The path of rolling elements in defective bearings:
observations, analysis and methods to estimate spall size

Alireza Moazen ahmadi\textsuperscript{a,1}; Carl Q. Howard\textsuperscript{b}; Dick Petersen\textsuperscript{b}

\textsuperscript{a}School of Mechanical Engineering, The University of Adelaide, Australia
\textsuperscript{b}Trackside Intelligence Pty Ltd, 17–19 King William Street, Kent Town, South Australia 5067, Australia

ABSTRACT

This paper describes the experimental investigation of the vibration signature generated by rolling elements entering and exiting a notch defect in the outer raceway of a bearing. The vibration responses of the bearing housing and the displacement between the raceways were measured and analyzed. These key features can be used to estimate the size of the defect and is demonstrated in this paper for a range of shaft speeds and bearing loads. It is shown that existing defect size estimation methods include assumptions that describe the path of the rolling elements in the defect zone leading to poor estimates of defect size. A new defect size estimation method is proposed and is shown to be accurate for estimating a range of notch defect geometries over a range of shaft speeds and applied loads.

Keywords: rolling element bearing, bearing defects, bearing condition monitoring, defect size estimation

1. Introduction

Rolling element bearings are widely used in rotating machinery and bearing failure is one of the common reasons for machinery breakdowns. Effective bearing condition monitoring systems should be able to detect and estimate the size of defects in bearings at early stages of a defect development to either enable remedial action to be taken, or schedule the replacement of the bearing at a convenient time. The typical vibration condition monitoring process involves the measuring vibration level on a bearing housing and trending the level over time. When the level exceeds a nominated threshold, the bearing is flagged for replacement. Alternatively, the philosophy used in this paper is to determine the geometric arc length of a bearing defect from the vibration signal, without using historical data.

Defects in bearings are commonly categorized into localized and distributed defects. Distributed defects, such as waviness, surface roughness or off-size rolling elements, are usually the result of manufacturing errors [1, 2]. Localized defects are often initiated by insufficient lubrication film between the contact surfaces. This causes metal-to-metal contact between the rolling elements and the raceways. This in turn generates stress waves, leading over time to the formation of sub-surface cracks. The large forces between bearing parts cause sub-surface cracks to grow into surface defects, a phenomenon called pitting or spalling [3]. This paper considers vibrations and displacements between raceways generated in bearings with an outer raceway line-spall of various dimensions.

Previous studies on the vibration signature of defective bearings with raceway spalls show that the passage of a rolling element over the spall generates two main components [4-7]. The first component, which has low frequency content, results from the entry of a rolling element into the spall. The second component, which has higher frequency content, results from the impact of a rolling element into the trailing edge of the spall [6, 8-10]. Several defect size estimation methods have been previously suggested based on the time separation between these two components from the vibration signal [6, 11]. These algorithms are based on...
assumptions of the path of a rolling element as it traverses a defect, and led to less accurate estimates of the defect size, as shown in Section 0 of this paper.

In the work conducted here, a bearing test rig was used to measure the vibration of a “floating” bearing housing, and eddy current proximity probes were used to measure the relative displacement between the inner and outer raceways. Deep-groove ball bearings with electro-discharged-machined (EDM) notches with various dimensions were installed on the test rig and the vibration signals were measured for a range of speeds and loads. The design of the test rig made it possible to investigate the hypotheses proposed in previous studies and findings from previous simulation results. There are several studies on the effect of the high frequency vibration responses of defective bearings in the literature [12, 13]. Epps and McCallion [14] have investigated the effect of speed and load on the amplitude of the low frequency entry and exit events. However, investigation of the effect of speed and load on the angular extent of the low frequency entry and exit transient events has not been done previously. Based on the findings in this study, a defect size estimation method is proposed. The proposed method is compared with previous defect size estimation methods for various defect sizes, speeds and loads. The proposed method is shown to be more accurate and reliable than previous methods in estimating the defect size in bearings.

