7.2.1 Test Conditions

Upon validation of the overall test plan, by way of the outcomes from the previous test, a similar approach was taken with the subsequent experiment, and the operating parameters as defined previously were upheld. Figure 7.3 describes the full test in terms of both applied load and

bearing speed. A load of $1100kN$ was applied for the majority of the test duration, to be more in line with the predefined $1140kN$ set point than the $1000kN$ applied previously. As can be seen in the figure, this was later increased to $1200kN$ in an attempt to further accelerate failure.

Similarly to Figure 7.1, discontinuities in the curves are a result of the test being stopped intermittently. This occurs much more frequently in Figure 7.3 than in the previous test due to the need to manually assess bearing status throughout, so that the investigation could be ceased as soon as possible following any indication of failure. As such, the test was only performed under supervision during working hours, where around 50,000 cycles could be achieved in an 8.5 hour timeframe.
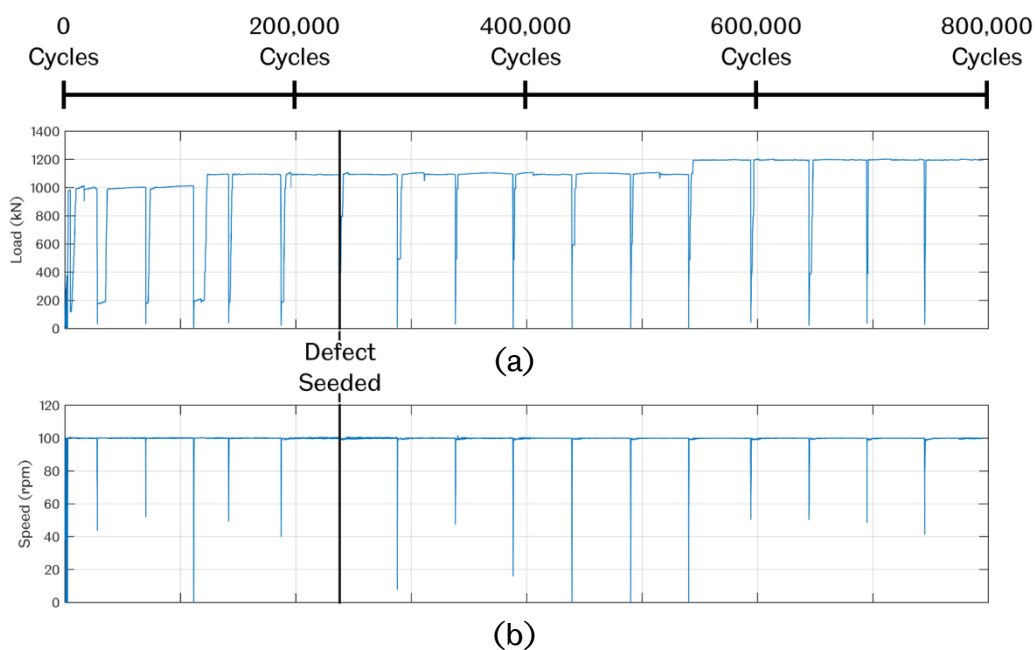


Figure 7.3: (a) Load and (b) Speed applied throughout the "Instrumented Bearing Test"

Defects were seeded in the inner raceway surface after around 240,000 test cycles, where the cycles accrued up to that point had been utilised to acquire baseline ultrasonic data, so as to characterise signals under "normal" bearing operation. Defects were introduced in accordance with the definition in §5.5, where three raceway notches were proposed at 148, 180 and 356° from bearing TDC. Figure 7.4(a) shows the inner raceway subsequent to the wire cut EDM process, where the defects at locations 1 and 2 are visible. Defect 2, positioned at the point of peak internal bearing load (BDC), is shown in more detail in Figure 7.4(b).
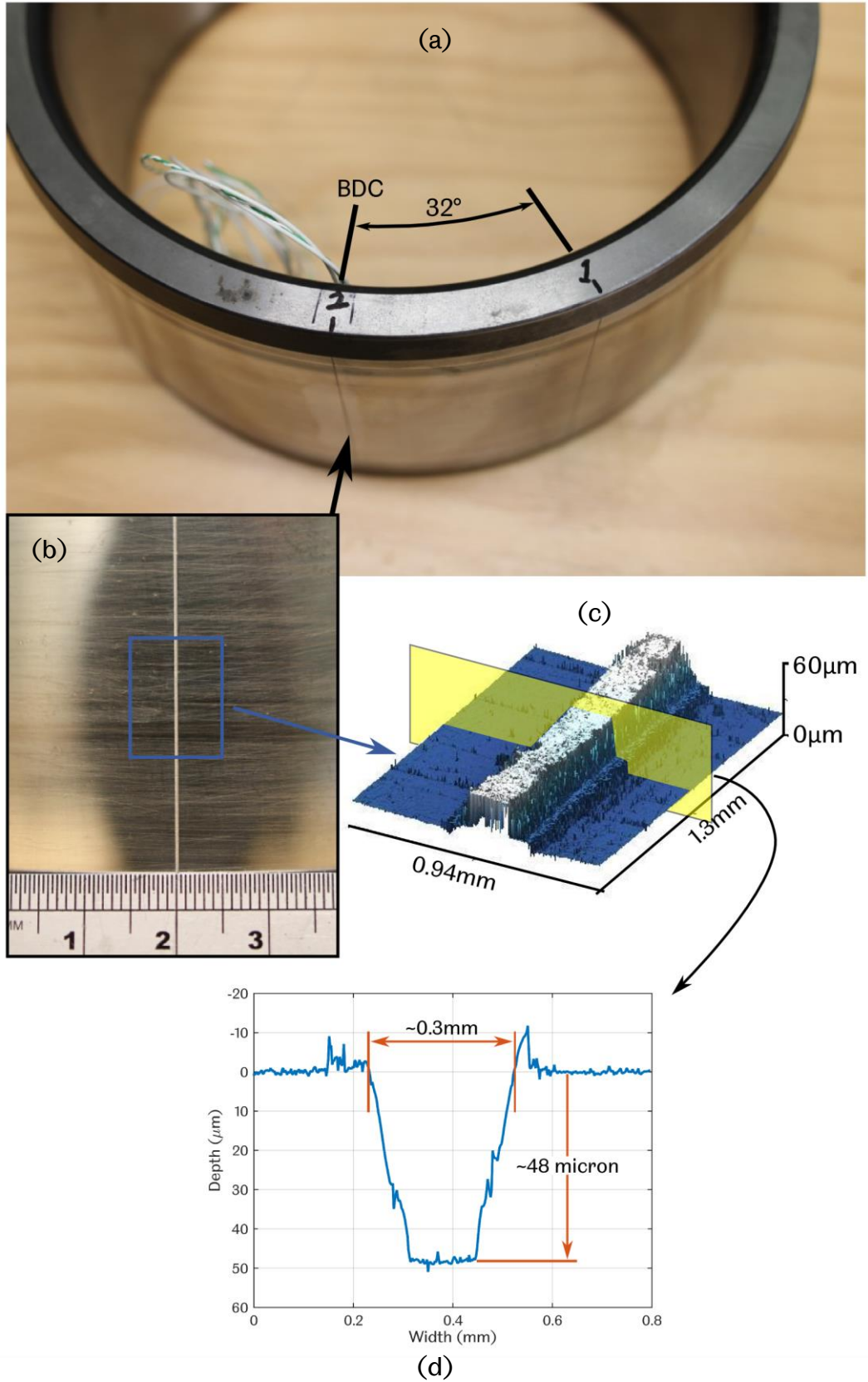
Figure 7.4: (a) Illustration of the test bearing inner raceway following the wire cut EDM process (b) Closer detail of the defect at BDC ("Defect 2") (c) 3D map of the defect (d) 2D profile to obtain defect width and depth

A machined notch of 0.2*mm* width and 0.1*mm* depth was again requested at each location across the length of the rolling surface so as to be representative of the earlier analytical work. This time however, in order to assess the accuracy to which the defect had been generated, an impression was taken of the raceway surface at each location following the machining process.

Impressions were measured via optical interferometry, allowing 3-dimensional maps to be generated of each of the geometric features. Figure 7.4(c) shows the map of defect 2. Each 3-dimensional map was then sectioned at three locations across the sample (parallel to the notch length) to obtain a 2D profile at each point. An example of one such section from the map of Defect 2 is highlighted in yellow in Figure 7.4(c) and the resultant profile is shown in Figure 7.4(d), from which the width and depth of the machined feature could be obtained. Table 7.1 details the average notch dimensions obtained from each of the defect impressions.

| Defect Location | Average Width (mm) | Average Depth (μm) |
|---|---|---|
| 1 | 0.361 | 38.85 |
| 2 | 0.353 | 50.54 |
| 3 | 0.356 | 65.32 |

Table 7.1: Average dimensions of the machined raceway defects, obtained from 3D maps of silicon impressions taken of the raceway surface and measured via optical interferometry

As evident from Table 7.1, machined width was found to be reasonably consistent across each of the defects, yet depth measurements were found to be slightly more variable. Overall however, defect geometry was found to differ significantly from that originally requested, where actual width and depth were around 175% and 50% of the respective specifications. Consequently, the bearing life prediction made in §5.5, based on the requested defect geometries, would not be entirely representative of the current test.

The analytical work performed in Chapter 5 did not extend to the investigation of bearing life with alternative defect geometries, however, a further investigation was made in [102] to explore this within the context of the same bearing and test platform. Here, through similar finite element methods, it was found that a 50% reduction in defect depth would in fact result in a 5% reduction in bearing life due to a smaller geometric discontinuity generating higher stresses within the bearing material, and thus increasing the resultant stress concentration factor. In a similar

manner, doubling the width of a defect whilst maintaining constant depth was found to result in a 12% increase in bearing life. The effects that the actual defect geometries would have on the outcomes of the current test were therefore unclear, yet given the greater relative deviation in defect width from the original specification, compared to that of depth, it was expected that the observed bearing life may be somewhat longer than the original 35.9 hour prediction made in §5.5 (equivalent to 215,000 test cycles).

## 7.2.2 Processing of Ultrasonic Data

### 7.2.2.1 Initial Observations

Figure 7.5(a) presents the first ultrasonic reflection from the inner raceway rolling surface, as obtained from the raceway mounted ultrasonic sensor. Frequency content of the signal is shown in Figure 7.5(b), where a clear bell shaped profile can be observed, centred around a $10.5MHz$ resonant frequency, and with an $8.8 - 11.9MHz$ bandwidth.
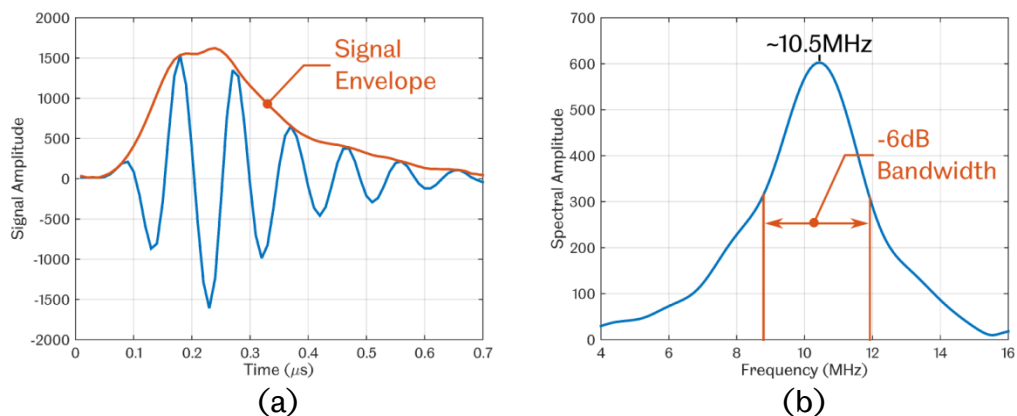


Figure 7.5: First reflection from the inner raceway rolling surface as obtained from an ultrasonic element installed at bearing BDC (a) Time domain signal (b) Frequency content

Signals were configured on the FMS system according to the settings outlined in Table 7.2.

| Excitation Voltage | $30V$ |
|---|---|
| Excitation Pulse Width | $10ns$ |
| Pulse Repetition Rate (PRR) | $30kHz$ |
| Amplitude Multiplier (Gain) | $\times 100$ |
| Delay | $6.9\mu s$ |
| Range | $0.7\mu s$ |

Table 7.2: Ultrasonic signal settings initially utilised for the Instrumented Bearing Test

### 7.2.2.2 Determination of the Average Minimum Reflection Coefficient per Roller Pass

As discussed in §6.6, ultrasonic data would be obtained in 3 second bursts at approximately 10 minute intervals throughout the test. Bearing condition would be assessed through observation of trends in the minimum reflection coefficient (MRC) obtained during each recorded roller passage. Based on initial inspection of the signals obtained from the ultrasonic instrumentation, as shown in Figure 7.5, the $10.5MHz$ component of the spectral reflection coefficient was considered the most appropriate feature to extract from the measurement data. An automated process was thus developed to achieve this, as illustrated in Figure 7.6 and briefly explained in the following paragraphs.
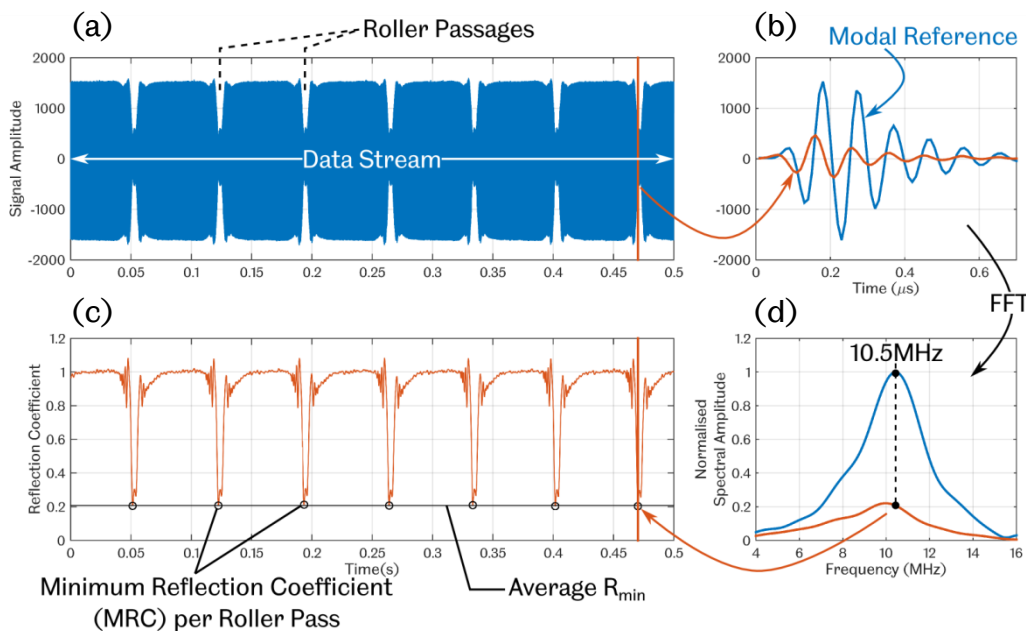


Figure 7.6: Illustration of the processing of ultrasonic data (a) A segment of a raw time domain "Data Stream" (b) Modal reference for the data stream and an example ultrasonic pulse recoded at $t \approx 0.47s$ (c) Reflection Coefficient vs Time for the data stream obtained at the resonant frequency of $10.5MHz$ (d) Frequency content of the modal reference waveform and the example ultrasonic pulse

A reference signal would be obtained for each 3 second data file such that any gradual signal changes occurring over the duration of the test were compensated for (assuming no significant change would occur over the three second data capture window). Figure 7.6(a) illustrates a 0.5 second snapshot of a 3 second data stream, and the blue curve in Figure 7.6(b) shows the corresponding reference waveform calculated for the same data set. The reference was obtained using the modal approach detailed in §3.4.3.1 and the $10.5 MHz$ component was then compared to the same frequency component of each ultrasonic pulse spectrum within the data stream, demonstrated in the figure through the use of the pulse obtained at $t \approx 0.47s$. Through analysis of all pulses within the data stream a reflection coefficient vs time curve was obtained (according to Equation (3.20)) as shown in Figure 7.6(c).

Each individual roller pass was detected by identifying the segments of the reflection coefficient curve that fell below a pre-defined threshold ($R < 0.5$ in this case). The minimum value in each segment was subsequently identified, as circled in black in Figure 7.6(c), and from these the average MRC was calculated, therefore providing a single $R_{min}$ value to represent the entire data stream.

## 7.2.3 Results

As shown in Figure 7.3, the test was ceased after 800,000 cycles due to a general rise in audible noise. This was indeed found to be a result of an inner raceway failure, induced around 650,000 cycles after the defects had been introduced into the rolling surface, thus roughly representing a 3-fold exceedance of the post-defect $L_{10}$ life predicted in §5.5. As discussed earlier, this was anticipated to a certain extent due the variation between actual defect size and the modelled geometry. Ultimately however, the vast degree of variation is likely a result of the statistical nature of the bearing life predictions and the fact that the defined bearing life was actually only representative of a 10% failure rate. Indeed, the achieved bearing life, based on Figure 2.6(b), was in fact found to represent a 76% failure rate, and therefore the predictions made in the aforementioned section were still considered valid.

