Analysis of bearing stiffness variations, contact forces and vibrations in radially loaded double row rolling element bearings with raceway defects

Dick Petersen\textsuperscript{a}, Carl Howard\textsuperscript{a}, Nader Sawalhi\textsuperscript{b}, Alireza Moazen Ahmadi\textsuperscript{a}, Sarabjeet Singh\textsuperscript{a}

\textsuperscript{a}School of Mechanical Engineering, The University of Adelaide, Adelaide SA 5005, Australia
\textsuperscript{b}Mechanical Engineering Department, Prince Mohammad Bin Fahd University, Al Khobar 31952, Kingdom of Saudi Arabia

Abstract
A method is presented for calculating and analyzing the quasi-static load distribution and varying stiffness of a radially loaded double row bearing with a raceway defect of varying depth, length, and surface roughness. The method is applied to ball bearings on gearbox and fan test rigs seeded with line or extended outer raceway defects. When balls pass through the defect and lose all or part of their load carrying capacity, the load is redistributed between the loaded balls. This includes balls positioned outside the defect such that good raceway sections are subjected to increased loading when a defect is present. The defective bearing stiffness varies periodically at the ball spacing, and only differs from the good bearing case when balls are positioned in the defect. In this instance, the stiffness decreases in the loaded direction and increases in the unloaded direction. For an extended spall, which always has one or more balls positioned in the defect, this results in an average stiffness over the ball spacing period that is lower in the loaded direction in comparison to both the line spall and good bearing cases. The variation in bearing stiffness due to the defect produces parametric excitations of the bearing assembly. The qualitative character of the vibration response correlates to the character of the stiffness variations. Rapid stiffness changes at a defect exit produce impulses. Slower stiffness variations due to large wavelength waviness features in an extended spall produce low frequency excitation which results in defect components in the velocity spectra. The contact forces fluctuate around the quasi-static loads on the balls, with rapid stiffness changes producing high magnitude impulsive force fluctuations. Furthermore, it is shown that analyzing the properties of the dynamic model linearized at the quasi-static solutions provides greater insight into the time-frequency characteristics of the vibration response. This is demonstrated by relating the low frequency event that occurs when a ball enters a line spall to the dynamic properties of the bearing assembly.

Keywords: Rolling element bearing, bearing stiffness, defect, varying compliance, vibration, condition monitoring

1. Introduction

1.1. Background
Rolling element bearings are used in a wide variety of rotating machinery and their failure is one of the most common reasons for machine breakdowns. Vibration-based condition monitoring of bearings forms an essential part of any condition monitoring program, and accurate diagnosis improves safety and reduces machinery downtime and maintenance costs. Excessive bearing vibrations can be caused by distributed defects, such as surface roughness [1, 2], waviness [3–13], misaligned races, and off-size rolling elements [3], or localized defects caused by pitting or spalling of the raceways or rolling elements [14–30]. Raceway spalls generally develop due to a fatigue mechanism [31] and are a common failure mode in rolling element bearings [32]. This paper considers the vibrations generated by line and extended raceway spalls.

Line spalls are normally initiated by sub-surface fatigue cracks which appear after some time even when the bearing is properly lubricated, aligned, and loaded. The sub-surface cracks generate stress-waves that may be detectable in the ultrasonic frequency range using acoustic emission [33], Shock Pulse Method [34], or Spike Energy sensing techniques [35]. Eventually, the sub-surface cracks grow and break through to the surface causing a spall or crack. As the rolling elements pass through the spall, abrupt changes in the contact forces occur which excite the high-frequency resonances of the bearing. This results in impulses in the vibration response which can be detected using envelope
which will provide additional insights into the vibration response predicted by these models [21–29].

response, because a method for doing so was not available. The method developed here can be used for this purpose, non-linear contact springs, the actual sti
tently considered in previous multi-body nonlinear dynamic models of defective bearings [21–29], i.e. via inclusion of
damping ratios of the linearized dynamic model, greater insight into the time-frequency characteristics of the
linearizing the nonlinear dynamic model at the calculated quasi-static solutions and analyzing the natural frequencies
of the dynamic contact forces predicted by the model is presented for the line and extended spall cases, which has
not been reported in previous research into multi-body nonlinear dynamic models of defective bearings [21–29].

The majority of vibration models of defective bearings that have been developed [14–30] consider the line spalls
that occur in the early stage of bearing failure. The extended spalls that occur during the later stage of failure, which
cover a larger section of the raceway but not the entire raceway like distributed defects [1–13], have received much less
attention [23, 24], perhaps because early detection at the line spall stage is more important. Although it is commonly
known in the vibration condition monitoring industry that defect components start to appear in the velocity spectrum as
the spall grows from a line into an extended spall [38], the mechanism by which this occurs has not been investigated.