The main contributions of this paper are (1) an improved method of estimating the size of a line spall defect based on vibration signals on the bearing housing, (2) comparing and highlighting the relationship between the vibration signature and the shaft displacement signal as a rolling element passes through a line spall defect, and (3) the effect of speed and load on the duration of the low frequency entry and exit transient vibration signals.

This paper is structured as follows: Section 2 reviews the previous experimental and modelling studies of the vibration response of defective bearings in regards to the path of rolling elements in the defect zone. Section 3 describes the test rig used in the experiments. Section 4 presents the experimental outcomes and the detailed analysis of the vibration and displacements between raceways to indirectly study and explain the probable path of the rolling element in the defect zone for various speeds and loads. Section 0 presents a reliable defect size estimation method resulting from the analysis presented in the previous sections. The method is tested on three bearings with various defect sizes, for a range of speeds and loads. The estimated bearing defect size using the method proposed in this paper is compared with previously proposed estimation methods. Section 6 summarizes the findings of this work.

2. Previous work

Figure 1(a) illustrates a typical defective rolling element bearing with a localized line-spall defect on the outer raceway. A typical measured vibration response generated by the defect is shown in Figure 1(b). Previous experimental studies show that the entry of a rolling element into a line-spall defect produces a vibration signal with low frequency content, while the exit of the rolling element excites a much broader range of frequencies, including the high frequency bearing resonances [4, 6, 14]. These resonances are excited by the impact of the rolling element on the exit point of a defect. Furthermore, parametric excitations are caused by rapid changes in the bearing stiffness which occur when the rolling element re-stresses between the raceways [8]. The high frequency event observed in the experimental results [10, 14] often appears to have been caused by multiple impacts rather than a single impact. Numerical simulation results of defective bearings presented by Singh et al. [9] and the analytical simulation results presented by Moazen Ahmadi et al. [4] indicate that multiple impacts occur when the rolling element re-stresses at the exit point.
Limited experimental studies have examined the relationship between the path of a rolling element and the characteristics appearing in the vibration signal. Sawalhi and Randall [6] measured the vibration of several defective ball bearings and observed the appearance of low and high frequency vibration signatures and related these events to the entry and exit points of the defect. Those authors suggested that when the low frequency vibration event is at its maximum local amplitude (see point entry event point in Figure 1(b)), the center of a rolling element passes the entry point. Furthermore, these authors assumed that the high frequency exit event is associated with the time that the center of a rolling element is in the middle of the defect. These hypotheses were examined indirectly in an experiment which aimed to estimate the defect size in a bearing using a method based on these hypotheses. However, the estimation results of different bearings often showed large errors and standard deviations [6]. Their assumption that the exit-impact event occurs when the center of a rolling element is in the middle of the defect zone might be valid for very small defects. However the authors did not specify the size limit of the defect for which their suggested method is valid.

Numerous multi-body dynamic models have been developed to understand the relationship between the vibration characteristics and the path of the rolling elements in the defect zone for line-spall defects [15-21]. Although some of these models include the mass of the rolling element, they do not include both, the mass and the finite size of the rolling element. In these models, the path of a rolling element is modelled such that its center follows the geometry of the modelled defect. Harsha [22-25] considered the mass and centrifugal forces acting on a rolling element but not the finite size of the rolling element in a multi-body dynamic model. This model was initially developed to predict the vibration response of defective bearings with distributed defects [26]. The model was subsequently improved to include the mass of the rolling elements to predict the nonlinear dynamic behavior of a rolling element bearing, due to waviness and unbalanced rotor support [22-25]. The improved version of the model was further modified by Tadina [21] to predict the vibration response of bearings with localized spall defects on raceways. All of the aforementioned models were designed for defects with curvatures larger than the curvature of the rolling element, which maintains the contact between the raceways in the load zone. Therefore, none of the above mentioned multi-body dynamic models are suitable for modelling the path of a rolling element in the defect zone or determining the effect of the entry and exit points on the vibration signal of a defective bearing with a line-spall defect.