### 7.2.3.1 Failure Analysis

Figure 7.7 shows the inner raceway rolling surface at the locations of each seeded defect. As expected, bearing failure had occurred at the location of Defect 2, where a small spall was found to have developed on the rolling

surface adjacent to the surface groove, as shown in Figure 7.7(b) and (d). It was apparent that the test had been stopped in a reasonably early stage of failure, evident from the size of the spall; which at this stage had not progressed as far as that observed from the previous test, and spanned approximately $5mm$ across the rolling surface and less that $1mm$ circumferentially. Similarly to the previous test, some misalignment of applied load was apparent, due to the initiation point again arising at the edge of the rolling surface. Nevertheless, it is likely that damage would have progress to a similar magnitude as that observed in Figure 7.2(a) within a very short space of time should the test have been continued.
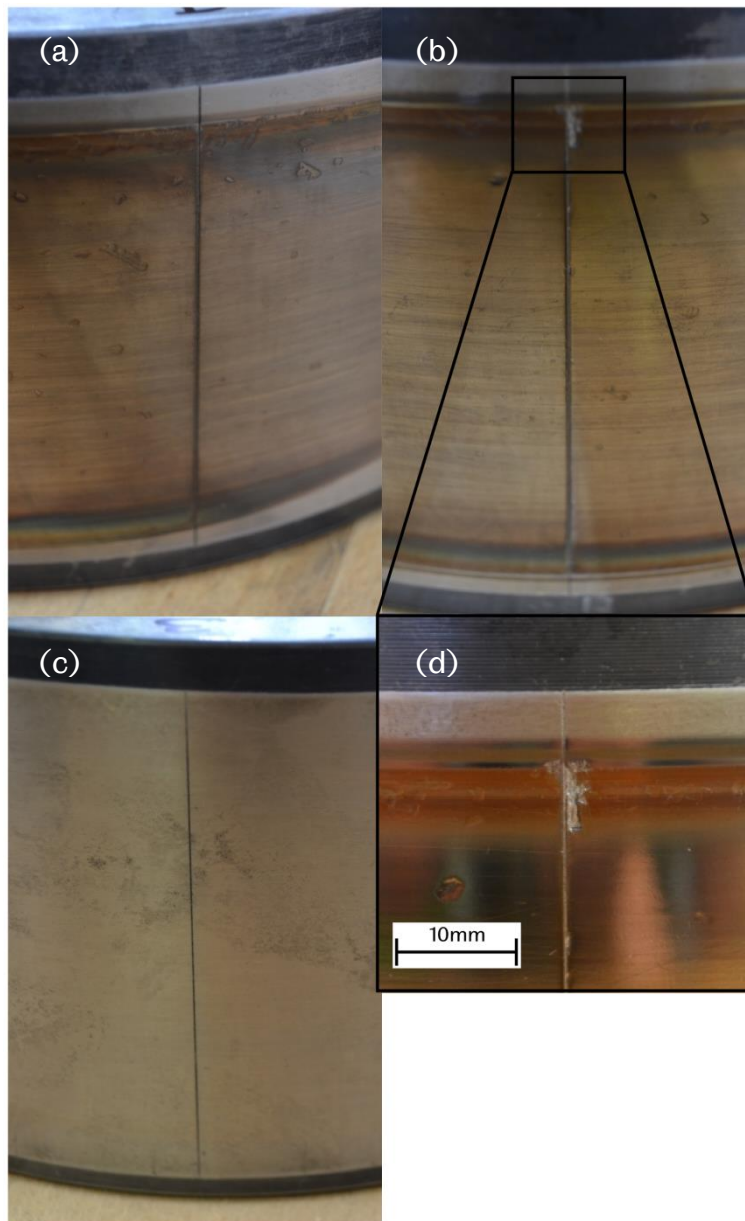


Figure 7.7: Status of the inner raceway rolling surface post-test at each of the defect locations ((a) "Defect 1," 148° (b) "Defect 2," 180° (c) "Defect 3," 356° from bearing TDC) (d) Closer detail of the spalling failure at defect location 2

Impressions were again taken of the raceway defects after the test so as to assess any dimensional changes. As shown in Figure 7.8, while the critical spalling damage to the bearing raceway was located at the edge of Defect 2, further evidence of early stage spalling was also observed at alternative points along the length of the notch at bearing BDC. No such observations were made at any other defect location.
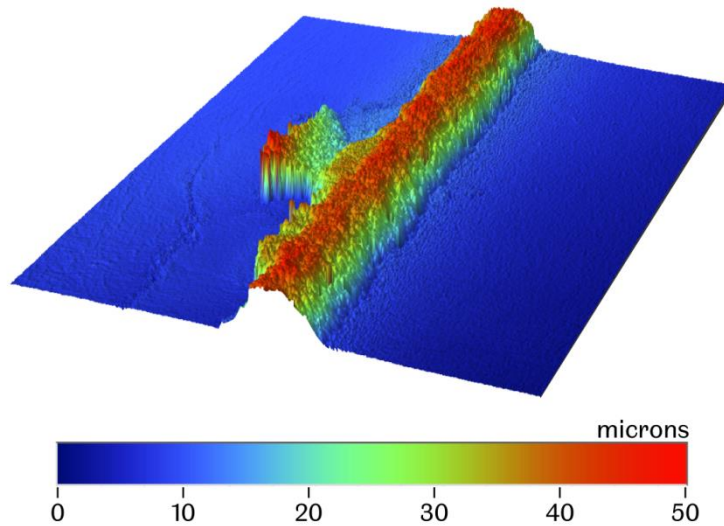


Figure 7.8: 3D map of Defect 2 post-test, illustrating additional early stage spalling damage around the defect

Table 7.3 compares the post-test measurements to the original measured dimensions as presented in Table 7.1 earlier. In some cases a negative percentage change in dimension is recorded, which is due to the pre and post-test impressions having been taken in different regions along the length of each notch, and thus slight dimensional variability along the defect length is observed. As such, variations in the order of $10 - 20\%$ were generally considered to represent negligible dimensional change overall.

| Defect Location | Defect Width | | | Defect Depth | | |
|---|---|---|---|---|---|---|
| | Pre (mm) | Post (mm) | % Change | Pre (µm) | Post (µm) | % Change |
| 1 | 0.361 | 0.551 | +53% | 38.85 | 94.03 | 142% |
| 2 | 0.353 | 0.517 | +47% | 50.54 | 43.60 | -14% |
| 3 | 0.356 | 0.323 | -9% | 65.32 | 56.61 | -13% |

Table 7.3: Post-test defect dimensions and comparison to pre-test measurements

It is clear from the table that no significant change had occurred to Defect 3, residing outside the loaded zone, while Defects 1 and 2 – having experienced cyclic loading throughout the test – show clear evidence of growth. This is further evident through visible inspection of Defect 3 (Figure 7.7(c)) when compared to either of the loaded zone defects, where both of the latter are visibly wider and darker in appearance.

No spalling damage is observed around Defect 1, lying 32° away from peak internal bearing load, and thus a reduced rate of damage progression is evident as a result of lower contact loading at this point.

### 7.2.3.2  Ultrasonic Sensor Changes

At two instances during the test the ultrasonic response was altered as a result of issues with the instrumentation, rather than changes to bearing operation. These are noted within the following paragraphs so as to be considered when subsequently analysing the overall trends in ultrasonic reflection throughout the test.

At around 110,000 cycles into the test the original ultrasonic element was broken as the bearing was removed from the rig during an inspection, and subsequently a new sensor had to be installed. The reference signals presented in Figure 7.5 represent the response from this second iteration ultrasonic element, which was utilised for the majority of the test.

A change in the response of the ultrasonic element was observed at around 595,000 cycles due to a partial de-bonding of the sensor, likely resulting from combined mechanical and thermal fatigue of the adhesive layer. During initial installation of the sensor clamping pressure was known to have not been maintained throughout the entire curing process, and therefore the observed weakness in the bond layer was not altogether surprising. Fortunately, the ultrasonic sensor was not de-bonded entirely, and a signal could still be obtained, however a significant reduction in peak-to-peak amplitude was observed. As such, it was necessary to increase both the excitation voltage and the channel gain settings so as to boost signals back to a reasonable amplitude. The partial de-bonding was also found to alter the response of the ultrasonic reflection signals slightly, as shown in Figure 7.9(a), where the response was found to be less damped, evident from the reduced rate of decay in the signal oscillations following the major amplitude peak. In addition, the resonant frequency of the signal was found to have reduced to $9.3MHz$, and the overall bandwidth of the signal was both reduced and shifted; now occupying a

$7.9 - 10.6MHz$ range, as shown in Figure 7.9(b). Subsequently, the $9.3MHz$ component of reflection coefficient was utilised in the extraction of the MRC from this point onwards.
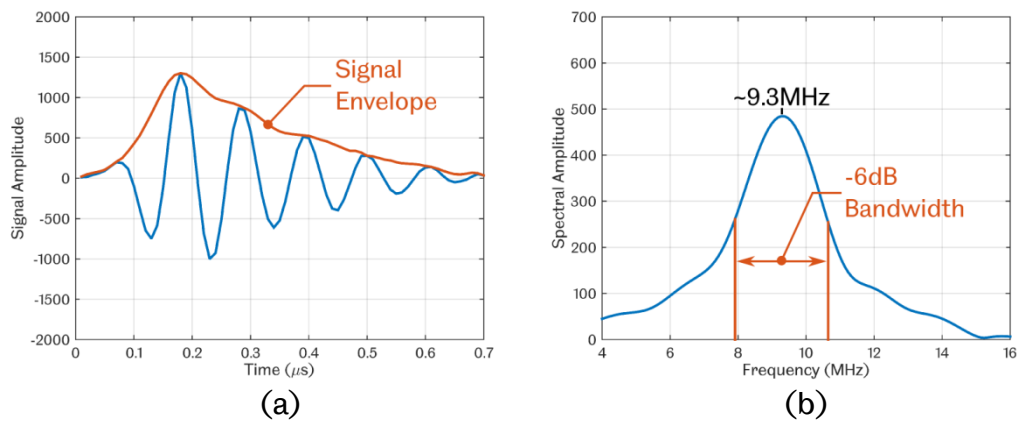


Figure 7.9: First reflection from the inner raceway rolling surface as obtained from the same ultrasonic element after partial de-bonding and a subsequent alteration to signal configurations (a) Time domain signal showing a reduced level of damping (b) Frequency content illustrating a reduced bandwidth, lower resonant amplitude and shift in the location of the resonant peak to a lower frequency level

### 7.2.3.3   Analysis of Ultrasonic Data

Figure 7.10(a) plots the progression of ultrasonic reflection throughout the duration of the test. Overlaid upon this is the applied bearing load so that any changes in reflection as a result of changes to contact loading can be accounted for where necessary. Also shown in Figure 7.10(b) alongside ultrasonic data is the temperature measurements recorded from the raceway-mounted thermocouple, installed adjacent to the ultrasonic element as shown in Figure 4.11. Due to the more discontinuous nature of this particular test, as discussed earlier, temperature would drop during downtime, and a subsequent heat-up phase is observed over a period of a few hours following recommencement of the test each morning. The temperature over-shoots occasionally observed on the figure are a result of the cooling fan on the oil supply system not being enabled from the outset of the test on such occasions.

The first point to note from the figure is the gap in ultrasonic reflection data that is evident across **Region B**, following the installation of the new ultrasonic element, and prior to the point at which defects were seeded. An ultrasonic data capture error here led to all acquired data becoming corrupted, hence the gap in the curve across this region. Unfortunately

therefore, a comparison could not be drawn between ultrasonic signals immediately before and after the defect at BDC was seeded.
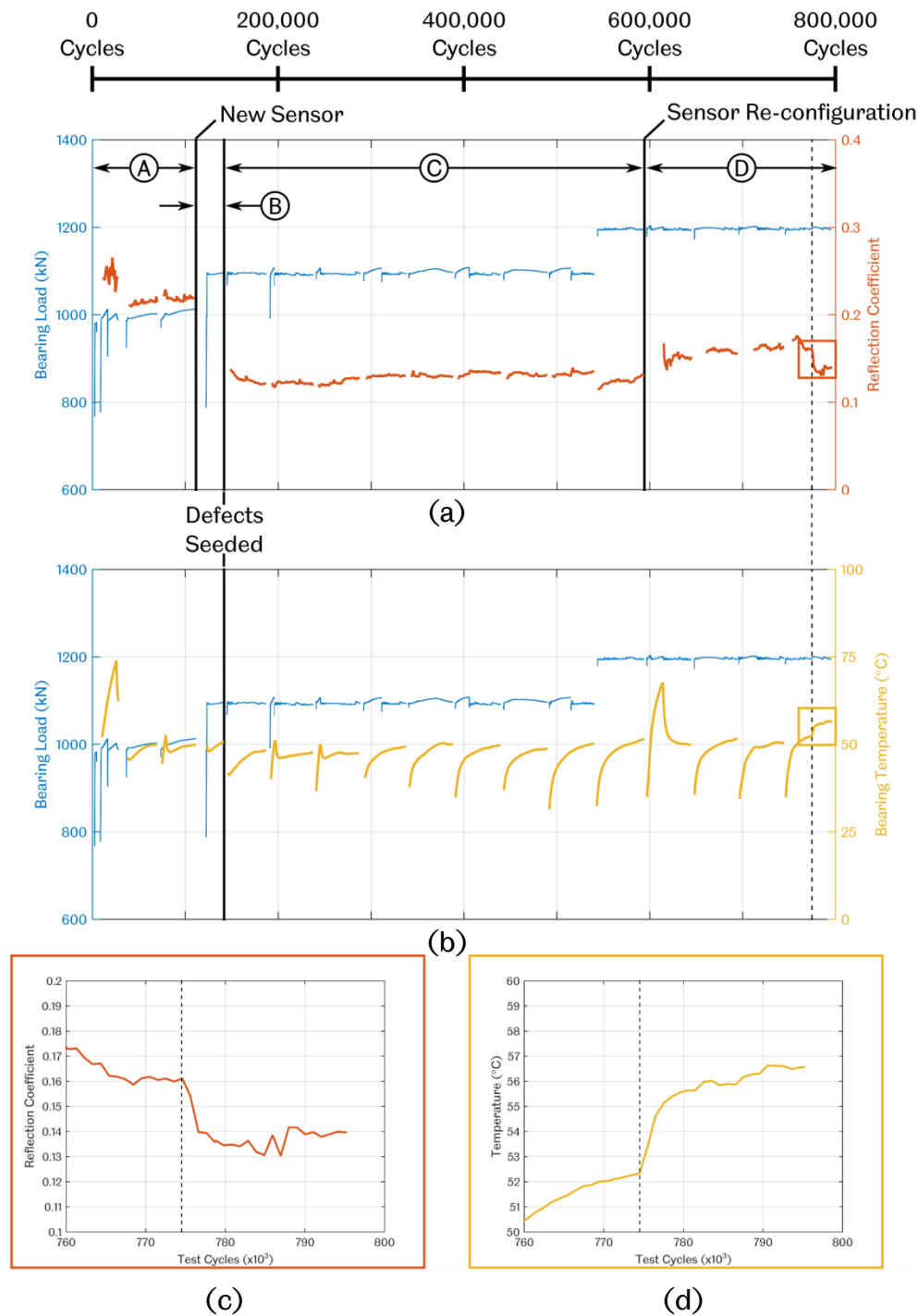


Figure 7.10: (a) Average Minimum Reflection Coefficient (MRC) and (b) inner raceway temperature vs test cycles throughout the full "Instrumented Bearing Test" with applied load overlaid (c) and (d) more detail of MRC and temperature traces respectively in the 45,000 cycles prior to the end of the test

Ultrasonic reflection tended to remain stable across **Region C** in Figure 7.10(a) and reduced – as expected – with the increase in applied load at

178

the end of this phase of the test. Ultrasonic reflection was generally higher across **Region D** than that observed throughout the earlier stage however, despite a lower load being applied throughout the majority of **Region C**. This is thought to be due to the change in the selected frequency component of $R$ at this point, which was changed from 10.5 to $9.3MHz$ due to the observed shift in the sensor frequency response, as discussed above. $R$ however, is known to be frequency dependent, as illustrated by Equations (3.10) and (3.11), and therefore the observed reflection coefficient response would inevitably be different at an alternative frequency. Indeed, a more direct comparison may have been possible had analysis been performed at $10.5MHz$ throughout. As observed in Figure 7.9 however, this would represent the upper bandwidth of the signal subsequent to the partial de-bonding, and it is therefore likely that the reflection coefficient would have displayed a lower sensitivity, as well as increased noise, at this frequency in any case.

Indeed, even at peak resonance, reflection coefficient was generally noisier throughout the latter part of the test, which was clearly a result of the weakened bond layer causing a more variable sensor response throughout. On average however, reflection coefficient generally fluctuated within the range of $0.16 - 0.17$.

The broken line in the figure, located at 775,000 cycles (20,000 cycles prior to rig shutdown) indicates the point at which a considerable drop in ultrasonic reflection was observed, shown in more detail in Figure 7.10(c). This represents a considerable deviation from the general trend observed throughout the earlier part of this test stage, and therefore appears to indicate a sudden alteration to the conditions at the bearing contact. This corresponds directly with a sudden rise in the inner raceway temperature, as illustrated in Figure 7.10(d), which was identified in the previous test as a key indicator of bearing damage.

Indeed, it could be argued that the ultrasonic response has been induced as a direct response to the local temperature rise, given that the step response appears to occur simultaneously in both measurements. The two can be decoupled however through consideration of the general trends in both curves throughout the test. As discussed earlier, the discontinuous nature of the test led to the recurrent heat ramps that are observed in the temperature curve of Figure 7.10(b), yet ultrasonic reflection is seen to remain constant throughout these variable temperature regions. This was achieved through the use of a "real-time ultrasonic reference" for each obtained data stream, which was discussed in §3.4.3, and was successful in compensating for any local temperature

changes that may have altered the piezoelectric and/or bond layer response over time. Based on this, the ultrasonic response can be confidently attributed to alterations in the observed contact, rather than simply arising in response to the observed temperature rise.

## 7.3  Discussion

The bearing life tests presented within this chapter provide sufficient evidence to validate the MultiLife principle. A significant fault has been induced to a bearing rolling surface, yet further operation of the bearing has been possible by relocating such damage away from the bearing loaded zone, thus halting any further progression. Indeed, the levels of damage induced during the test presented in §7.1 were extreme, and if such a degree of failure were to arise in an operational WTGB, the presence of wear debris alone would ultimately result in failure to the bearing and adjacent components if the system was to remain active and no maintenance was immediately carried out. The key therefore is to detect the onset of such damage prior to catastrophic failure, or to relocate loaded segments of raceway material regularly so as to distribute fatigue appropriately. To support this, it has also been shown, through the test in §7.2, that relocation of a weakened bearing region by as little as 32° from the point of maximum bearing load is sufficient to delay the onset of failure.

The onset of failure was identified through observation of a general rise in temperature, which was found to occur around 10,000 cycles prior to the end of the initial "Un-instrumented" bearing test, but a further 10,000 cycles earlier in the second "Instrumented" test. Given the degree of damage induced during the initial test, the local temperatures at the bearing inner raceway were likely much greater than those observed from the indirect oil and rig casing measurements. Furthermore, the amount of time required for local inner raceway temperatures to affect those of the oil and rig casing would be considerable, hence the relatively late detection in this case. It is therefore expected that if the bearing inner raceway had been instrumented with a thermocouple at this stage, a temperature rise would have been observed much earlier. A much greater temperature rise was indeed observed from the direct raceway measurements made within the second test, which were a result of a much smaller surface failure, thus highlighting the vastly increased sensitivity in this case.

Ultrasonic reflection was found to reduce, as predicted in §6.6, with the onset of damage, albeit at a much faster rate than that originally anticipated (hypothesis presented in Figure 6.23). This reduction is thought to have been induced by a disruption to the lubricant layer at the point at which damage had initiated, which is likely to have occurred as soon as the first spall was released from the rolling surface, subsequently allowing oil to flow into the resultant surface pit. Indeed, the tendency for this to occur with the test bearing lubricant utilised here would already be high due to its low viscosity grade, therefore causing it to present very little resistance to flow. It was previously identified in §6.3.1 that the inner raceway-rolling element contact under observation by the installed ultrasonic element would be boundary lubricated, and therefore any slight change to the bearing surfaces would likely have a profound effect upon the lubricity of the oil across the entire contact. It is therefore postulated that the damage observed at the edge of the rolling contact has resulted in a reduction in the presence of oil across the entire contact length, causing the relative ratio between local metal-to-metal contact areas and oil separated regions to shift further towards a more solid-solid dominated contact. Subsequently then, there has been an increase in the transmission of ultrasonic energy through the interface, thus reducing the amount of reflected acoustic energy received by the ultrasonic source element.

The observed ultrasonic response could be positively linked to the failure of the raceway surface due to it coinciding directly with an increase in inner raceway temperature; known to be an indicator of such phenomena as discussed within §2.3.1, but also demonstrated convincingly through the initial bearing life test in §7.1. Given that the observed ultrasonic response was triggered at the same time as the temperature rise, it would appear that the technique offers no significant advanced warning of failure over more classical temperature monitoring techniques. However, as has already been established in §4.3.3.1, the area of the rolling surface that is monitored by the bonded ultrasonic element is small, equating to a width of around $1.1mm$ in the rolling direction, and no more than $6 - 7mm$ along the contact length. Clearly then, as illustrated in Figure 7.7(b) and (d), the failure induced within this particular test has occurred well outside the observable area of the ultrasonic beam, which may explain the delayed detection in this particular case. It is therefore considered likely that the onset of damage may well have been detected much sooner should failure have occurred within the detectable area of the sensor.

It is of further interest to note that temperature measurements were found to be vastly more responsive to bearing failure when installed directly upon the component itself, rather than on external casings. This alone illustrates the advanced warning capabilities that would be possible through use of even basic monitoring techniques if used more within the wind turbine system. This ultimately highlights a need for such measurements to be considered in the early stages of design of new WTGB's such that accommodation is made for as many bearings as possible to be individually monitored even on a basic level.

# 7.4 Conclusion

- Failures have been simulated on the test platform, which are representative of those known to occur in planetary WTGB bearings. This was achieved through accelerated life testing of two NU2244 test bearings based on the test conditions proposed in Chapter 4, and with geometric faults seeded in the inner raceway to simulate the stress concentrations induced in the bearing raceways in-service as a result of impact events in the WTGB.

- The MultiLife concept has been validated by illustrating that such failures do not exhibit any further propagation once moved away from the bearing loaded zone. Furthermore, it has also been demonstrated that the onset of a failure can be effectively reduced by locating a fault seeded region of an inner raceway as little as 32° away from the point of maximum internal bearing load.

- Ultrasonic reflection was found to reduce as a result of an inner raceway spalling failure, as predicted in §6.6. This was found to be a result of the breakdown in the oil layer, which subsequently increased direct metal-to-metal interactions within the contact.

- However, the ultrasonic response was found to provide no advanced warning of failure in this case when compared to standard temperature measurements.

- This was attributed to the fact that the failure did not occur across the area of the rolling surface that was observable by the ultrasonic element, and it is therefore expected that much sooner warnings may have been obtained, should this have been the case.

# Chapter 8

# APPLICATION OF SENSOR SYSTEMS TO AN OPERATIONAL WIND TURBINE

Convincing evidence was presented in the previous chapter to not only validate the MultiLife principle, but to also highlight the potential for the ultrasonic technique as a condition monitoring tool; providing indication of oil film breakdown during the early stages of failure. Clearly however, there exists a crucial link between both the above elements to ensure the successful implementation of a MultiLife system in an operational wind turbine. Indeed, MultiLife actuation could be optimised if bearing raceways are only indexed as and when required, based on feedback from reliable sensors that can detect when bearing condition is becoming critical to performance. Great importance therefore lies in the further development of such condition monitoring tools and associated algorithms to determine remaining useful bearing life and prompt raceway indexation autonomously, where the analytical work developed in Chapter 5 provides a starting point for such analysis.

Initially therefore, a key challenge was to extend the laboratory based work performed within this thesis to an operation wind turbine. This would not only provide further validation of the MultiLife concept, but would also provide insight into the typical response of ultrasound – and other sensor systems – to real turbine operating conditions, with the added possibility that a gearbox bearing failure might, at some stage, also be observed. All of this would ultimately aid in the development of the aforementioned algorithms for intelligent MultiLife actuation.

This chapter therefore presents the authors' involvement in the Ricardo-led Offshore Wind Drivetrain Innovation (OWDln) research programme, which was primarily focussed around successful application of the MultiLife technology, but also aimed to supplement this work with advanced condition monitoring techniques for the overall wind turbine system.

## 8.1 Background

A combined Wind Turbine Condition Monitoring System (CMS) was to be developed as part of the OWDln project through a consortium formed between Ricardo, the University of Strathclyde and the University of Sheffield. Figure 8.1 illustrates the proposals for the CMS suite, and describes the individual responsibilities of each project partner. Generally, The University of Strathclyde were to implement a system to monitor overall turbine parameters such as temperatures and vibration on key turbine components ("CMS1"), and Ricardo would develop a system largely focussed on MultiLife actuation and monitoring ("CMS3").
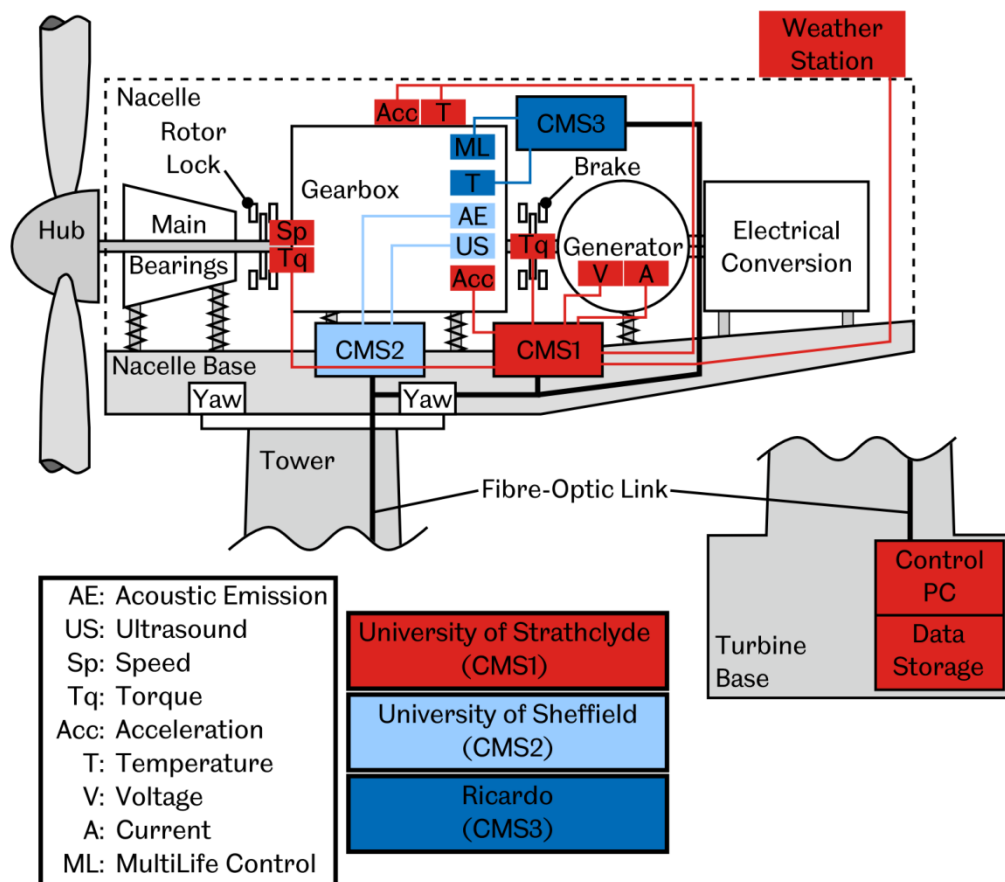


Figure 8.1: Full CMS sensor definition and breakdown of individual system requirements (adapted from [103])

The author, as part of the University of Sheffield, would implement an advanced sensor system for bearing monitoring at a critical gearbox location utilising the Ultrasound technique as discussed throughout this work, and AE sensors (this system is herein referred to as "CMS2").

### 8.1.1  Wind Turbine Specification

A Vestas V42 ($600kW$) wind turbine was made available for this work by Scottish Power Renewables. This was located at Barnesmore Wind Farm in the Republic of Ireland, and was proposed due to its exposure to the harshest wind conditions of all turbines across the Barnesmore site, which has subsequently resulted in a high propensity to failure in recent years. As such, the probability that a gearbox/turbine failure would be observed was high, and the opportunity to demonstrate planetary bearing life enhancement through application of the MultiLife mechanism was also significantly increased.

### 8.1.2  Gearbox Bearing Specification

A MultiLife actuation mechanism was to be installed upon each of the three planetary bearings within the epicyclic stage of the $600kW$ gearbox. Ideally then, instrumentation of AE and ultrasound sensors would have also been applied to one or all of the static inner raceways of such bearings so as to monitor the onset of failure and support operation of the MultiLife systems. However, instrumentation of such bearings was particularly challenging since, as illustrated in Figure 1.4, each bearing would rotate about the rotor axis during operation (through rotation of the planet carrier), thus the use of slip rings or wireless systems would have been necessary to facilitate the extraction of signals. In the case of the former method, this would have required significant alterations to the gearbox, which was already the subject of modifications to enable application of the MultiLife systems. The latter "wireless telemetry" method may have been possible, but this had not been widely utilised for transmission of such high frequency signals, and would therefore have required significant hardware development and testing. It was therefore decided that an alternative bearing should be instrumented at a location that was still known to observe high failure rates (as identified in §1.2.2), but which was more easily accessible for application of conventional signal cable routing methods.

The High Speed Shaft (HSS) was thus identified as being the most appropriate area for bearing instrumentation and thus further analysis was carried out to identify the appropriate bearing within the HSS assembly. Figure 8.2 shows the bearing configuration supporting the HSS. A tapered roller bearing (TRB) pair is employed to support the shaft at the front (input) side of the gearbox, accommodating both radial and axial load, and a single CRB is employed at the gearbox rear (output) side for support of radial load only.
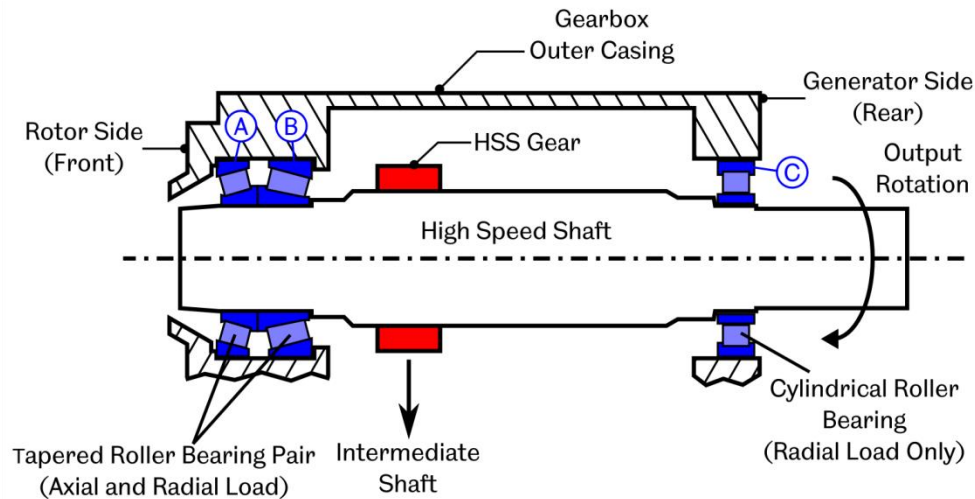


Figure 8.2: High speed shaft bearing configuration

An analysis of gearbox loading was performed by Ricardo engineers using in-house gearbox modelling software "SABR," to identify the expected loads upon each of the HSS bearings, as well as the associated fatigue life, based on the maximum rated torque applied to the gearbox ($600kW$) at its maximum rated input speed ($30rpm$). Through this it was found that **Bearing B** in Figure 8.2 would be subject to the highest equivalent radial load of all three bearings within the HSS assembly ($P = 45.8kN$), and this would subsequently possess the lowest expected fatigue life. The bearing utilised here was a 32222 Type Tapered Roller Bearing, general details of which are provided in Table 8.1.

186

| | | |
|---|---|---|
| Bore diameter, $d$ | 110$mm$ | |
| Outer diameter, $D$ | 200$mm$ | |
| Width, $b$ | 56$mm$ | |
| Roller Diameter Min, $d_e^{min}$ | 22.68$mm$ | |
| Roller Diameter Max $d_e^{max}$ | 25.76$mm$ | |
| Roller Length, $L$ | 37$mm$ | |
| Outer Race Contact Angle, $\phi$ | 17.77° | |
| Dynamic load rating, $C$ | 402$kN$ | |
| Number of rollers, $Z$ | 20 | |
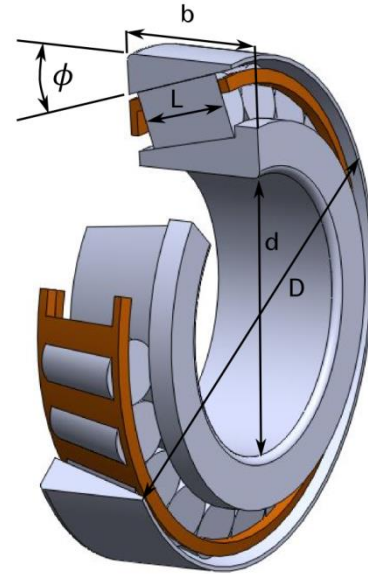
Table 8.1: Specification and illustration of the 32222 Tapered Roller Bearing to be instrumented

## 8.2 Methodology

### 8.2.1 Bearing Instrumentation

Upon identification of the critical HSS bearing, SABR modelling was further utilised to gain a more detailed understanding of loads that would be induced upon the 32222 TRB. Figure 8.3(a) shows a map of the normalised stress distribution around the bearing circumference. From this, the positioning of two ultrasonic elements was defined, as denoted by the vertical broken lines in the figure, which would be installed onto the static outer raceway of the bearing; one at the point of maximum bearing load, 201° from bearing TDC, and the other located 40° further around the raceway circumference, at 241° from TDC. The latter was chosen to represent the outer boundary of the peak loaded area, thus permitting observation of the rolling contact moving in and across this zone. Figure 8.3(b) illustrates the assumed distribution of internal bearing load ($Q_\varphi$) through the centre of the outer raceway, to approximate the expected loads that would be "visible" to ultrasonic elements installed at the proposed locations. From this the definitions in Equation (8.1) were established for contact load at sensor locations 1 ($Q_{US1}$) and 2 ($Q_{US2}$), where $Q_{max}$ is established from Equation (2.1) considering that bearing load would equal 45.8$kN$ at gearbox rated power, as detailed in the previous section.
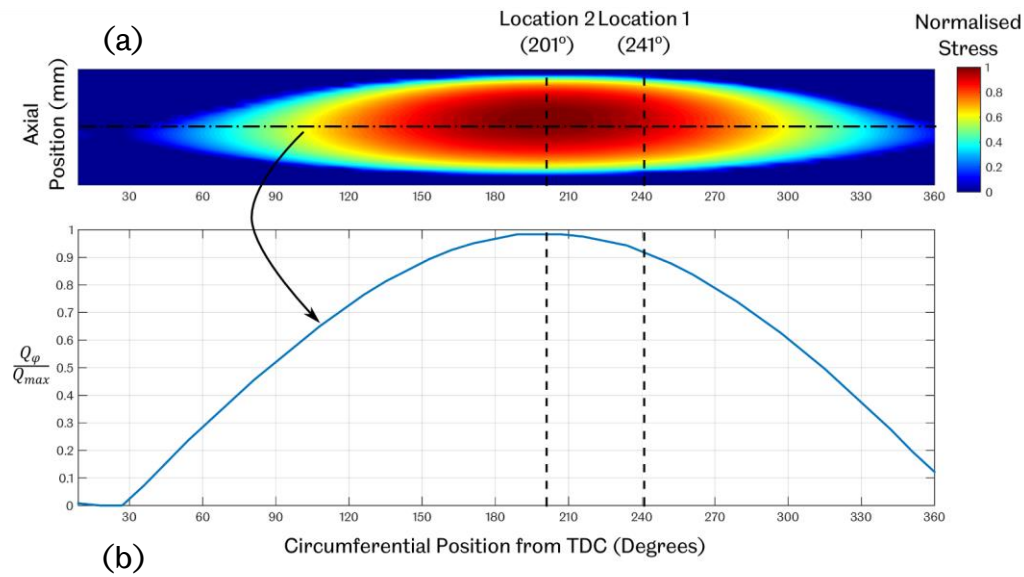
Figure 8.3: (a) Expected stress distribution (normalised) around the maximum loaded HSS bearing (Bearing B, Figure 8.2) (b) Inferred distribution of internal load $Q_\varphi$ through the centre of the outer raceway (normalised to $Q_{max}$) based on a cross section of the stress distribution map in (a)

$$Q_{US1} = 0.92Q_{max}$$

$$Q_{US2} = 0.98Q_{max} \tag{8.1}$$

Figure 8.4 illustrates the arrangement of such instrumentation upon the 32222 bearing, as well as all additional sensors that were to be employed. As can be seen in the figure, accommodation was also made for both acoustic emissions and accelerometer sensors, however, given the focus of this study so far, the remainder of this section will focus primarily on the installation, acquisition and data analysis of ultrasonic signals only.