A more comprehensive multi-body dynamic model, developed by Moazen Ahmadi et al. [4, 10], considers the mass and centrifugal forces acting on rolling elements and the finite size of the rolling elements. These authors show that in order to predict the path of a rolling element and the corresponding features in the vibration response correctly, it is crucial to include the finite size of rolling elements when modelling localized defects. Their analytical simulation revealed that the local maximum of the low frequency signature in the vibration response corresponds to the instant that a rolling element completely de-stresses between the two raceways upon entering the defect. Moreover, the process of de-stressing starts well before the entry point. Therefore, based on the analytical simulation of Moazen Ahmadi et al. [4, 10], it can be concluded that the entry event, shown as the entry point in Figure 1(b), does not correspond to the time that
the center of the rolling element is at the entry point of the defect. This conclusion contradicts the observations and analyzes of an earlier study [6].

In this paper, the relationship between the characteristics of the vibration response and the relative displacement between the raceways to study and explain the probable interactions of a rolling element with raceways in the load zone of defective bearings is explored. Eddy current proximity probes were used to measure the relative displacement between the inner and outer raceways. Experimental measurements show the relationship between entry and exit events and vibration response in defective bearings. The effects of the rotational speed and the applied load on the magnitude of the entry and exit events on the acceleration signal are investigated. These findings are used to propose an accurate defect size estimation method.

This study investigated the assumptions used in previously described defect size estimation methods and identifies discrepancies of the assumed path of a rolling element into a defect and the corresponding features on the vibration response to the entry and exit events. These discrepancies contribute to inaccuracies found in the defect size estimation algorithms proposed previously. This paper demonstrates that impacts at the exit do not necessarily occur midway through a defect and previous assumptions made regarding the relationship between the vibration response and the entry and exit events are not accurate. The defect size estimation method proposed in this paper is based on numerical modelling [4] and supported by the experimental findings presented in this paper.

3. Test equipment

The test rig used in this study was manufactured by Spectra Quest, Inc., as shown in Figure 2. The bearing housing supplied with the test-rig was replaced with a custom-made housing to enable testing of the hypotheses proposed in this study. The test rig includes an electric motor controlled by a variable frequency drive. The motor was fitted onto a shaft via a coupling and supported by two back-to-back tapered roller bearings. Test bearings were fitted into a custom-made bearing housing at the end shaft that has a radial load applied by a hydraulic jack.

![Figure 2: Top view of the test rig used in this study.](image)

Figure 3 shows the custom-made bearing housing that was designed to accommodate accelerometers to measure the vibration of the housing, and eddy current proximity probes to measure the relative displacement between the inner and outer raceways of the test bearing.
Two stud mounted accelerometers (Bruel & Kjaer type 4393) were attached to the bearing housing to measure vibrations. Eddy current proximity probes (Micro-Epsilon type EPU05-C3) were mounted on the bearing housing to measure the relative displacement between the inner and outer raceways. A tachometer was used to measure the shaft speed and a load cell measured the force applied to the bearing housing. The data acquisition system consisted of a National Instruments (NI) CompactDAQ system with two NI 9234 modules. Data was acquired and post-processed using MATLAB software. All signals were acquired with a sampling frequency of 102.4 kHz.

The test bearings used in this study were ball bearings manufactured by Rexnord (ER16K), incorporating 9 balls, ball diameter of 7.94mm, pitch diameter of 39.32mm, and a contact angle of 0°. Line spall defects with various circumferential lengths, which were less than the angular separation of the rolling elements, and across the full axial extent of the outer raceway, were machined using electro-discharge machining (EDM) on the outer raceway of the bearings. Figure 4 shows microscopic images of the three test bearings that were tested. Table 1 lists the geometry of the defects tested in this study that was measured using an Olympus model BX60M microscope fitted with a digital camera and software for obtaining calibrated geometry measurements. The defective bearing was installed into the bearing house such that the center of the defect was aligned with the radial load.
Figure 4: Microscopic images of three defective bearings tested in this study. Pictures show the width of the defect and the scale in pixels measured on the digital photo. (a) 3.11mm (b) 2.24mm (c) 1.61mm.
Table 1: Geometry of the defects in the three test bearings.

<table>
<thead>
<tr>
<th>Measured defect circumferential length (mm)</th>
<th>Nominal angular extent (rad)</th>
<th>Radial depth (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.11</td>
<td>0.138</td>
<td>100</td>
</tr>
<tr>
<td>2.24</td>
<td>0.103</td>
<td>100</td>
</tr>
<tr>
<td>1.48</td>
<td>0.068</td>
<td>100</td>
</tr>
</tbody>
</table>

4. Measurements

This section presents an analysis of the measured signals using the test rig described in the previous section. Measurements were conducted at rotational speeds of 5Hz, 10Hz and 15Hz and at radial loads of 2.5kN and 5kN. The aim of the analysis was to provide insights into the relationships of the vibration and displacement responses to indirectly study the probable path of the rolling element in the load zone and use the results as a basis for developing a method to estimate the size of a defect. This study distinguishes the low frequency entry and exit transient events and the low frequency response of the pedestal after the transient events. There are several studies on the effect of the high frequency vibration responses of defective bearings exists in the literature [12, 13]. However, investigation of the effect of speed and load on the duration of the low frequency entry and exit transient events has not been done previously. The effects of load and speed on the characteristics of the entry and exit acceleration signal are also investigated in this section. These analyses are important to develop a reliable and accurate defect size estimation method.

4.1 Analysis of entry and exit points

Figure 5(a) and (b) present two revolutions of the measured vibration and outer-to-inner race relative displacement measured on the bearing with a 0.103 rad angular extend defect at a rotational speed of 5Hz and 2500N load. The measured relative displacement between the two raceways for a new (undamaged) bearing is presented in Figure 5(c) for comparison.
Figure 5: Vibration signals generated by a defective bearing with a 0.103rad angular extent defect at a rotational speed of 5Hz with a 2.5kN load showing (a) acceleration of the bearing housing, (b) relative displacement between the outer raceway to inner raceway for the defective bearing, and (c) relative displacement of the outer raceway to inner raceway for an undamaged bearing.

Compared with the undamaged bearing, the relative displacement of the faulty bearing clearly shows sharp and large magnitude impulses each time an event is occurs on the acceleration signal, as shown in Figure 5(a) and (b). These vibration impulses are caused when a ball bearing enters and exits the defect zone. Note that the oscillation of the displacement signal at the run speed frequency is the result of the small out-of-roundness measured on the inner raceway, which could be the result of quenching the inner raceway.

Figure 6 shows a zoomed-in section of the acceleration signal and the corresponding relative displacement signal. This zoomed-in section corresponds to the period when a roller approaches and leaves the defect. The key characteristics observed in Figure 6 are described below.
Figure 6: (a) Zoomed measured signals at a 5kN load and rotational speeds of 5Hz (a) Vibration response (b) Relative displacement of the outer raceway to inner raceway signal.
**Rolling element entry event**

The entry of the rolling element into the defect generates predominantly low frequency content in the vibration response. Observations of the relationship between the relative displacement signal and the acceleration signal reveal that the entry event is composed of two components which are summarized in the following Table 2:

Table 2: Description of the entry event components and the related schematic view of the location of the rolling element

<table>
<thead>
<tr>
<th>Schematic view of the location of the rolling element in the defect zone</th>
<th>Description of the entry event</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Schematic view" /></td>
<td><strong>Entry-transient phase:</strong> In this phase, the acceleration amplitude displayed a small decrease followed by a sharp increase due to the gradual de-stressing of the rolling element in the defect zone. This phase started at entry point (a_1) on the acceleration and (a_2) on the relative displacement signals. This phase ended when the rolling element loses contact with both raceways at point (b_1) on the acceleration and (b_2) on the relative displacement signals, as shown in Figure 6. In this phase, the center of the rolling element travelled from the entry point of the defect to a point through the defect at which the contact forces acting on the rolling element become zero.</td>
</tr>
<tr>
<td><img src="image2.png" alt="Schematic view" /></td>
<td><strong>Entry-low-frequency response:</strong> Low frequency modes of the system were excited when the rolling element in the defect zone completely unloaded. At this stage, the number of the load carrying rolling elements decreases from 9 to 8 balls and consequently the bearing assembly stiffness decreases. Therefore, the low frequency modes of the system are excited as suggested by Petersen et al. [8].</td>
</tr>
</tbody>
</table>
Rolling element exit event