It was necessary to modify the bearing outer raceway with a shallow, tapered circumferential recess to provide a surface that was parallel to the outer-raceway rolling surface. This ensured that ultrasonic acquisition could be performed in "pulse-echo" mode, where the same ultrasonic element could be used to both emit and receive ultrasonic waves transmitted through the bearing raceway. As such, a single ultrasonic element could be utilised at each measurement location, thus reducing the total number of elements required, and simplifying the instrumentation considerably. $10MHz$ piezoelectric elements were utilised, as introduced in §3.2.3, which were cut down to rectangular strips of $1 \times 5mm$; a $1mm$ width being the smallest that could feasibly be applied

to maintain a reasonable performance from the element. Thermocouples were also installed directly upon the outer raceway, adjacent to the ultrasonic elements, so as to provide the closest approximation of the temperatures occurring at the contact under observation by each ultrasonic element.
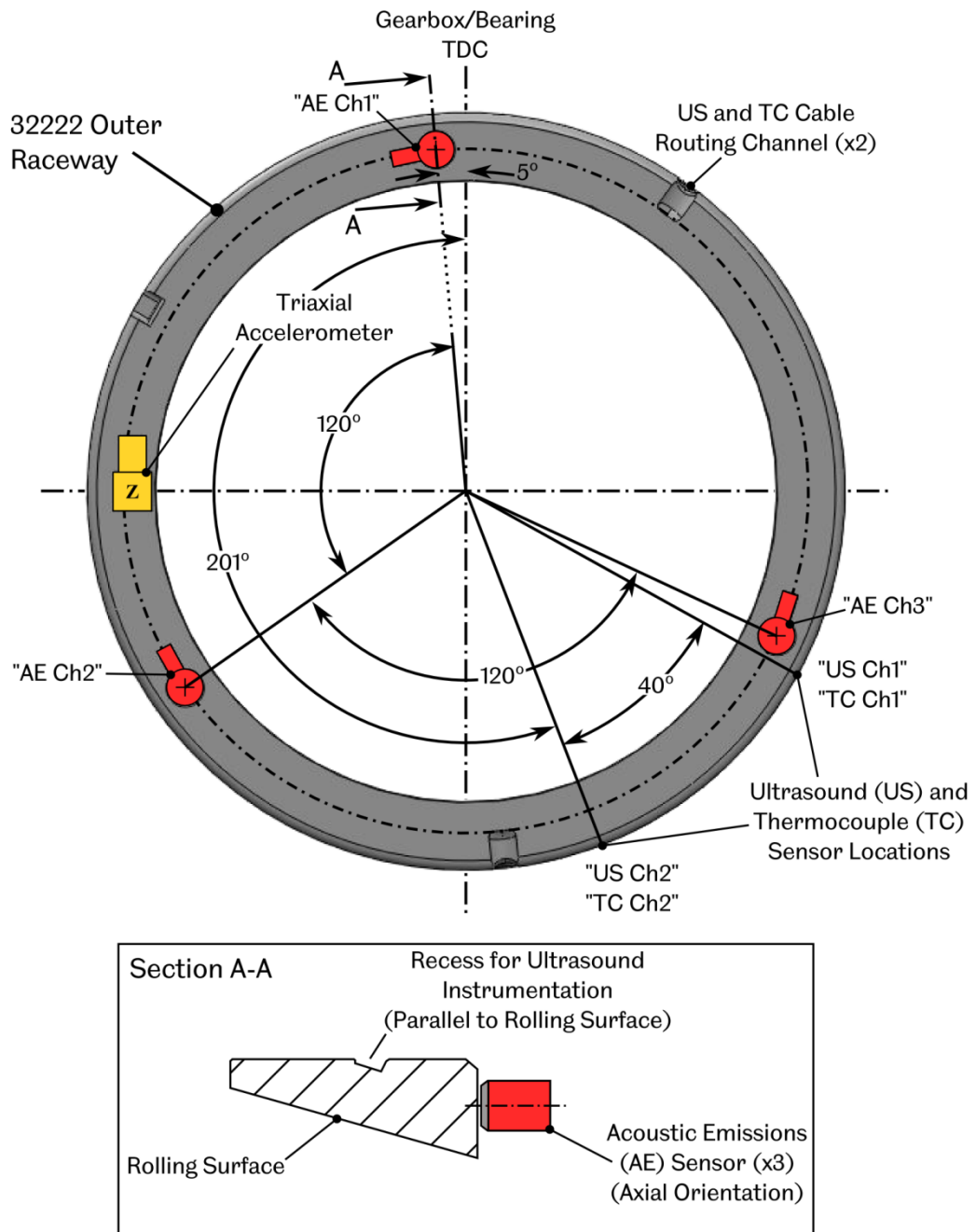


Figure 8.4: Location of sensors installed on the 32222 HSS bearing outer raceway
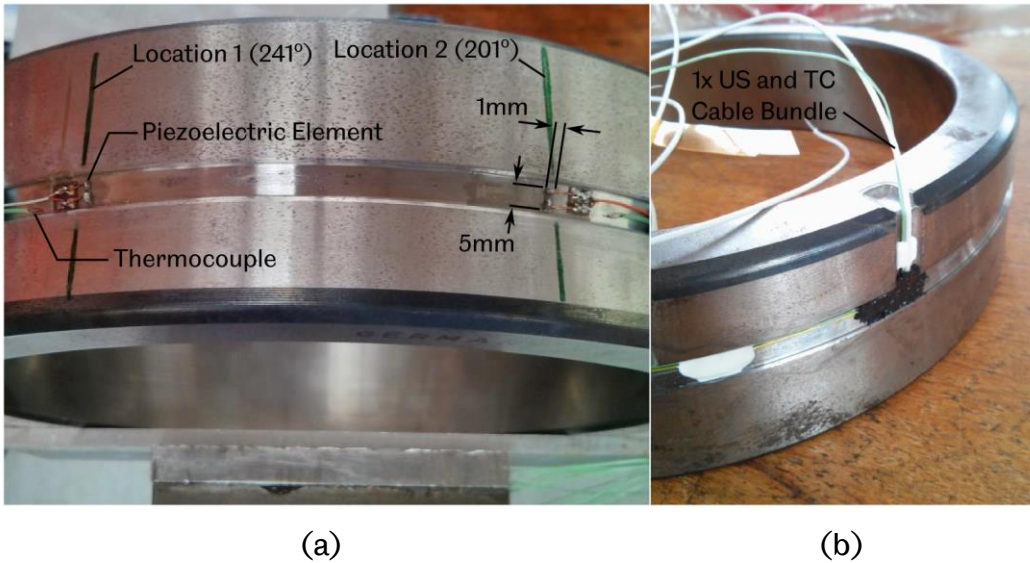
Figure 8.5: (a) Illustration of piezoelectric elements and thermocouples bonded to the HSS bearing outer raceway (b) Routing of signal cables axial to the bearing through additional recesses machined into the outer raceway

The acquisition of all signals from the HSS bearing sensors would not be performed solely by the University of Sheffield CMS2, since some of the measurement types detailed in Figure 8.4 had also been defined within the specification of other CMS devices. As such, alternative systems would already have in place the data acquisition equipment required for such measurements, and thus resources could be shared to avoid over specification of individual systems. Figure 8.6 summarises all HSS instrumentation and details the specific CMS device that would be responsible for each measurement type.
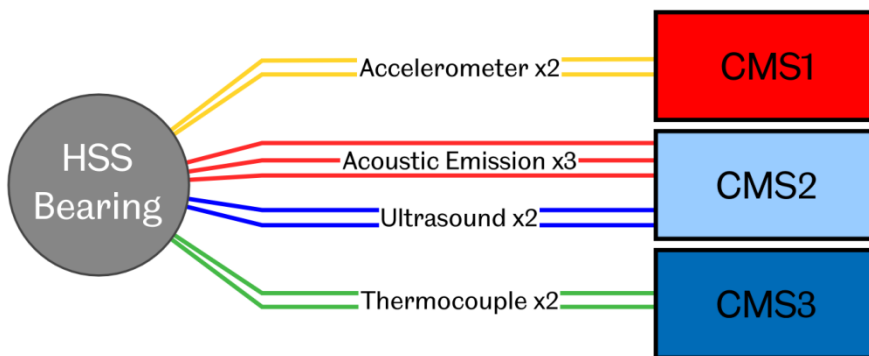


Figure 8.6: All measurement signals from the HSS bearing and the associated data acquisition device for each

As shown in the figure, measurements from the thermocouples shown in Figure 8.5 would not be directly logged by the CMS2 unit. As illustrated in

Figure 8.1 however all systems would have access to shared data storage at the turbine tower base, and thus the thermocouple data would be readily available as required.

## 8.2.2 Hardware Development

To facilitate ultrasonic measurements an FMS system (as described in §3.4.2) was required, which subsequently dictated that the CMS2 system would have to be PC based. The minimum footprint for the CMS would therefore be approximately $500 \times 430 \times 175mm$, as based on the outer dimensions of the PC. Consequently, the system would be reasonably large in relation to available space within the nacelle, and therefore the orientation, location and mounting of the system within the nacelle were key considerations at the early stages.
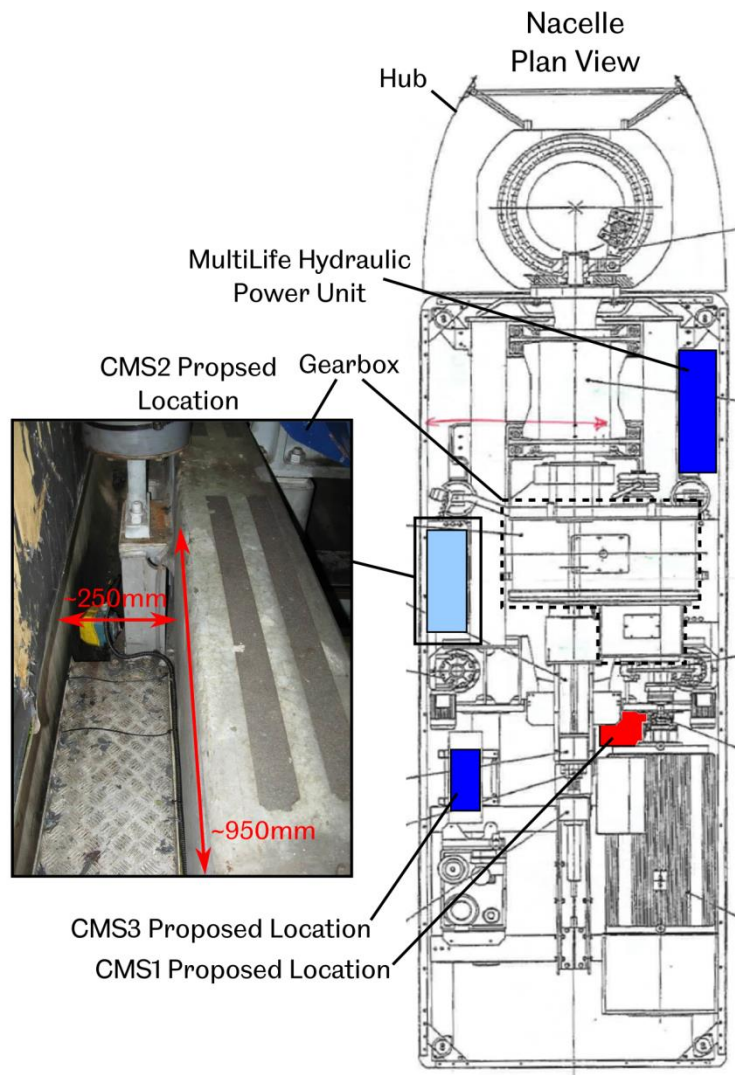


Figure 8.7: Proposed location of CMS2 and other key CMS devices within the nacelle

Assessment of available space within the nacelle was therefore made, considering the minimum dimensions of the system as defined above, and the requirement for its location to be as close as possible to the gearbox so as to minimise signal cable lengths. The most appropriate location was thus identified as shown in Figure 8.7, where it could be positioned on a checker plate panel in between a walkway and the nacelle wall to the left of the gearbox.

The system would therefore need to be installed in a vertical orientation so as to fit in the available space, and thus a method of mounting and securing the system in place was considered. Furthermore, the system would need to be "ruggedised" so that the base PC was appropriately protected from the nacelle environment. The following were identified as having the potential to cause damage to the system once installed.

- **Vibration:** Vibration levels within the nacelle were known to be high during turbine operation, which could cause damage to system components such as hard disk drives, and also result in loosening and disconnection of PCI DAQ cards from the motherboard.
- **Moisture:** The nacelle was known to not be completely watertight, and there was a risk that water may drip onto components from above.
- **Mechanical:** During installation of the CMS equipment, replacement of the gearbox, and the regular turbine servicing and inspections that were likely, it was inevitable that technical staff would often be working near to CMS equipment. There was a risk therefore that the system may sustain physical damage in such events, signal cables may be pulled and sensitive connectors could be damaged.

To mitigate such risks accordingly, the PC was packaged within a steel enclosure, as shown in Figure 8.8.
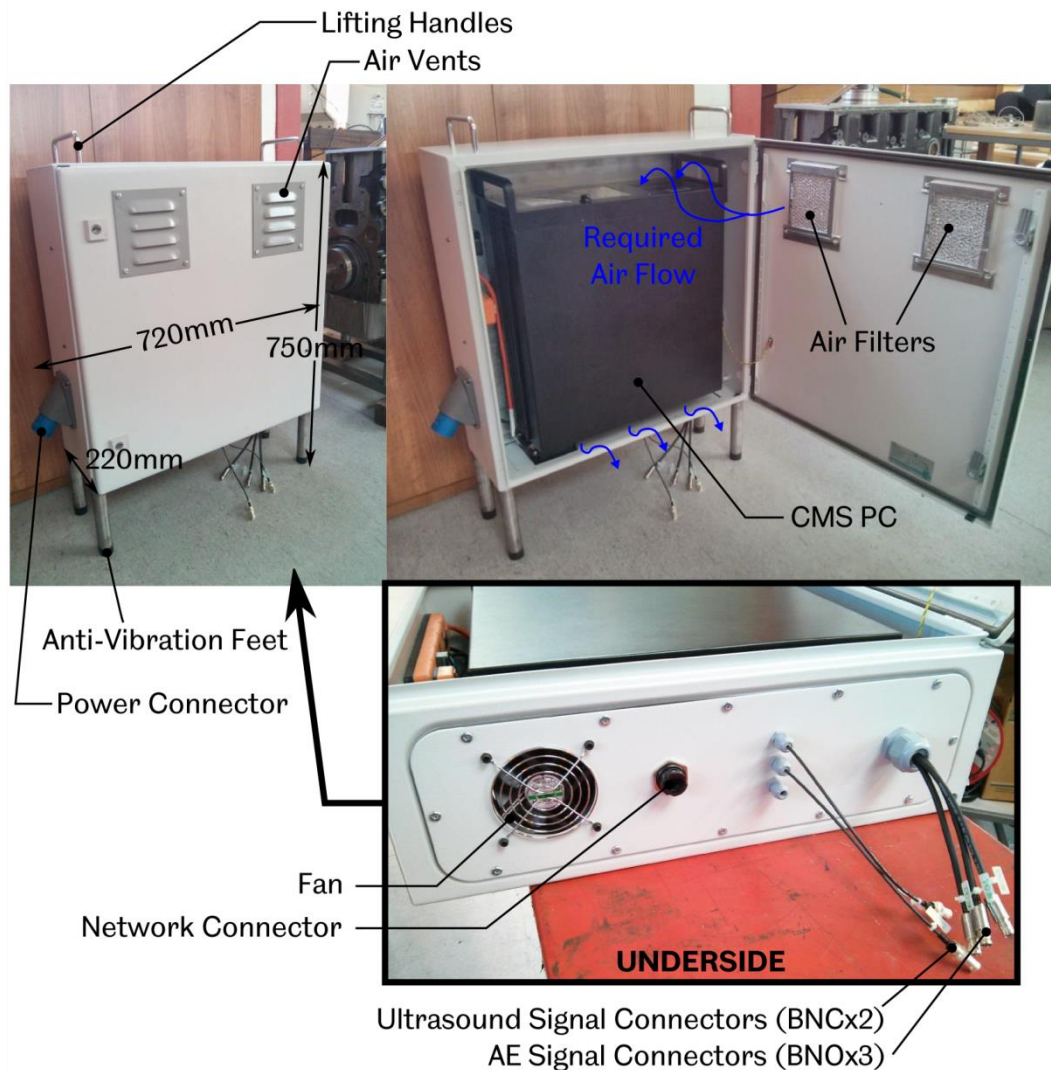
Figure 8.8: Final CMS2 system prior to installation

As shown in the figure, steel legs were installed onto the enclosure so that the unit could be installed onto the checker plate panel in its proposed nacelle location as shown in Figure 8.7. Anti-vibration feet were also installed to reduce the transmission of mechanical vibrations to the CMS system. Given the orientation of the PC within the cabinet, an internal bracket was installed (not visible in the figure) across the motherboard to secure all PCI cards in place and stop them from becoming disconnected. To further limit the potential damage to the system from vibration, the PC utilised a solid state hard drive (SSD), as conventional disk drives were known to often fail when subject to sustained vibration.

The cabinet provided an environmental seal to reduce exposure of the PC to moisture and water drips within the nacelle. A circulating air flow was still required by the PC however so as to avoid overheating, where the standard enclosure was designed to provide a flow of air through the

system in the direction illustrated in the figure. To maintain this, air vents were installed as shown, which were louvered so as to still restrict ingress of moisture from above. A fan on the underside of the unit was oriented so as to draw air out of the enclosure, thus maintaining the correct direction of air flow within the cabinet as required by the PC. Air filters were installed behind the air vents so as to limit the entrainment of particles into the enclosure and provide the PC with a clean air supply.

The sensitive signal connections on the underside of the PC system were protected through the use of jump leads linking the internal connections to more robust BNC and BNO connectors (for US and AE signals respectively) external to the cabinet. This also simplified installation within the nacelle, since connections could be made easily without having to open the cabinet and interact directly with the sensitive PC connections. Cable glands were used to provide strain relief for the jump leads, thus limiting the risk that cables would become disconnected inside the cabinet. The external BNC and BNO connections were selected due to their robustness and the bayonet type fastening system employed, which would also ensure that intermediate "external" connections could not be detached should either end of the cable be placed under tension.

### 8.2.3 Software Development

The full CMS suite would be installed within the turbine for 6-12 months, and therefore bespoke control software was required that was fully autonomous, requiring only minimal manual intervention. This section describes the key requirements and capabilities of the final developed software.
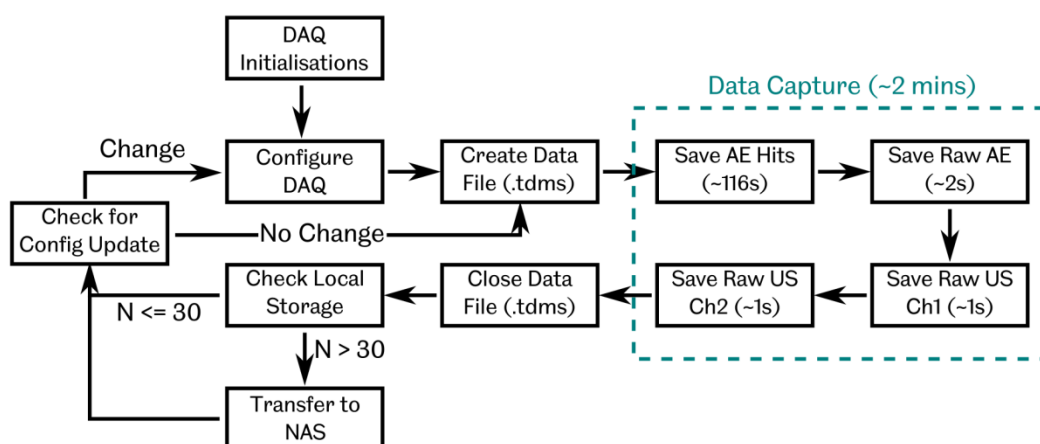


Figure 8.9: General operating sequence for the CMS

The controlling software for the CMS2 system was developed to maintain a sequence of operation as illustrated by the block diagram in Figure 8.9. All software was written in the LabVIEW environment due to hardware compatibility, and previous experience with this programming method. Software was configured to automatically run whenever the system was active and would subsequently maintain the operational sequence illustrated in Figure 8.9 until the software was manually stopped. The system was designed to run "headless," that is, with no interactive user interface, or indeed a requirement for any user input. Interaction with the system was however possible, if necessary, via a remote desktop connection achieved by connecting a computer to the internal LAN from the base of the tower. This was used to commission and debug the system initially and configure signal and data capture settings.

On start-up, a one-off initialisation would firstly be performed to activate the DAQ cards. The system would then be configured based upon settings defined in a configuration file stored locally on the system hard drive. This contained information relating to signal capture settings and data capture durations as detailed in Table 8.2.

| Ultrasound Settings | Acoustic Emissions Settings | Capture Settings |
|---|---|---|
| Excitation Voltage | Full Scale Range | AE raw data capture duration |
| Pulser Repetition Rate | Hit threshold | AE hit capture duration |
| Signal Range | Sampling Rate | US raw data capture duration |
| Gain | DC Offsets | |
| Pulse Width | Data capture segments | |
| Delay | | |
| Signal Filters | | |

Table 8.2: Configurable acquisition settings stored within the system configuration file

Following this, a data file would be created in a LabVIEW binary file format (.tdms) in preparation for the acquisition and storage of data. The file name would be created using the current date and time, and metadata would be written as a header to the file to describe all the current system settings as defined within the configuration file.

The system would then enter its main operational state and cycle through a sequence of capture regimes. It was identified that when applying the excitation voltage to the ultrasonic elements significant levels of noise would be observed on acoustic emissions measurements, and therefore the two different data sets could not be acquired simultaneously.

Primarily, it was desirable to capture data from both ultrasound and AE sensors in their raw digitised form, i.e. without any filtering or post-processing applied. However, given the high sample rates of each ($100MHz$ and $2MHz$ respectively); this had the potential to generate vast amounts of data extremely quickly. Therefore capture durations and the time delay between each capture instance were set so as to minimise data volumes but still acquire measurements regularly. Furthermore, to still provide insight into bearing operation in the time between raw data capture, AE "hit based" acquisition was performed in the interim; where only certain parameters of the raw data were stored.

The capture sequence was thus defined to occur over a total duration of approximately 2 minutes, where AE hit based capture would be performed for the most part, followed by a 2 second duration of both AE and Ultrasonic raw data capture. As illustrated in Figure 8.9, ultrasound data capture was set up to acquire on each channel sequentially rather than in "simultaneous" mode. This was implemented due to the requirement to acquire the data at the highest applicable pulse rate of $80kHz$. This was necessary given the high rotational speed of the bearing ($1500rpm$) and subsequently the speed at which each roller would move past the ultrasonic elements. High PRR settings were thus required to capture the passage of each roller with sufficient resolution and to ensure the contact centre point and thus minimum film thickness was observed. Given the fact that all FMS input/output channels were switched, as discussed in §3.4.2.1, the pulse rate would have been halved if acquisition was performed with both channels active, and thus independent acquisition was necessary.

Once completed, the system would close the current data file and then check the total number of files currently stored on the local hard drive. If this value exceeded a pre-defined threshold, equating to roughly 1 hours' worth of data, a transfer sequence would be triggered where all local files would be moved across the network to a NAS device at the turbine tower base. A secondary $2TB$ hard disk drive was also installed within the CMS2 PC to act as a temporary storage location, so that if no NAS drive could be found, data would be temporarily moved to this location to free up space on the system SSD.

It was possible to change the system settings defined in Table 8.2 by changing the relevant values in the system configuration file. As such, after performing local storage checks and transferring data if necessary, the system would assess the last modified date of the system configuration file. If this was found to have changed since it was last checked, the

system would carry out a reconfiguration of the data acquisition accordingly, prior to re-entering the data capture loop.

# 8.3 Gearbox Testing

The gearbox used for this work was an ex-service unit from a turbine within the same Barnesmore site as the turbine under investigation. This gearbox had previously been removed due to a fault and was subject to refurbishment and reassembly by *ZF Technical Services* in Nottingham, at which point it was possible for the MultiLife planetary bearings and HSS 32222 instrumented bearing to also be installed. Following refurbishment and prior to reinstallation of the gearbox within the turbine, stability tests would be performed on a test platform, also housed within the ZF facility, to ensure that the gearbox was fit to be reintroduced into service. This provided a good opportunity to set up the full CMS suite to simulate the turbine setup and confirm that all software was functioning as intended. In addition, this also provided the opportunity to check all HSS bearing instrumentation to ensure that no damage had occurred during the installation process, signals were received with reasonable clarity, and that responses were appropriate based on gearbox test parameters.

## 8.3.1 Test Conditions

The test platform utilised a motor on both the input and output side of the gearbox so as to vary both the load and speed applied to it. The input-side motor permitted control of gearbox input speed up to its $30rpm$ rating, whilst the output-side motor could be used to apply a resistive torque to simulate that which would be induced by the generator during turbine power production. The amount of resistive force that could be applied was however limited, equating to gearbox load of around $63kW$, thus just over 10% of the maximum gearbox power rating. Figure 8.10 illustrates the operating conditions under which the gearbox was placed during a 5 hour endurance test, throughout which signals from HSS instrumentation were acquired continuously. Gearbox speed is quoted in terms of the HSS output in the figure where, due to a $50:1$ gearbox ratio, a $30rpm$ input speed would equate to a $1500rpm$ output.
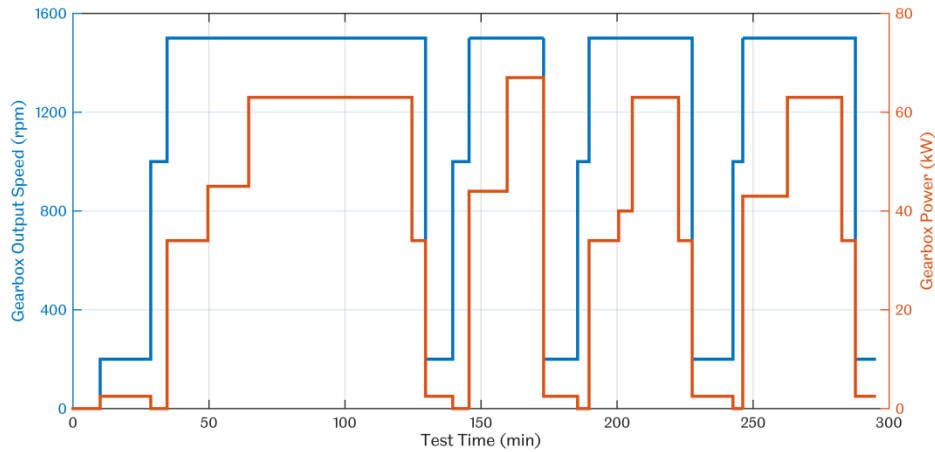
Figure 8.10: Gearbox endurance test conditions

The test sequence was defined largely for testing of the MultiLife systems, which were required to actuate only when gearbox speed was below $200rpm$, and input torque (measured via load pins on the gearbox suspension mounts) was below a level equivalent to a power of $2kW$. The test sequence could therefore be utilised to check that the control software would correctly actuate all MultiLife systems only during periods when such conditions were met. At all other times it was essential to ensure that the MultiLife planetary bearing inner raceways were suitably restricted from any further, undesirable indexation. The remainder of the test sequence was therefore defined to place the bearings under a range of speed and load conditions to ensure that this was always the case. From an instrumentation standpoint however, this test was extremely useful in providing a range of measurement conditions over which data could be obtained to identify associated changes in ultrasonic sensor signals, as will be discussed throughout the remainder of this section.

## 8.3.2 Results

### 8.3.2.1 Initial Observations

Cable pathways within the nacelle were still unknown at this stage, and therefore cable lengths were specified at the absolute maximum that may have been required within the nacelle, thus representing the worst-case scenario throughout the gearbox endurance testing. At such lengths signal quality was found to be degraded somewhat, and significant excitation voltage had to be applied to the ultrasonic elements (in the order of $50V$)

198

in order to obtain a reasonable response in terms of the peak-to-peak amplitude of the obtained raceway reflections.

Figure 8.11(a) and (b) illustrate the modal reference waveform signals obtained from both ultrasound sensors at around 60 minutes into the gearbox endurance test. Here it is clear that time domain reflection pulses did not possess a clear bell-shaped envelope as observed from typical ultrasonic reflection signals presented previously within this work (Figure 7.5 for example). Similarly, when observing the frequency content of each reference waveform, the resonance characteristics were also found to be unusual (again, by comparison to that exemplified in the aforementioned figure) where, rather than a bell shaped curve with a single peak around $10MHz$, as would be expected for the piezoelectric elements utilised, two peaks were apparent at $\sim 5MHz$ and $\sim 14MHz$, but with no such resonance observed at the expected $10MHz$ region. This was therefore attributed to the capacitive nature of the coaxial signal cables, which were found to filter certain frequency components of the transmitted signal when long cable lengths were utilised.
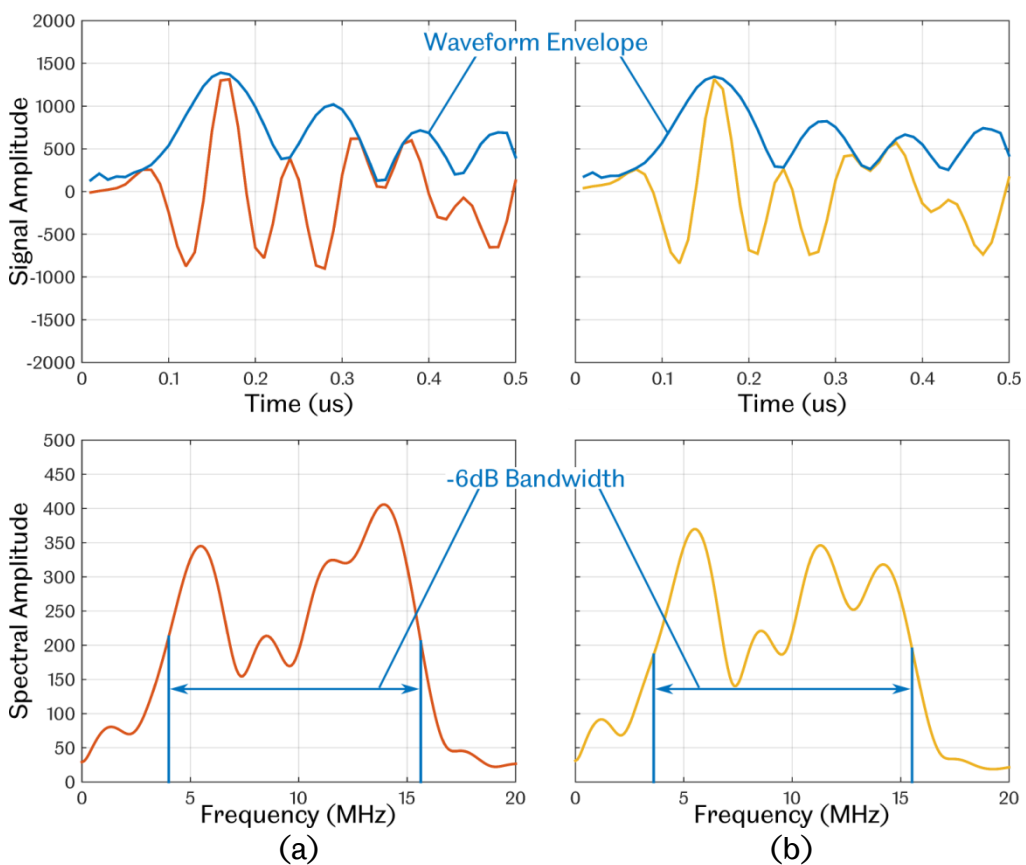


Figure 8.11: Reference signals obtained from HSS ultrasound sensor locations (a) 1 and (b) 2 in both the time (top) and frequency domain (bottom), obtained at $t = 60min$ into the gearbox endurance test

### 8.3.2.2 Determining the Metrics for Condition Monitoring

In the analysis of data obtained from the test platform in Chapter 6 it was identified that only very thin oil films would form between the contacts of the NU2244 test bearings due to high bearing loads and low viscosity lubricating oils that were used. A boundary lubricated contact would thus be observed by the ultrasonic sensor, the spring model would therefore not be applicable to the contact under observation, and a meaningful film thickness result could not be obtained from the measurement data. For this reason, the ultrasonic reflection coefficient, $|R|$, was utilised as the metric for characterising bearing condition in Chapters 6 and 7. The gearbox bearing under observation here however would likely observe much thicker oil films, owing to the lubricant having ten times the base viscosity rating as that utilised on the test platform, in addition to the fact that bearing loading would be substantially lower. Indeed, as established in §8.1.2, at gearbox rated power ($PWR_{max} = 600kW$), the expected equivalent load applied to the 32222 bearing would equate to $P_{max} = 45.8kN$, representing only around 11% of its dynamic load rating (loads applied on the test platform on the other hand were equivalent to $\sim 0.75C$). A reasonably thick oil layer may therefore form at the ultrasonic measurement points, and the contact may subsequently be more closely aligned to an ideal spring model contact, thus potentially allowing a representative film thickness to be determined from ultrasonic measurements.

Theoretical film thickness formation within the 32222 bearing contacts could be approximated through use of Equation (6.3), and would vary throughout the endurance test according to changes in both gearbox speed and power, as a result of the subsequent alteration to contact loading, $Q$, and oil entrainment velocity, $\bar{U}$.

Loads at each sensor location, $Q_{US1/2}$, were established by firstly estimating the equivalent load on the bearing under each gearbox power level tested during the endurance test, $PWR_t$.

$$P_t = P_{max} \left( \frac{PWR_t}{PWR_{max}} \right)$$

(8.2)

Through this, the internal bearing load at each ultrasound sensor location could be estimated according to Equations (8.1), and subsequently the theoretical film thickness, $\bar{h}$, at the outer race rolling element interface determined.

The gearbox was lubricated with a VG320 gear oil, and calculations were therefore based on oil properties as defined in Table 8.3, where oil temperature was assumed to be approximately $50°C$ throughout the test, based on measurements obtained from the HSS bearing thermocouples at the time.

| | |
|---|---|
| Viscosity at $20°C$ | $335cSt$ |
| Viscosity at $100°C$ | $38.3cSt$ |
| Density at $15.6°C$ | $860kg/m^3$ |
| Pressure-Viscosity Coefficient, $\alpha$ | $1.796 \times 10^{-8}Pa^{-1}$ |

Table 8.3: Viscosity and density reference data for the VG320 gearbox lubricant

In determining the reduced radius of the contact, given the tapered geometries of both the roller and outer raceway, the mean radius of each surface was utilised, thus effectively representing the contact geometry at the centre of the rolling surface.

Film thickness curves could thus be obtained, as shown in Figure 8.12, based on the speed and power curves in Figure 8.10, and the estimated normal contact load at each sensor location, $Q_{US1/2}$.
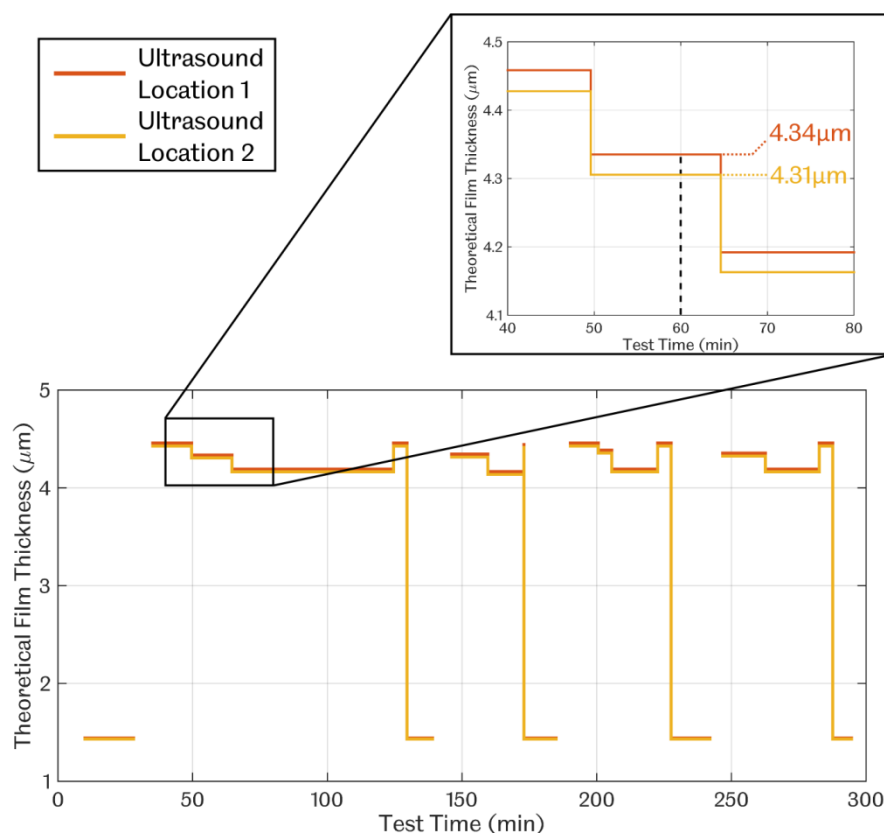


Figure 8.12: Theoretical film thickness formation within the 32222 HSS bearing based on gearbox power and speed throughout the endurance test as defined in Figure 8.10

Particular focus is made in the figure to the theoretical film thickness results obtained at a test time of 60 minutes, so as to be comparable to the ultrasonic signals presented in Figure 8.11. As shown in Figure 8.12, at $t = 60min$ test conditions were such that film thicknesses of $4.34\mu m$ and $4.31\mu m$ were expected at sensor locations 1 and 2 respectively. Ultrasonic data was thus examined further to identify the measured film thickness at the same test time so a comparison could be made to the above theoretical expectations.