The rolling element exit event is generally associated with the excitation of high and low frequency modes of the bearing assembly. Observation of the relationship between the relative displacement signal and the acceleration signal suggest that the exit event exhibits a transient phase similar to the entry transient which has previously received little attention in the literature. The exit events consist of the components summarized in Table 3:

Table 3: Discerption of the exit event components and the related schematic view of the location of the rolling element

<table>
<thead>
<tr>
<th>Schematic view of the location of rolling element in the defect zone</th>
<th>Description of the exit event component</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="a" alt="Schematic view of the location of rolling element in the defect zone" /></td>
<td><strong>High-frequency response:</strong> The rolling element strikes the exit point of a defect and excites the high frequency resonance modes of the bearing assembly. The high-frequency event in the acceleration response observed at point c1 at angle 108.8 rad as shown in Figure 6(a), is associated with this event. Note that there is no corresponding event in the relative displacement response at time 108.8 rad in Figure 6(b). Further, the impact to the outer raceway is well before points d1 and d2, where the re-stressing phase between the raceways starts.</td>
</tr>
<tr>
<td><img src="b" alt="Schematic view of the location of rolling element in the defect zone" /></td>
<td><strong>Exit-transient phase:</strong> When the rolling element re-stresses between the raceways at the exit point of the defect, it re-stresses back to its normal load carrying capacity. The starting point of re-stressing is marked as points d1 and d2 in Figure 6 (a) &amp; (b) respectively, and is associated with the beginning of when the magnitude of the relative displacement signal decreases sharply.</td>
</tr>
<tr>
<td><img src="c" alt="Schematic view of the location of rolling element in the defect zone" /></td>
<td><strong>Exit-low-frequency response:</strong> The exit-transient phase ends at points e1 and e2. Between points d1 to e1 and d2 to e2, low frequency oscillations occur during this re-stressing phase. Depending on the damping characteristics of the bearing assembly, the exit low frequency response continues after points e1 and e2.</td>
</tr>
</tbody>
</table>
When a rolling element re-stresses between the raceways at the exit point of the defect, the rolling element may alternately strike both raceways as it re-stresses back to its normal load carrying capacity [27]. Multiple impacts at the exit of the defect have been observed in previous studies [9, 10]. Maximum local amplitude of the high frequency response can be either the result of the summation of several different excited natural frequencies or a superposition of the excited high frequency modes and the low frequency component excited in turn by the exit-transient event. Depending on dynamics of a bearing assembly and its pedestal, the two events, namely the high-frequency impact response and the low-frequency response, could be separated or superimposed. Therefore, detection of the maximum local amplitude of the high frequency response cannot be used as a reliable marker to indicate the exit point of the defect.

4.2 The effect of load

In this section, the effect of varying the applied load on a defective bearing is investigated in regards to the change in the amplitude and the angular extent of the low-frequency acceleration response. Figure 7 shows a comparison of the acceleration response and the relative displacement signal of the bearing with 0.103 rad angular extent defect at a rotational speed of 10 Hz for two radial loads. The angular positions of two signals were adjusted to match the entry point of the defect at which the entry-transient event starts, shown as point A in Figure 7. The constant offset of the relative displacement signal was also adjusted to coincide with the entry point.
Figure 7: Comparison of the measured signals for the defective bearing with 0.103 rad angular extent defect at a rotational speed of 10 Hz and two different loads, 5 kN and 2.5 kN (a) Acceleration signal (b) The relative inner-to-outer raceway displacement signal.