Figure 8.13 shows reflection coefficient and subsequent film thickness vs frequency results from a single roller pass measured within the same data stream as that used to obtain the modal reference waveforms of Figure 8.11. Results are shown only for frequencies lying within the bandwidth of each reference spectrum, as again shown in Figure 8.10. It should be noted that – in calculation of film thickness – the sound velocity through the lubricant and its variation with temperature were unknown and
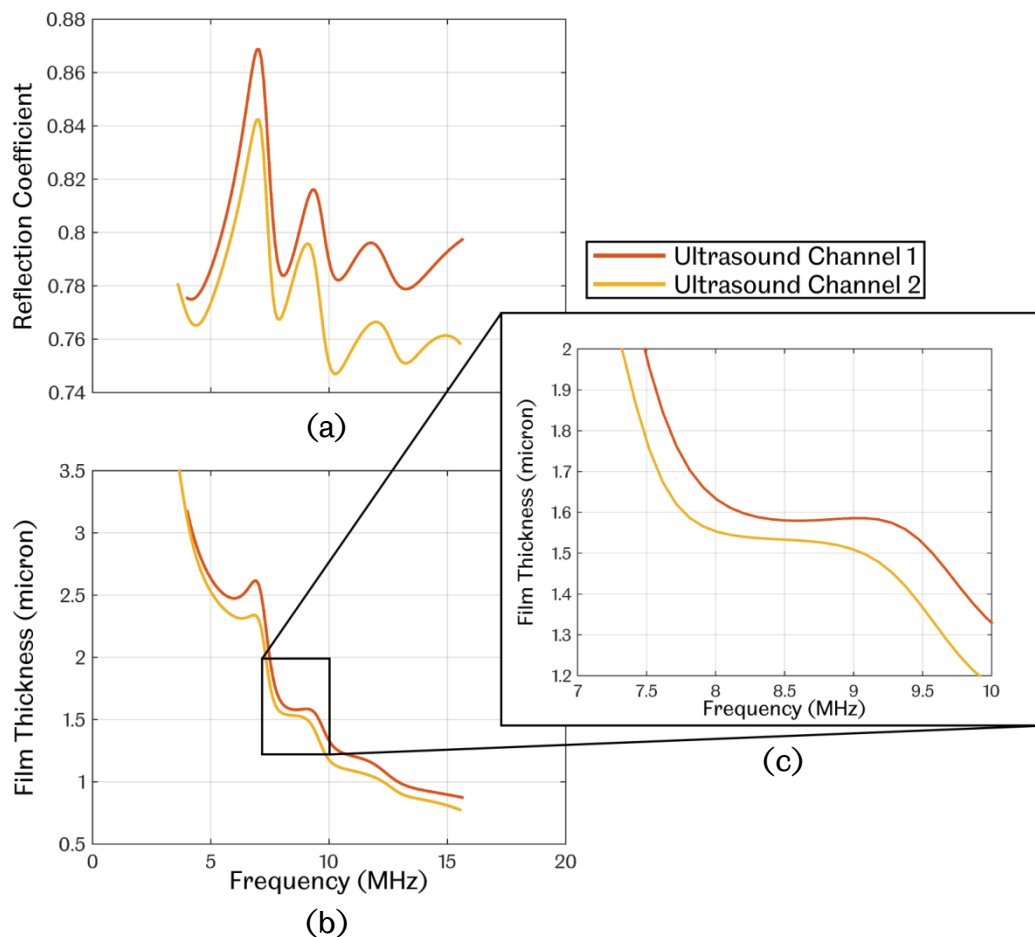


Figure 8.13: (a) Reflection Coefficient vs Frequency within the bandwidth of the two ultrasonic sensors (as defined in Figure 8.11) and (b) subsequent film thickness based on Equation (3.16). (c) shows the frequencies over which a constant film thickness output is achieved

therefore the relationship defined in Equation (6.6) was utilised, which was considered a reasonable approximation.

As shown in Figure 8.13(b), film thickness varies considerably with respect to frequency across the majority of the bandwidth of the sensor, thus making the obtained results questionable. Curves do appear to remain constant over an $8 - 9MHz$ frequency range, yet this corresponds to oil layers in the order of $1.55 - 1.6\mu m$, thus equating to around 35% of the theoretical expectations defined above. Similar film thickness variations with frequency were highlighted in Figure 6.13 with regards to measurements made on the test platform. Here, it was found that the observed film thickness variation with frequency would occur to a greater extent at lower load and, in combination with similar observations made in [80], this was concluded to be a consequence of the width of the ultrasonic beam being larger than the area of contact under observation. It would be reasonable to assume that this is also the case for the current measurements, given that the ultrasonic source element width is only 1/3$^{rd}$ smaller than that implemented on the experimental rig, yet a much smaller bearing is utilised here with a much lower load applied to the bearing contacts.

It is concluded therefore that — similarly to the earlier laboratory based work — application of film thickness models cannot be applied to this measurement scenario, and the use of reflection coefficient is again justified as the most appropriate condition monitoring metric to extract so that long term trends in bearing behaviour can be observed. Referring back to the spectral reference signals presented in Figure 8.11, it is assumed that the peaks observed at 5.5 and $14MHz$ are spurious, and appear to be a function of the use of long cable lengths. A lower amplitude peak at $8.5MHz$ is evident within the spectral response of both ultrasonic sensors, which is therefore considered the most appropriate resonance on which to base all further analysis of ultrasonic data obtained throughout the entire endurance test.

### 8.3.2.3 Ultrasonic Trends across the full Endurance Test

As discussed earlier, ultrasonic data capture was performed at 2 minute intervals and for a 1 second duration per channel as configured within the CMS2 software. An automated data analysis process was thus required to analyse all data obtained throughout the test since the volume of data amassed over the 5 hours was significantly large ($\sim4GB$). The feature extraction process is illustrated in Figure 8.14(b) (with reference to the

raw data stream shown in Figure 8.14(a)) and was similar to that used in §7.2.2.2, where the minimum reflection coefficient per roller pass (at $8.5MHz$ in this case) was detected and from this the average $R_{min}$ per data stream was identified.

For this case it was also of interest to note the time between each sequential roller passage within every data stream, so that roller pass frequency could then also be determined, thus permitting further calculation of bearing speed through rearrangement of Equation (2.18) (considering $BPF_O$ as the relevant bearing frequency in this case). Autocorrelation was therefore subsequently performed on the reflection coefficient curves, to obtain the periodicity and thus establish the relative time spacing of each roller pass, as shown in Figure 8.14(c).



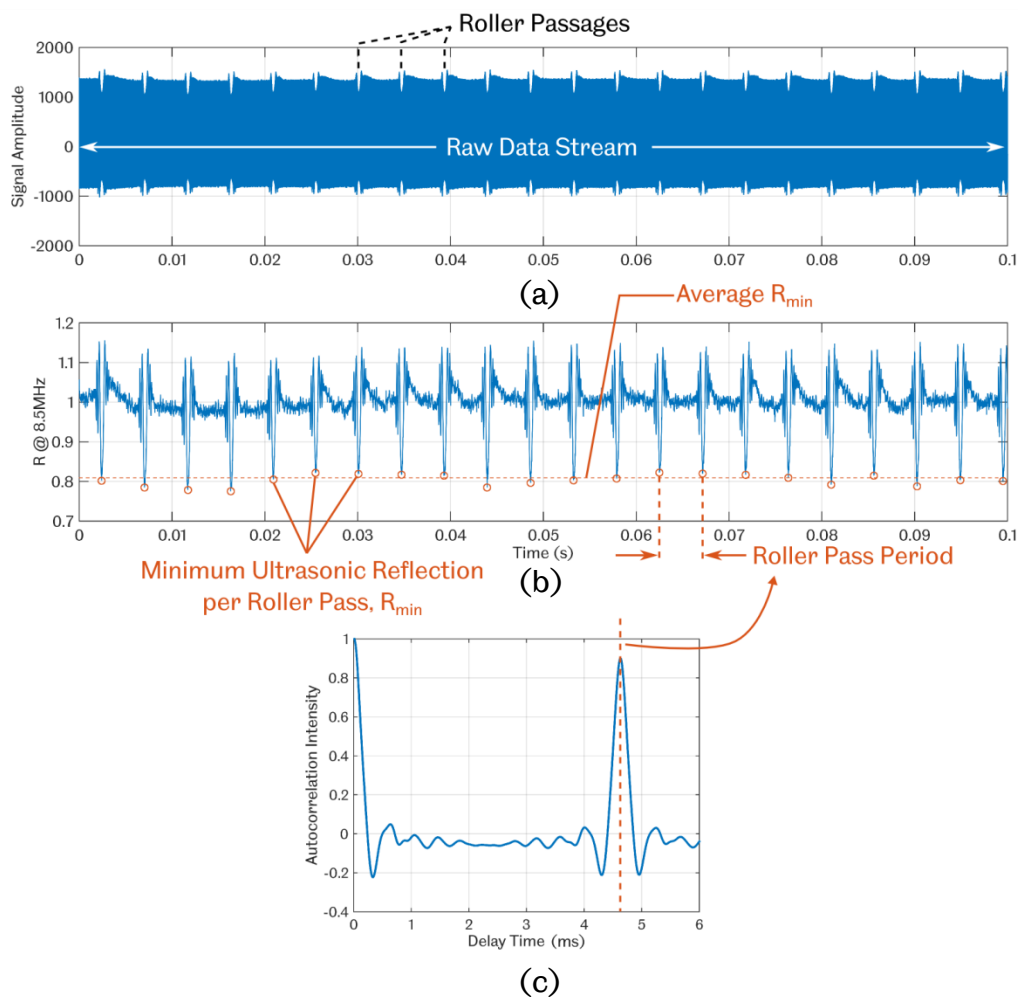Figure 8.14: (a) Segment of raw ultrasonic data obtained from Sensor 1 during gearbox operation (b) refection coefficient calculated from the raw data using the $8.5MHz$ frequency component (c) Autocorrelation of the reflection coefficient curve to obtain the roller pass period

Figure 8.15 plots the inferred HSS bearing speed based on the periodicity of the reflection coefficient vs time curves, as discussed above. Given that no deviation was expected to exist between both ultrasound data channels with regards to the observed roller pass frequency, both data vectors have been merged to form the single red line in the figure. Overlaid is the actual test speed curve of Figure 8.10. It should be noted that the data used to generate the curve of Figure 8.10 was obtained from written notation of the associated time at which each change in a test parameter was made. Subsequently, ultrasonic measurements and the overlaid test parameters are not exactly synchronised due to inevitable errors induced in the manual recording of the latter.
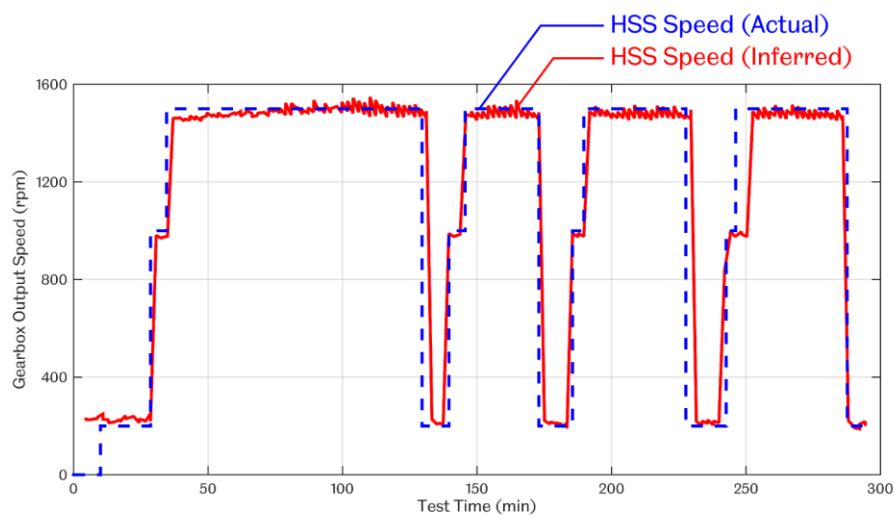
Figure 8.15: Comparison between actual gearbox output speed and that inferred from the roller pass frequency measured from ultrasound data

In general, inferred bearing speed can be seen to agree well with the known test set point throughout the plot, excluding the slight phasing errors observed in places, as discussed above. The curve obtained from ultrasonic measurements does tend to become noisier at higher speed, which is attributed to known issues with ultrasonic data capture at the high pulse repetition rates utilised ($80kHz$). This is known to occasionally result in some slight loss of data, thus causing discontinuities to occasionally arise within the acquired data stream, which then lead to irregularities when the roller pass period is calculated. This has a much lesser impact at lower speed, where time spacing between roller passages is larger, and the associated error in determining the roller pass period becomes less significant.

Figure 8.16 plots the average ultrasonic reflection recorded per roller pass from each sensor with the gearbox test parameters of Figure 8.10 overlaid.

Throughout the duration of the endurance test a greater ultrasonic reflection is measured at sensor location 1, which follows with the lower loading that is expected at this position, as illustrated in Figure 8.3. Correlation between the changes in gearbox operating conditions and measured reflection coefficient are particularly clear within the low power and speed regions, where a large drop in reflection can be observed in both sensor traces as a result of the significant drop in film thickness that will occur over the same regions, as illustrated in Figure 8.12.
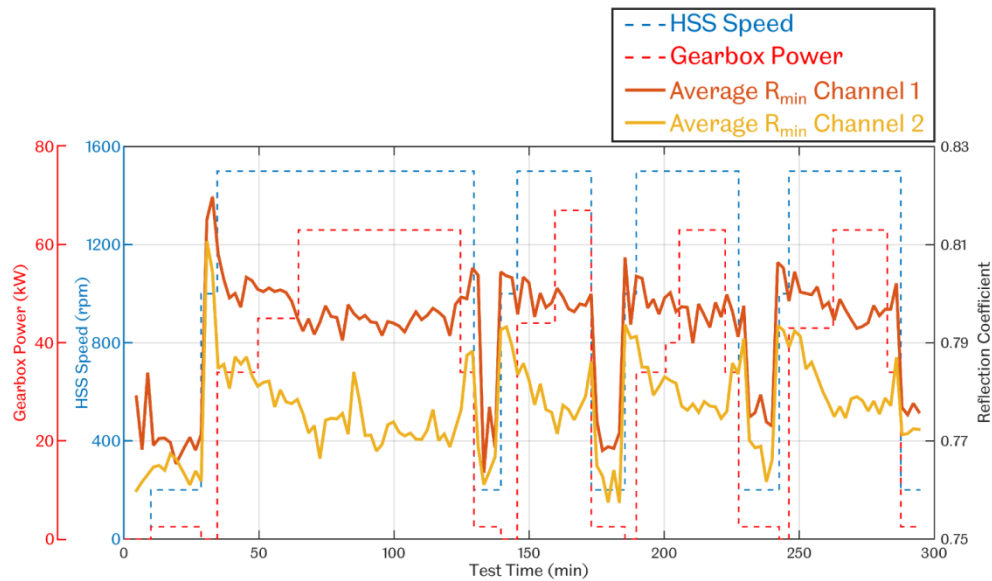


Figure 8.16: Comparison of the average ultrasonic reflection per roller pass per channel against gearbox power and speed

During the high power/speed stages of the test, reflection curves can both be seen to generally reduce as power is incrementally increased, which follows with the reduction in film thickness that is expected. However, such changes are not instantaneous, and do not reduce in a stepwise manner according to the changes in gearbox operating conditions. It is therefore likely that that the observed trends are a result of more progressive phenomena occurring within the bearing contacts, rather than as a direct function of the applied bearing load. Such observations may be a result of the general inertia of the gearbox system, or a gradual rise in temperature at each test set point. The latter, for instance, would cause a gradual reduction in film thickness at the bearing contacts and an increase in the relative approach of the surfaces, thus reducing the reflection of acoustic energy at the measurement interfaces.

A spike in reflection is observed at the end of each high loading cycle, as rotational speed is maintained but power is reduced, which also agrees

well with the expected increase in film thickness during such instances, as identified in Figure 8.12.

Overall, the ultrasonic response appears to be repeatable, where similar reflection trends are observed across comparable operational zones. Signals appear to be more responsive to changes in gearbox speed, than power. This is likely due to the limited range of power levels that could be tested, which would subsequently only induce very slight changes to the loading of the 32222 bearing. Such changes appear to be too slight to directly and immediately induce responses in ultrasonic signals, thus perhaps illustrating a limitation in the detection capabilities of the sensors at such low gearbox power ratings. It is therefore anticipated that further variation in ultrasonic reflection would be observed once the gearbox is in-service, due to the higher magnitude and more widely varying loads that will inevitably be induced. The detection of slight changes such as that highlighted in the previous chapter as a result of rolling surface failure may therefore be difficult to detect given the high level of variability that is expected to exist during general operation.

## 8.4 Installation Testing

Final commissioning of the systems within the turbine took place in May 2015. Figure 8.17 shows the CMS2 unit installed within the nacelle in its proposed position next to the gearbox, and fixed down to the checker plate panel with ratchet straps.



Figure 8.17: CMS2 unit installed within the nacelle

### 8.4.1 Test Conditions

Following installation it was not possible to return the turbine to service immediately due to additional repair work that needed to be performed before power production could be resumed. It was however possible to perform the necessary checks to ensure that all software was functioning as intended, and that all signals from HSS instrumentation could be received.

Wind speeds were low at the time of commissioning and so most data obtained from the HSS instrumentation was recorded when the gearbox was static. At brief periods wind speeds were sufficient to induce slight rotation of the turbine rotor at a rate of approximately $0.5 - 1rpm$, and a portion of data obtained during this period is presented in the following section. However, given that no power was being produced, the generator was idle, and therefore loading through the gearbox was negligible.

### 8.4.2 Results

#### 8.4.2.1 Initial Observations

During the installation of the CMS unit it was found that the cable lengths utilised within the endurance test were in fact a significant over estimate of what was actually required, and a reduction by almost half of the original estimate was possible (to around $3m$). This was found to have a profound effect upon ultrasonic signal clarity, as illustrated in Figure 8.18.

As can be seen from the figure, by comparison to Figure 8.11, the time domain reference signals are comprised of a much more stable series of oscillations, and waveform envelopes are much smoother, with a more typical bell-shaped profile. Frequency content is also contained within a much narrower bandwidth, and a single resonance is once again observed, with a peak lying around $11MHz$. This frequency component was therefore utilised in determination of reflection coefficient from the obtained dataset.
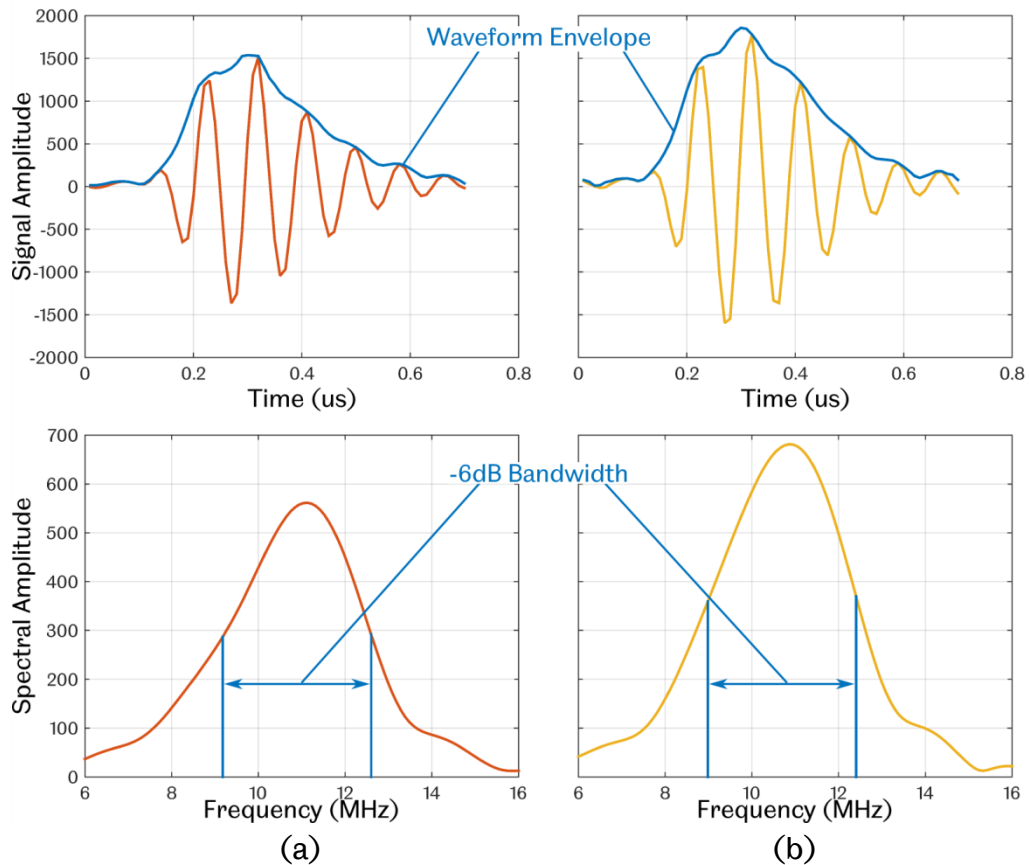
Figure 8.18: Ultrasonic reflection signals obtained from (a) ultrasound sensor 1 and (b) ultrasound sensor 2 in both the time (top) and frequency domain (bottom) following final installation within the wind turbine nacelle

### 8.4.2.2  Reflection Coefficient Measurements

Figure 8.19 illustrates the ultrasonic data obtained from both HSS sensors during the brief period when the turbine was rotating. 4-5 roller passes were observed over the 1 second capture window, equating to a gearbox output speed, at the time, of approximately $30rpm$. It should be noted that, due to the sequential capture of both ultrasonic channels, the time axis on the two plots is not synchronised.