The increase in the relative displacement magnitude $\Delta$ shown in Figure 7(b) is the result of the increased relative contact deformation between the rolling elements and both raceways in the load zone due to the increase in the applied load. The maximum total elastic contact deformation on a rolling element can be calculated using the load-deflection relation [3]:
\[ \delta_{\text{max}} = \left( \frac{Q_{\text{max}}}{k_n} \right)^{1/n} \]

where \( n \) is the load-deflection exponent (\( n = 1.5 \) for point contact), \( Q_{\text{max}} \) is the maximum radial distributed load in the direction of the applied load and \( k_n \) is the total load-deflection factor of a bearing which depends on the curvature of rolling elements and raceways. Detailed descriptions and formulae for \( Q_{\text{max}} \) and \( k_n \) for ball bearings can be found in Harris [3]. For small defects which are smaller than angular separation of the rolling elements (with only one rolling element in the load zone), it can be assumed that the relative contact deformation between a rolling element and both raceways at the entry and exit points of the defect is equal to \( \delta_{\text{max}} \). The average increase of the relative displacement magnitude, \( \Delta \), is 14.5\( \mu \)m and the calculated increase of the elastic contact deformation due to the load increasing, using Eq. (1), is 12.1 \( \mu \)m. The agreement between the measured and calculated relative displacement justifies the adjustments made to plot the signals shown in Figure 7.

Relatively flat sections in the displacement signals, \( \theta_1 \) and \( \theta_2 \), are the angular extents where a rolling element is unloaded and travels through the defect before the start of the exit-transient phase. Figure 7(b) shows that the angular extent \( \theta \) decreases with increasing applied load, which may seem counter-intuitive as the defect size is identical for both tests. This is due to the additional angular extent that a rolling element requires to travel in order to lose contact between the raceways, as the elastic contact deformations are greater in bearings with a larger load. Points \( A_1 \) and \( A_2 \) show the end of the entry transient phases on Figure 7(b). Similarly, the period of time for the rolling element to re-stress between the raceways increases with increasing applied load. Since the elastic contact deformations between the load-carrying rolling elements and raceways are also greater in the bearing with higher load, the rolling element in the defect zone engages the raceways at the exit points \( (B_1 \) and \( B_2 ) \) of the defect earlier, as shown in Figure 7(b).

The presented analysis highlights the importance of including the effect of the applied load when attempting to develop a defect size estimation algorithm for rolling element bearings. The proposed defect size estimation method in this study includes the effect of the applied load on the duration of the entry and exit transient events to achieve accurate estimation of the defect size.

4.3 The effect of speed

In this section, the effect of varying the rotational speed is investigated in regards to the change in the angular extent of the low-frequency entry transient response. Figure 8 shows a comparison of the acceleration and the relative displacement signals of the bearing with a 0.103rad angular extent defect under 5kN load, at three rotational speeds. The angular positions of the signals were adjusted to match at the local maxima of the entry event. The relative displacement signals were also adjusted to coincide at the matched cage angular position as shown at point B in Figure 8(a) & (b).
Figure 8: Comparison of the measured signals for the bearing with a 0.103 rad angular extent defect at a 5kN load and rotational speeds of 5Hz, 10Hz and 15Hz (a) acceleration signal, and (b) relative displacement signal between the inner and outer raceways.

Figure 8 (b) shows that the magnitude of the relative displacement does not change significantly in the entry transient phase (A-B in Figure 8) with increasing rotational speed. This is because the relative contact deformation between a rolling element and both raceways in the load zone depends only on the applied load at the considered speed [3], as was shown in Figure 7(b). Therefore, the angular extent that the center of a rolling element travels from the entry point of the defect until it de-stresses completely is not dependent on the rotational speed. The increase in the acceleration response due to the increasing rotational speed can be explained by the fact that the rolling element has to de-stress and re-stress faster in the transient phases at the entry and exit points of the defect when rotational speed increases.
Investigation of the effect of speed and load on the angular extents of the low frequency entry and exit transient events has not been done previously. The experimental results show strong dependency of the angular extent of the low frequency entry and exit transient events on the applied load. It is evident that changes in the speed have almost no effect on angular extents of the low frequency entry and exit transient events.