As shown in the figure, the average MRC per roller pass was measured at around 0.8 from both sensor locations, which is somewhat larger than the values obtained under similar conditions within the gearbox endurance test $(R \approx 0.76 - 0.77)$. However, given that the turbine was idle, the gearbox lubricant system was not active and therefore sensors would likely be observing a boundary lubricated contact, where only residual oil on bearing surfaces would become entrained within the contact. Increased asperity interactions would therefore be likely, with an

increased likelihood that air voids would exist across the interface, thus inducing a higher degree of ultrasonic reflection.
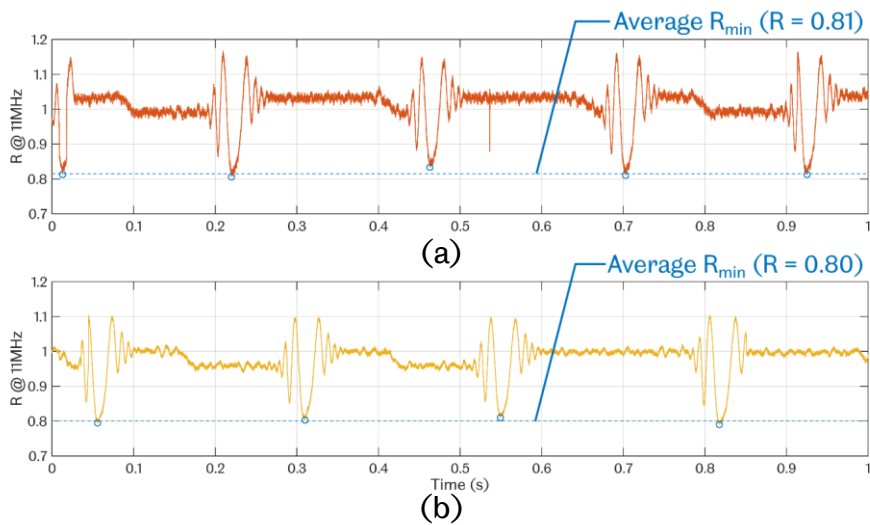


Figure 8.19: Reflection coefficient vs Time (at the $11MHz$ resonant frequency as identified in Figure 8.18) from a 2 second data stream obtained from (a) ultrasound sensor 1 and (b) ultrasound sensor 2

# 8.5  Discussion

In the previous chapter it was identified that ultrasonic measurements can indeed provide indication of a bearing rolling surface failure through a general reduction in the measured minimum reflection coefficient. The same measurement technique has successfully been implemented here within a WTGB bearing, potentially allowing similar monitoring techniques to be performed.

The work presented in the previous chapter however was performed within a controlled environment, where bearing load and speed were generally constant and directly measurable throughout. The observed trends pertaining to bearing degradation might therefore prove difficult to detect in the wind turbine gearbox due to the wide variation in operating conditions that are likely given the extremely transient nature of the wind. While attempts have been made to characterise the ultrasonic response under a range of gearbox operational regimes, the full range of loading conditions have so far not been observed, and therefore full knowledge of the normal signal variations that will occur in-service are unknown. Certain operational changes investigated within the gearbox endurance test were found to induce only gradual changes in the ultrasonic reflection

response, and therefore a reduction in the reflection coefficient that exceeds the typical rate of change may provide an indication of unusual operation. Alternatively however, the endurance test was defined with a degree of repetition, and many of the test set points were therefore observed multiple times. From this it was identified that the ultrasonic response was generally repeatable across comparable test regions. Similar techniques to those applied to SCADA data as discussed in §2.4 ([62, 64, 65] were provided as examples) may therefore be applicable, and through an initial learning phase in the early stages of turbine operation it may be possible to characterise the ultrasonic response across the full gearbox operational range. Any subsequent deviations from the expected response could therefore be used as an indicator of bearing failure.

Above all, it is clear that any detection algorithms that are developed based on this work will need to consider as many additional wind turbine operational parameters as possible, and therefore a vast array of additional measurements will be necessary in order to detect bearing failure robustly. Considering the full CMS suite that has been developed within the OWDIn consortium, a multitude of system-wide measurements are indeed available, thus the opportunity for robust algorithms to be developed as part of this work is high.

## 8.6 Conclusions

- A condition monitoring system was developed in collaboration with Ricardo UK Ltd and the University of Strathclyde to be installed within a $600kW$ Wind Turbine on Barnesmore Wind Farm in the Republic of Ireland.
- A high speed shaft gearbox bearing (known to exhibit poor reliability) was instrumented with ultrasonic sensors to monitor oil film formation within the bearing. Acoustic emissions sensors, as well as more conventional temperature and vibration sensors, were also installed upon the same bearing for comparison of the more experimental condition monitoring techniques to those more widely utilised.
- The condition monitoring system and instrumentation were developed to maximise robustness so as to mitigate the risks of damage in a harsh wind turbine nacelle environment.
- Bespoke software was developed in LabVIEW to manage the acquisition of data and operate autonomously.

- Early bench testing of the gearbox was performed prior to installation within the turbine whilst data was acquired from ultrasonic instrumentation.

- Similarly to the experimental application presented in the previous chapters, the installed sensors were found to be inappropriate for the determination of oil film thickness.  Measurement of the minimum ultrasonic reflection however was found to still be appropriate as a means to assess bearing condition, as applied in Chapter 7.

- Signals were found to respond as expected to the various test loads placed upon the gearbox, and appeared to generally be repeatable.

- The extraction of ultrasonic reflection trends pertaining to bearing failure will be significantly more challenging to achieve in this case – by comparison to that demonstrated in Chapter 7 – due to the highly variable bearing operating conditions that will be observed in-service. More complex algorithms will thus be required, drawing upon a variety of turbine measurements in order to detect such features robustly.

- The system has since been successfully commissioned within the turbine but, due to the requirement for additional repair work, the turbine has yet to be returned to service.

# Chapter 9

# CONCLUSIONS

Poor reliability of wind turbine gearboxes can very often be linked to the failure of rolling element bearings, where those supporting the high and intermediate speed shafts and those which support the planetary gears have been found to be particularly failure prone. The latter operate via outer raceway rotation, and loading is generally unidirectional, which subsequently results in a localised segment of the stationary inner raceway experiencing high cyclic stress throughout the entirety of its operational life. Overloading of the bearings will also occur due to the transient nature of the wind, and subsequently sub-surface cracks and stress concentrations are induced which weaken the material. The combination of these two factors means that failures tend to propagate rapidly within planetary support bearings via spalling of the localised segment of the inner raceway rolling surface that is placed under load.

A method to improve the reliability of such bearings has been proposed by Ricardo UK Ltd, whereby the nominally stationary inner raceway is periodically rotated or "indexed." As such, material fatigue can be distributed around the entire circumference of the race, thus delaying the onset of failure. This concept has been termed the "MultiLife$^{TM}$" system. This thesis has focussed upon the development and validation of the MultiLife concept through bearing life testing on a bespoke experimental platform. Simultaneously, the capability of the ultrasonic technique has been explored as a means to detect the onset of failure within such bearings and thus support the application of the MultiLife system.

## 9.1  The Detection of Bearing Failures

Various techniques have been developed over the years to assess the condition of rolling element bearings without the need for machine disassembly and direct inspection. Oil analysis, temperature and vibration monitoring have been identified as the most common, and are employed widely in a variety of applications. Acoustic emissions has also been identified as a further method to detect bearing failure, but this has not yet seen as wide spread adoption. The technique however shows potential in providing early warning of incipient bearing failure through its apparent capability to detect sub-surface crack initiation and propagation.

### Application to Wind Turbine Gearboxes

Wind turbines generally only employ very basic monitoring techniques as standard, which do not extend to the monitoring of specific bearings within the gearbox. This is due to the additional added complexity of instrumenting specific gearbox bearings, and the modifications to gearbox architecture that would be required. All of this adds cost to the wind turbine system, thus resulting in limited uptake of such advanced CM systems to date. Ultimately however, the advanced warning of bearing failure that may be provided by such systems could potentially enable wind farm operators to schedule maintenance more effectively, which may act to reduce costs and downtime in the long-term.

### The use of Ultrasound for Bearing Condition Monitoring

Previous researchers have demonstrated the capability of ultrasound in determining the thickness of the oil films that form within bearing contacts. Given the fundamental link between bearing lubrication and operational performance it was identified that the monitoring of long term ultrasonic trends obtained from a bearing could be used to indicate lubricant breakdown, thus potentially providing an additional means to detect an incipient bearing failure.

## 9.2  Development of a Bearing Life Test

An experiment was developed to enable life-to-failure testing of NU2244 cylindrical roller bearings; chosen to be representative of those utilised within the planetary stage of a $2-3MW$ WTGB. Test conditions were identified, through utilisation of the ISO bearing life rating procedures, so as to accelerate the fatigue life of the test bearings. In addition, and in an

attempt to further simulate the in-field failure mode, a means of simulating a transient impact event was identified and implemented. This was achieved through the introduction of geometric discontinuities to the inner raceway rolling surface, implemented via a wire cut EDM process, and employed to act as a stress raiser from which a representative spalling failure would propagate.

### Life Predictions for a MultiLife Bearing

A bearing life rating model was developed, based on the ISO bearing life rating principles, but considering a bearing elementally, such that the relative consumption of life at localised raceway regions could be assessed. This was based on the NU2244 bearings to be investigated on the test platform and considered the test parameters identified for accelerated life. Various strategies of indexation were analysed and it was identified that up to a five-fold increase in bearing life was possible through raceway indexation, where frequent raceway indexation at a low angle was found to be the most advantageous approach to ensure high life.

The model was then further developed to consider the effects to bearing life when the proposed seeded defects are introduced. This was achieved through finite element modelling and assessment of the stress concentrating effect of the geometric discontinuity. As such, the expected bearing life on the test platform could be more accurately predicted, but more importantly; the performance of the MultiLife system could be assessed for cases that are more representative of the WTGB planetary bearings, where failures are understood to not arise purely as a result of rolling contact fatigue. As expected, relative bearing life enhancements were lower in the presence of stress raising features. It was further identified however, that much more care must be taken to avoid the relocation of "defected" raceway regions; thus highlighting the importance of detecting the occurrence of such faults in the field so as to maximise effectiveness of the MultiLife system.

## 9.3  MultiLife Validation

Evidence was obtained to indicate that the MultiLife concept was indeed valid in extending the life of wind turbine planetary support bearings, or indeed any bearing whose inner raceway is stationary, and where load is applied uni-directionally. This was demonstrated through the performance of two bearing life tests. In the first, a failure was induced at the point of maximum bearing load following the seeding of a surface

defect. This was then indexed by 180° and no further wear progression was observed. In the second test, multiple stress raising sites were introduced and damage was found to have progressed only from the site located at maximum bearing load. Indeed, the defect located out of the loaded zone showed no progression, whilst another defect located within the loaded zone, but away from the peak internal bearing load showed slight growth, but no evidence of spalling failure, thus illustrating a reduced rate of progression here.

# 9.4 Ultrasonic Monitoring

A means of ultrasonically monitoring the NU2244 bearings throughout the accelerated life testing was identified. Ultrasonic elements were bonded to the test bearing inner raceway to monitor the contact at the point of maximum bearing load; that being the point at which failures were expected to emanate.

A novel scanning method was utilised to map the ultrasonic wave-front across the rolling surface and thus identify the divergence of ultrasonic energy through the bearing raceway. Through this it was identified that the installed ultrasonic element could suitably monitor the rolling contact at bearing loads above $300kN$, this being the threshold at which the ultrasonic beam would be narrower than the Hertzian contact width. Ultimately however, it was identified that the determination of film thickness from the ultrasonic measurements was not possible due to the high contact loads and low viscosity lubricant utilised, thus resulting in a boundary lubricated contact. It was concluded therefore, that the ultrasonic reflection coefficient would be the most appropriate metric to extract for the purposes of condition monitoring.

**General Observations**

Ultrasound data was initially discussed in terms of the general features that are frequently observed within bearing applications, but that are not yet particularly well understood. The "fringe effects" that are often observed at the entry and exit to the contact were attributed to the divergence of ultrasonic energy. This was identified as occurring due to the reflection of acoustic energy from the contact before it is directly within the line-of-sight of the ultrasonic beam, thus causing the measured reflection coefficient to rise above 1. An increased waviness in the envelope of the raw ultrasonic data stream was found to result from the abundance of oil within the bearing contact. This was identified through

comparison of two data sets, one obtained under "fully-flooded" lubricant conditions, and the other obtained when the bearing was known to have been run dry due to a lubricant system failure. As such this was identified as a means to assess the status of a bearing lubricant. Finally, the observation of a "through-roller reflection" was discussed as a potential means to determine the width of the rolling contact, which could subsequently be used to infer applied load. The need for an accurate measurement of cage speed was highlighted however, since any changes in bearing speed would also alter the rate at which a roller would move past the ultrasonic element, thus affecting the observed width.

### Ultrasonic Reflection vs Bearing Life

Ultrasonic data was acquired during a full bearing life test so as to observe any changes in the degree of ultrasonic reflection induced at the inner raceway/rolling element contact. A reduction in reflection coefficient was observed which could confidently be linked to the release of a spall from the inner raceway surface. This is thought to have been induced due to a general breakdown in the oil film, leading to an increase in the metal-to-metal contact occurring across the observed interface. Unfortunately however, the ultrasonic response was found to occur at the same time as a rise in inner raceway temperature, and therefore appeared to provide no advanced warning of failure over the more conventional condition monitoring approach in this case. What is encouraging however, is that the spalling failure was minor and occurred a significant distance outside the line of sight of the sensor. It was therefore concluded that a similar response may have occurred sooner should the failure have initiated closer to the ultrasonic sensor.

## 9.5 Industrialisation

As part of this work, the ultrasonic technique has also been applied to an operational wind turbine gearbox so as to monitor bearing condition in a similar manner to that employed on the test platform. Ideally, instrumentation would have been applied to planetary bearing inner raceways, but the complexities involved in instrumenting a rotating component were considered too challenging for the initial investigation. Subsequently a high speed shaft bearing – also known to exhibit high failure rates – was instrumented. A data acquisition system was developed to log data from the sensors, as well as signals from acoustic emissions sensors mounted to the same bearing. Overall therefore, the monitoring system represented a significant advancement to anything

known to have been implemented within a wind turbine gearbox previously.

Data obtained from an early bench test of the gearbox prior to installation was analysed and data was found to respond as expected to the range of operational loads under which the gearbox was placed. The system has since been commissioned within the turbine and has been active since mid-June 2015.

# 9.6 Further Work

**MultiLife Development and Optimisation**

Within this work the fundamental principles of the MultiLife concept have been tested analytically and experimentally to illustrate that raceway indexation could successfully extend WTGB planetary stage bearing life. In addition to this, the analytical methods developed in Chapter 5 have been further utilised to assess the effects that stress-raising impact events could have upon MultiLife performance, which in-fact could be utilised to further enhance MultiLife operation in service, such that raceway indexation is intelligently initiated based on the location and severity of inner raceway damage. This of course requires knowledge of real-time bearing condition, and the ultrasonic technique – as well as more direct temperature measurements – have been demonstrated as a means to provide such insight. A key element of further work will therefore involve the development of robust algorithms which utilise such bearing condition monitoring measurements to establish remaining useful life and determine bearing indexing requirements appropriately and in real-time.

This will require extensive analysis of the data obtained from the CM system presented in Chapter 8, and further development of the analytical methods discussed in Chapter 5. Additionally, a degree of further experimentation would also be recommended, to build upon the work presented in Chapter 7, particularly in regards to the observed ultrasonic response. Regrettably, due to time constraints and rig issues it was only possible to perform one complete bearing life test with ultrasonic instrumentation installed. A number of repeated experiments would thus ideally be carried out in order to show that the observed trends are indeed repeatable, so as to gain additional confidence in the conclusions that have been made.

**Technical**

It was clear from the outcomes of both the bearing tests presented in Chapter 7 that load misalignment was a recurrent issue, which was found to be very difficult to prevent. In the time between the two tests the main rig bearing was changed from a self-aligning SRB (as shown in Figure 4.1) to a CRB which – it was hoped – would restrict any natural inclination for the rig to tilt, thus forcing the load to act more evenly across the entire width of the test bearing. This was however found to be unsuccessful, as evidenced by the outcomes of the second test in §7.2, and therefore a further design review is essential so as to identify additional methods to prevent this issue during future experimentation.

The method of fault seeding utilised during the bearing life testing was selected as a straight forward and directly measurable way simulate the stress raising effect of the impact events known to occur in a WTGB. In order to be entirely representative of the in-field failure mechanism however, the induced stresses should be seeded sub-surface through controlled impact to the rolling surface. A brief investigation into such methods was discussed in §4.2.1, yet the required sub-surface damage was not thought to have occurred under the maximum impact energy applicable on the test platform utilised. As also discussed in the aforementioned section, compression of a roller against a raceway is thought to have similar effects to the material sub-structure, and thus may provide an alternative approach to achieving more representative faults. A further advantage of the latter technique is that this can be performed utilising equipment located within the University Of Sheffield Department Of Civil Engineering, and is thus more accessible for further, more detailed investigation. It is therefore recommended that a range of compressive tests are performed on the NU2244 bearings, and accelerated life testing subsequently performed, to identify if a similar localised spalling failure can be reproduced as a result of this – more representative – sub-surface fault seeding approach.

**The Ultrasound Technique**

Various options were explored in §4.3.2 to utilise ultrasound sensors without the need to permanently bond piezoelectric elements to the bearing raceway. Clearly, this would be a key requirement if the technique were to be an integral part of the MultiLife system, so as to ensure the failure critical raceway segment is always observed. Methods involving an ultrasonic coupling rubber were investigated within this work, which was eventually deemed inappropriate for various reasons. As discussed in

§4.3.2.3 however, silicone based materials may well provide a better solution, or indeed alternative coupling mediums such as liquid, provided sufficient sealing can be maintained.

Alternatively to the above, a more robust method might involve the permanent bonding of an array of ultrasonic elements, covering the entire circumference of the inner raceway bore. In turn then, signal cabling could be fixed within the bearing shaft at the required measurement location, and an electrical connection would be made (via spring loaded "pogo" pins for instance) to the nearest available ultrasonic element on the inner raceway bore following each indexation instance. Of course, instrumentation on such a scale would be time consuming, and would therefore perhaps be unfeasible for tests of multiple bearing on the test platform, but may provide a viable option if it were to be supplied as part of a commercial "MultiLife Bearing" unit.

Ultimately, the localised nature of the ultrasonic monitoring technique was highlighted as a limitation, based on the outcomes of the test presented in §7.2. This further highlights the need for higher resolution arrays of ultrasonic sensors so as to enable monitoring across the entire length of the contact.

# References

1. Department of Energy & Climate Change [DECC] (2013). *UK Renewable Energy Roadmap Update 2013*. Available from: https://goo.gl/ldK6YO

2. European Wind Energy Association [EWEA] (2011). *Pure power, wind energy targets for 2020 and 2030*. Available from: http://goo.gl/0yg3mK

3. Helsen, J., et al. *Some trends and challenges in wind turbine upscaling*. in *Proceedings of ISMA International Conference On Noise And Vibration 2012*.

4. Low Carbon Innovation Coordination Group [LCICG] (2012). *Technology Innovation Needs Assessment (TINA) - Offshore Wind Power Summary Report*. Available from: https://goo.gl/DOUo8T

5. Ribrant, J. and L. Bertling. *Survey of failures in wind power systems with focus on Swedish wind power plants during 1997-2005*. in *Power Engineering Society General Meeting, 2007. IEEE*. IEEE.

6. GCube (2012). *Top 5 US Wind Energy Claims*. Available from: http://goo.gl/qGqgmO

7. Marsh, G., *Offshore reliability*. Renewable Energy Focus, 2012. **13**(3): p. 62-65.

8. Evolution - the business and technology magazine from SKF. *Premature bearing failures in wind gearboxes and white etching cracks (WEC)*. 2013; Available from: http://goo.gl/K6miKW.

9. Oyague, F., *Gearbox modeling and load simulation of a baseline 750-kW wind turbine using state-of-the-art simulation codes*. 2009: National Renewable Energy Laboratory.

10. Musial, W., S. Butterfield, and B. McNiff. *Improving wind turbine gearbox reliability*. in *European Wind Energy Conference, Milan, Italy*.

11. Wheals, J.C. *Improved Wind Turbine Gearbox Reliability Through Novel Concepts for Bearings - MultiLife$^{TM}$*. in *European Offshore Wind*. 2009.

12. ISO 281:2007. Rolling bearings - Dynamic load ratings and rating life. (International Organisation for Standardisation, 2007)

13. Evans, M., *White structure flaking (WSF) in wind turbine gearbox bearings: effects of 'butterflies' and white etching cracks (WECs).* Materials Science and Technology, 2012. **28**(1): p. 3-22.

14. Windpower Monthly. *Wind turbine gearboxes and the effort to improve their reliability.* 2010; Available from: http://goo.gl/CXhelP.

15. Evans, M.H., et al., *Serial sectioning investigation of butterfly and white etching crack (WEC) formation in wind turbine gearbox bearings.* Wear, 2013. **302**(1–2): p. 1573-1582.

16. Wheals, J. *OWDIn - MultiLife Impact Tests V03*, 2014, Ricardo UK Ltd. Unpublished Presentation.

17. Bruce, T., et al., *Characterisation of white etching crack damage in wind turbine gearbox bearings.* Wear, 2015. **338**: p. 164-177.

18. Kotzalas, M.N. and G.L. Doll, *Tribological advancements for reliable wind turbine performance.* Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2010. **368**(1929): p. 4829-4850.

19. Scott, K.G., et al. *Effects of Extreme and Transient Loads on Wind Turbine Drive Trains*. in *50th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*. Nashville, USA.

20. Naumann, J., *Acoustic emission monitoring of wind turbine bearings*, 2015, PhD Thesis, University of Sheffield

21. Harris, T.A. and M.N. Kotzalas, *Essential concepts of bearing technology: rolling bearing analysis*. 5th ed. 2006: CRC Press.

22. Harris, T.A. and M.N. Kotzalas, *Advanced concepts of bearing technology: rolling bearing analysis*. 5th ed. 2006: CRC Press.

23. Khonsari, M.M. and E.R. Booser, *Applied tribology: bearing design and lubrication*. Vol. 12. 2008: John Wiley & Sons.

24. Eschmann, P. and L. Hasbargen, *Ball and roller bearings: theory, design, and application*. 1985: John Wiley & Sons.

25. Palmgren, A., *Ball and roller bearing engineering*. 3rd ed. SKF Industries Inc., Philadelphia. 1959.

26. Tallian, T.E., *Failure atlas for Hertz contact machine elements*. 1992: American Society of Mechanical Engineers.

27. The Timken Company (2014). *Timken bearing damage analysis with lubrication reference guide*. Available from: http://goo.gl/mQKqSR

28. NSK (2009). *New Bearing Doctor: Maintenance of Bearings*. Available from: http://goo.gl/nEYrgR

29. Williams, J., *Engineering tribology*. 2005: Cambridge University Press.

30. SKF. *Bearing Damage: Wear*. [cited 5th March 2015]; Available from: http://goo.gl/tn9ZIT.

31. Morales-Espejel, G.E. and V. Brizmer, *Micropitting Modelling in Rolling–Sliding Contacts: Application to Rolling Bearings.* Tribology Transactions, 2011. **54**(4): p. 625-643.

32. Sadeghi, F., et al., *A review of rolling contact fatigue.* Journal of Tribology, 2009. **131**(4): p. 041403.

33. Nélias, D., et al., *Role of inclusions, surface roughness and operating conditions on rolling contact fatigue.* Journal of tribology, 1999. **121**(2): p. 240-251.

34. Martin, J. and A. Eberhardt, *Identification of potential failure nuclei in rolling contact fatigue.* Journal of Fluids Engineering, 1967. **89**(4): p. 932-942.

35. Zaretsky, E., *STLE life factors for rolling bearings.* 2nd ed. 1999: Society of Tribologists and Lubrication Engineers.

36. Jacobson, B., *The Stribeck memorial lecture.* Tribology International, 2003. **36**(11): p. 781-789.

37. Harris, T. and R. Barnsby, *Life ratings for ball and roller bearings.* Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 2001. **215**(6): p. 577-595.

38. Leng, X., Q. Chen, and E. Shao, *Initiation and propagation of case crushing cracks in rolling contact fatigue.* Wear, 1988. **122**(1): p. 33-43.

39. Schlicht, H., E. Schreiber, and O. Zwirlein, *Effects of material properties on bearing steel fatigue strength*, in *Effect of steel manufacturing processes on the quality of bearing steels*, J.J.C. Hoo, Editor. 1986. p. 81-101.

40. Ioannides, E. and T. Harris, *A new fatigue life model for rolling bearings.* Journal of Tribology, 1985. **107**(3): p. 367-377.

41. Nagaraj, H.S., D.M. Sanborn, and W.O. Winer, *Direct surface temperature measurement by infrared radiation in elastohydrodynamic contacts and the correlation with the blok flash temperature theory.* Wear, 1978. **49**(1): p. 43-59.

42. ISO 4406:1999. Hydraulic fluid power - Fluids - Method for coding the level of contamination by solid particles. (International Organisation for Standardisation, 1999)

43. Springer, T.E., et al., *Assessing Bearing Performance Under Debris-Contaminated Lubrication Conditions*, in *Practicing Oil Analysis*. 2004, Noria Corporation.

44. Agoston, A., et al., *A concept of an infrared sensor system for oil condition monitoring.* e & i Elektrotechnik und Informationstechnik, 2008. **125**(3): p. 71-75.

45. Mauntz, M.R., et al., *A sensor system for online oil condition monitoring of operating components.* Tribology—Fundamentals and Advancement, 2013.

46. Lacey, S.J., *An Overview of Bearing Vibration Analysis*, in *ME: Maintenance & Asset Management*. 2008. p. 32-42.

47. McFadden, P. and J. Smith, *Vibration monitoring of rolling element bearings by the high-frequency resonance technique—a review.* Tribology international, 1984. **17**(1): p. 3-10.

48.     Tandon, N. and A. Choudhury, *A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings.* Tribology international, 1999. **32**(8): p. 469-480.

49.     Dyer, D. and R. Stewart, *Detection of rolling element bearing damage by statistical vibration analysis.* Journal of Mechanical Design, 1978. **100**(2): p. 229-235.

50.     Igarashi, T. and H. Hamada, *Studies on the Vibration and Sound of Defective Rolling Bearings: First Report: Vibration of Ball Bearings with One Defect.* Bulletin of JSME, 1982. **25**(204): p. 994-1001.

51.     Tandon, N., *A comparison of some vibration parameters for the condition monitoring of rolling element bearings.* Measurement, 1994. **12**(3): p. 285-289.

52.     Mathew, J. and R. Alfredson, *The condition monitoring of rolling element bearings using vibration analysis.* Journal of Vibration and Acoustics, 1984. **106**(3): p. 447-453.

53.     Antoniadou, I., *Accounting for nonstationarity in the condition monitoring of wind turbine gearboxes*, 2013, PhD Thesis, University of Sheffield

54.     Yoshioka, T., *Detection of rolling contact sub-surface fatigue cracks using acoustic emission technique.* Lubrication Engineering (United States), 1993. **49**(4).

55.     Li, C.J. and S. Li, *Acoustic emission analysis for bearing condition monitoring.* Wear, 1995. **185**(1): p. 67-74.

56.     Mba, D. and R.B. Rao, *Development of Acoustic Emission Technology for Condition Monitoring and Diagnosis of Rotating Machines; Bearings, Pumps, Gearboxes, Engines and Rotating Structures.* 2006.

57.     Al-Ghamd, A.M. 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, 2006. **20**(7): p. 1537-1571.

58.     Mba, D., *The use of acoustic emission for estimation of bearing defect size.* Journal of Failure analysis and prevention, 2008. **8**(2): p. 188-192.

59.     Rindorf, H.J., *Acoustic emission source location in theory and in practice.* 1981: Brüel & Kjær.

60.     Baxter, M.G., et al., *Delta T source location for acoustic emission.* Mechanical systems and signal processing, 2007. **21**(3): p. 1512-1520.

61.     Feng, Y., et al., *Monitoring wind turbine gearboxes.* Wind Energy, 2013. **16**(5): p. 728-740.

62.     Schlechtingen, M. and I. Santos, *Condition monitoring with ordinary wind turbine SCADA data–A neuro-fuzzy approach.* Offshore 2011, 2011.

63.     Wilkinson, M., et al., *Comparison of methods for wind turbine condition monitoring with SCADA data.* IET Renewable Power Generation, 2014. **8**(4): p. 390-397.

64.     Kim, K., et al. *Use of SCADA data for failure detection in wind turbines*. in *ASME 2011 5th International Conference on Energy Sustainability*. American Society of Mechanical Engineers.

65.     Zaher, A., et al., *Online wind turbine fault detection through automated SCADA data analysis*. Wind Energy, 2009. **12**(6): p. 574-593.

66.     Yang, W., et al., *Wind turbine condition monitoring: technical and commercial challenges*. Wind Energy, 2014. **17**(5): p. 673-693.

67.     Zhang, Z., A. Verma, and A. Kusiak, *Fault analysis and condition monitoring of the wind turbine gearbox*. Energy Conversion, IEEE Transactions on, 2012. **27**(2): p. 526-535.

68.     Olympus NDT (2011). *Ultrasonic Transducers - Technical Notes*. Available from: http://goo.gl/053eje

69.     Wan Ibrahim, M., D. Gasni, and R. Dwyer-Joyce, *Profiling a ball bearing oil film with ultrasonic reflection*. Tribology Transactions, 2012. **55**(4): p. 409-421.

70.     Harper, P., *Measurement of film thickness in lubricated components using ultrasonic reflection*, 2008, PhD Thesis, University of Sheffield

71.     Tattersall, H., *The ultrasonic pulse-echo technique as applied to adhesion testing*. Journal of Physics D: Applied Physics, 1973. **6**(7): p. 819.

72.     Mills, R., *Ultrasonic measurement of lubricant films generated at the piston-cylinder interface of internal combustion engines*, 2012, PhD Thesis, University of Sheffield

73.     Thomas, T. and R. Sayles, *Stiffness of machine tool joints: a random-process approach*. Journal of Manufacturing Science and Engineering, 1977. **99**(1): p. 250-256.

74.     Reddyhoff, T., et al., *The phase shift of an ultrasonic pulse at an oil layer and determination of film thickness*. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 2005. **219**(6): p. 387-400.

75.     Dwyer-Joyce, R., B. Drinkwater, and C. Donohoe, *The measurement of lubricant–film thickness using ultrasound*. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 2003. **459**(2032): p. 957-976.

76.     Povey, M.J., *Ultrasonic techniques for fluids characterization*. 1997: Academic Press.

77.     Dwyer-Joyce, R. and B. Drinkwater. *Analysis of contact pressure using ultrasonic reflection*. in *Experimental Mechanics, Proceedings of the 11th Annual Conference on Experimental Mechanics*.

78.     Hodgson, K., R. Dwyer-Joyce, and B. Drinkwater, *Ultrasound as an experimental tool for investigating engineering contacts*. TRIBOLOGIA-TAMPERE-, 2000. **19**(4): p. 9-17.

79.     Gonzalez-Valadez, M., R. Dwyer-Joyce, and R. Lewis, *Ultrasonic reflection from mixed liquid-solid contacts and the determination*

*of interface stiffness.* Tribology and Interface Engineering Series, 2005. **48**: p. 313-320.

80.   Dwyer-Joyce, R., T. Reddyhoff, and J. Zhu, *Ultrasonic measurement for film thickness and solid contact in elastohydrodynamic lubrication.* Journal of Tribology, 2011. **133**(3): p. 031501.

81.   Dwyer-Joyce, R., T. Reddyhoff, and B. Drinkwater, *Operating limits for acoustic measurement of rolling bearing oil film thickness.* Tribology transactions, 2004. **47**(3): p. 366-375.

82.   Dwyer-Joyce, R., P. Harper, and B. Drinkwater, *A method for the measurement of hydrodynamic oil films using ultrasonic reflection.* Tribology Letters, 2004. **17**(2): p. 337-348.

83.   Reddyhoff, T., R. Dwyer-Joyce, and P. Harper, *Ultrasonic measurement of film thickness in mechanical seals.* Sealing Technology, 2006. **2006**(7): p. 7-11.

84.   Mills, R., J. Vail, and R. Dwyer-Joyce, *Ultrasound for the non-invasive measurement of internal combustion engine piston ring oil films.* Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 2015. **229**(2): p. 207-215.

85.   Zhang, J., B.W. Drinkwater, and R.S. Dwyer-Joyce, *Monitoring of lubricant film failure in a ball bearing using ultrasound.* Journal of tribology, 2006. **128**(3): p. 612-618.

86.   Fuentes, R., et al., *Observations on the acoustic emissions from a line contact compressed into the plastic region*, in *42nd Leeds-Lyon Symposium on Tribology,* 7-9 September 2015, Lyon, France. (Currently Unpublished).

87.   Drinkwater, B., R. Dwyer-Joyce, and P. Cawley, *A study of the transmission of ultrasound across solid–rubber interfaces.* The Journal of the Acoustical Society of America, 1997. **101**(2): p. 970-981.

88.   Brotherhood, C., B. Drinkwater, and R. Freemantle, *An ultrasonic wheel-array sensor and its application to aerospace structures.* Insight-Non-Destructive Testing and Condition Monitoring, 2003. **45**(11): p. 729-734.

89.   Olympus IMS. *Application Notes: Ultrasonic Couplants.* [cited 10th August 2014]; Available from: http://goo.gl/uChp4z.

90.   NDT Resource Center. *Transduer Beam Spread.* [cited 26th October 2014]; Available from: https://goo.gl/Vlct5h.

91.   ISO 16281:2008. Rolling bearings - Methods for calculating the modified reference rating life for universally loaded bearings. (International Organisation for Standardisation, 2008)

92.   Pilkey, W.D. and D.F. Pilkey, *Peterson's stress concentration factors.* 2008: John Wiley & Sons.

93.   Budynas, R.G. and J.K. Nisbett, *Shigley's mechanical engineering design.* 8th ed. 2008: McGraw-Hill.

94.   MatWeb Material Property Data. *AISI E 52100 Steel.* [cited 3rd February 2015]; Available from: http://goo.gl/428eFb.

95. Darji, P. and D. Vakharia, *Development of Graphical Solution to Determine Optimum Hollowness of Hollow Cylindrical Roller Bearing Using Elastic Finite Element Analysis*, in *Finite Element Analysis Applications in Mechanical Engineering*, F. Ebrahimi, Editor. 2012, INTECH Open Access Publisher. p. 237-260.

96. Roark, R.J. and W.C. Young, *Formulas for stress and strain*. 5th ed. 1975: McGraw-Hill Kogakusha, Ltd.

97. Dowson, D. and G. Higginson, *A numerical solution to the elasto-hydrodynamic problem.* Journal of Mechanical Engineering Science, 1959. **1**(1): p. 6-15.

98. Marshall, M., et al., *Ultrasonic measurement of self-loosening in bolted joints.* Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2011. **226**(7): p. 1869-1884.

99. Zhu, J., *Simulation Model and Ultrasound Study for Engineering Interfaces*, 2012, PhD Thesis, University of Sheffield

100. Smith, R., *Stress-induced anisotropy in solids—the acousto-elastic effect.* Ultrasonics, 1963. **1**(3): p. 135-147.

101. Hikata, A., et al., *Ultrasonic attenuation and velocity data on aluminum single crystals as a function of deformation and orientation.* Acta Metallurgica, 1962. **10**(4): p. 423-429.

102. Abayev, Z., *Wind Turbine Bearing Tribology*, 2014, MSc Thesis, University of Sheffield

103. Hawke, D. *RD.14/308201.1 Wind Turbine Condition Monitoring System (CMS) Test Specification*, 2014, Ricardo UK Ltd. Unpublished Internal Document.