5. Proposed defect size estimation method

5.1 Formulation

In this section, a method is proposed to estimate the size of a defect in a bearing based on measured acceleration signals. The observations presented in the previous section revealed the following discrepancies in the assumptions used in defect size estimation algorithms in previous studies:

- It was generally accepted that the first peak of the low frequency oscillation on the vibration response corresponds to the time when a rolling element strikes the entry point of the defect. Thus, the peak of the low frequency oscillation was used in previous defect size estimation methods. However, this study shows that the corresponding point on the vibration response to the entry point of the defect is the rising point of the first low frequency event in the vibration response.

- It was assumed that the rolling element strikes the exit point when the center of a rolling element is mid-way through the defect. This assumption requires the rolling element to maintain contact with both raceways, where the center of the rolling element is in the middle of the defect. However, this study has revealed that the rolling element de-stresses long before it reaches the exit point. Although the previously held assumption might be valid for extremely small defects (smaller than the contact deformation patch), application of the assumption produces increasing errors in defect size estimation as the defect grows.

- It was assumed that the highest peak amongst the high frequency responses in the multiple impact region corresponds to the time that the rolling element is at the exit point of the defect. However, our results suggest that this assumption is not necessarily valid, as the highest high frequency event may happen at the time when the rolling element is not at the transient phase.

The above findings explain the origin of inaccuracies of previous algorithms and are instrumental for the development of an improved algorithm to accurately estimate the size of a defect. Rather than the high-frequency event used by others in previous works, the low-frequency event at the exit to estimate defect size is used in the proposed defect size estimation method in this study.

Figure 9(a) and (c) show the acceleration response and the relative inner-to-outer raceway displacement of the defective bearing with a 0.103 rad angular extent, respectively. To enhance the entry and exit events, the acceleration signal was pre-whitened and then low-pass filtered at 5kHz to remove high frequency resonances excited by the impact event. Finally, the squared enveloped signal was calculated using the Hilbert transform, as shown in Figure 9(b).
Figure 9: Experimentally measured response of a bearing with a 0.103 rad extent defect at 5kN load and a rotational speed of 10Hz (a) Acceleration signal (b) Enhanced acceleration signal (c) The relative displacement between the inner and outer raceways.

Locating the exact starting point of the entry transient phase on the acceleration signal, and locating the exact ending point of the exit transient phase, is often not possible due to the presence of background noise or the vibrations of other components. Therefore, it may not be practical to use these points to estimate the defect size in situation where there is significant signal noise. Thus, we propose the following algorithm to estimate the defect size accurately by detection of the local maxima in the acceleration signal.
The angular extent between the two maxima points of the first and second rising events on the enhanced acceleration signal was measured as shown as $\beta_1$ in Figure 9(c). This distance corresponds to the angular travel of the center of the rolling element between the time that it completely de-stresses at the entry of the defect and the time when it starts the re-stressing phase.

The de-stressing and re-stressing angular travel of the center of a rolling element $\beta_2$, depends on the maximum total elastic contact deformation that a rolling element requires to lose in the transient phase at the entry and exit. The angular travel $\beta_2$, can either be simulated using the mathematical model presented in Moazen Ahmadi et al.[4, 10] or estimated by geometrically relating the relative elastic contact deformation to the geometry of the bearing using the following equation:

$$\beta_2 = \sqrt{\frac{2r_b \delta_{max} - \delta_{max}^2}{(r_b + R)^2}}$$

where $R$ is the radius of the outer raceway, $r_b$ is the radius of the rolling element and $\delta_{max}$ can be calculated using Eq. (1). Therefore, the total length of a defect can be estimated by:

$$L = \sin(2\beta_2 + \beta_1) \times R$$

5.2 Validation

The method described in Section 5.1 was used to measure the angle $\beta_2$, and then estimate the defect size in three bearings with nominal angular extents of 0.138, 0.103 and 0.068 (measured circumferential lengths of 3.11mm, 2.24mm and 1.61mm defects respectively), as described in Section 3. The enhanced vibration response, described in Section 5.1, was used to measure the angle $\beta_1$. The angle $\beta_2$ is calculated using the Eq.(2). The size of the defect is calculated using Eq. (3). Mean and standard deviations for a number of events on the vibration response of the three defective bearings for various speeds and loads were estimated as shown in Figure 10. The same events were used to estimate the size of defects using the method suggested by Sawalhi and Randall [6] for comparison. In addition, the average distance (rad) between the low and high frequency local maximum levels measured on the vibration response ($\beta_2$), for each case, is shown for reference in Figure 10.
Figure 10: Comparison between the estimated mean and standard deviation of outer race defects using the method suggested by Sawalhi and Randall [6], the method proposed in this paper, and the angular extent between the two maxima points using the enhanced acceleration signal described in Section 5.1 for: a) Actual defect size of 3.11mm and load of 2.5kN, b) Actual defect size of 3.11mm and load of 5kN, c) Actual defect size of 2.24mm and load of 2.5kN, d) Actual defect size of 2.24mm and load of 5kN, e) Actual defect size of 1.61mm and load of 2.5kN, f) Actual defect size of 1.61mm and load of 5kN.

The defect size estimates based on the method presented in this paper are more accurate (closer to the actual defect size as shown in Figure 10) and have smaller standard deviations when compared with the previous methods. The importance of including the applied load in the estimation algorithms is evident by comparing the estimations of the two methods and the length of $\beta_2$ in the higher applied load cases. Since
the angular extent between the end of the entry-transient phase and the start of the exit-transient phase depends on the applied load for a defective bearing, estimation algorithms that do not consider the applied load exhibit greater errors when applied load increases.

Comparison of estimation results for the three defective bearings demonstrates that the effect of omitting the angular extents of the entry and exit transients is greatest for larger defects. In larger defects these angular extents are less significant than the angular extent between the end of the entry-transient phase and the start of the exit-transient phase. For extremely small defects, of an order smaller than the patch of the relative deformation between a rolling element and the raceways, where the rolling element does not completely de-stress before hitting the exit point, the method suggested by Sawalhi and Randall [6] could be used. The method of using only the measured time between the low and high frequency local maximums on the vibration response for the estimation and ignoring the entry and exit transient times could be sufficient for larger defects. The method proposed in this study is found to be more accurate for both small and large defects.

6. Conclusions

This paper describes the experimental measurements of acceleration and relative displacement between the raceways of rolling element bearings with machined defects of various sizes that are tested at various speeds and loads. The effects of entry and exit events on the vibration response were explained. The probable path of a rolling element in the defect zone was studied indirectly by experimental investigation of the vibration and displacement responses. The experimental results showed that:

- exit impacts do not necessarily occur at the moment that the rolling element in the defect zone is midway through a defect. The exit impact events shown by acceleration impulses can occur before the center of a rolling element reaches the end of a defect and are not apparent in the relative displacement measurements.

- the relative displacement measurements indicate that the entry transient event to the defect starts before the local maxima of the low frequency response.

- the angular extent for a rolling element to de-stress or re-stress between the raceways increases with increasing applied load. These angular extents are not speed dependent. Therefore the relative angular extents between the low frequency entry and exit transient events decrease with increasing load.

- the assumption used in previous defect size estimation methods is only valid for very small defects, of an order smaller than the patch of the relative deformation between a rolling element and the raceways. Ignoring the effect of the applied load, which is done in other defect sizes estimation methods, causes errors in the estimate of larger defect sizes.

Based on the experimental results, the characteristics of the vibration response in defective bearings are categorized into different events which are related to different stages of the probable travel path of a rolling element in the defect zone. A method for defect size estimation based on an explanation of the path of a rolling element was presented. It was shown that the proposed method estimates the bearing’s defect size more accurately compared to previously proposed methods.
References