It is generally accepted that for a bearing with a line spall, a high frequency impulsive event in the measured
vibration response corresponds to the excitation of a high-frequency bearing resonant mode which occurs when a
rolling element exits the spall. However, the low frequency event that occurs when a rolling element enters a line spall
reported in previous research [39] has not been related to the dynamic properties of the bearing assembly.

1.2. Contribution and method

The main contribution of this paper is the development of a method for calculating and analyzing the quasi-static
load distribution and varying stiffness of a radially loaded double row bearing with a raceway defect of varying depth,
length, and surface roughness. The method forms an extension to the method for defect-free single row bearings de-
veloped in previous research into so-called varying compliance vibrations [40–43]. The calculated stiffness variations
result in a parametric excitation of the defective bearing assembly, and play an important role in the generation of
the resulting vibration response. The stiffness variations are generally much larger than for a defect-free bearing, and
depend mostly on the defect geometry but also on the static load, the clearance, and the bearing macro-geometry.
Differences in the vibration signature produced by spalls of different geometry can be correlated to differences in the
character of the bearing stiffness variations.

The developed analysis method is illustrated by predicting and analyzing the stiffness variations for defective
 bearings on gearbox and fan test rigs, which are seeded with line or extended spalls. The predicted bearing stiffness
 variations are then correlated to the measured and modeled vibration responses. The modeled vibration responses for
the test rigs are generated using a previously developed multi-body nonlinear dynamic model of a defective single
row bearing [23, 24], which is enhanced and extended in Section 3 to a double row bearing. A detailed analysis
of the dynamic contact forces predicted by the model is presented for the line and extended spall cases, which has
not been reported in previous research into multi-body nonlinear dynamic models of defective bearings [21–29].
Greater insight into the contact forces is obtained by comparing them to the quasi-static load variations on the rolling
elements, which are calculated using the developed method. The analysis method can be applied to previous multi-
body nonlinear dynamic models of defective bearings [21–29], as well as more comprehensive explicit dynamic finite
element models [44], to improve the understanding of the predicted contact forces. Additionally, it is shown that by
linearizing the nonlinear dynamic model at the calculated quasi-static solutions and analyzing the natural frequencies
and damping ratios of the linearized dynamic model, greater insight into the time-frequency characteristics of the
modeled vibration response is achieved. This is demonstrated by relating the low frequency event that occurs when a
rolling element enters a line spall [39] to the dynamic properties of the bearing assembly.

The methods and results presented in this paper contribute to the understanding of the varying stiffness excitations
in a defective bearing and their effect on the measured vibration response. Although the varying stiffness was inher-
ently considered in previous multi-body nonlinear dynamic models of defective bearings [21–29], i.e. via inclusion of
non-linear contact springs, the actual stiffness variations were not analyzed and correlated to the measured vibration
response, because a method for doing so was not available. The method developed here can be used for this purpose,
which will provide additional insights into the vibration response predicted by these models [21–29].

Besides the varying stiffness excitation mechanism considered here, the other important excitation mechanism
in defective bearings is the impact force caused by a rolling element mass striking the bottom or edge of a defect.
Since rolling element inertia is not included in the model considered here, this excitation mechanism is not modelled. This simplification is warranted by previous research [45] which found that the effect of rolling element inertia was not significant at the low speeds considered here. Furthermore, the reasonable agreement between the measured and modelled vibration responses presented here, as well as in previous work that did not consider rolling element inertia [21–24], further suggests that this simplification is reasonable at lower speeds. Nevertheless, the relative importance of the two excitation mechanisms as a function of speed, load, clearance, defect geometry, etc., is an important scientific question that needs to be addressed to advance our understanding of vibrations generated in defective bearings. The analysis method developed here will be of great benefit to future research aiming to answer this question.

1.3. Outline

This paper is outlined as follows. Section 2 reviews the available literature on vibration models for rolling element bearings with raceway defects. The multi-body nonlinear dynamic model of a defective bearing that is used to model the vibration response measured on the gearbox and fan test rigs is introduced in Section 3. Section 4 presents the method for calculating the quasi-static load distribution and varying stiffness of a radially loaded double row bearing with a raceway defect of varying depth, length, and surface roughness. This section also discusses the linearization of the nonlinear model at the quasi-static solutions, and the dynamic properties of the resulting linearized dynamic system. In Section 5, the developed methods are applied to the gearbox test rig which includes a bearing seeded with a line or extended outer raceway spall. The predicted quasi-static load distributions and bearing stiffness variations are presented and analyzed, and the results are correlated with the differences in the modeled and measured vibration responses for the line and extended spalls. The dynamic contact forces predicted by the model are compared to the quasi-static loads on the rolling elements as this provides more insight into the generated contact forces. Section 6 compares the modeled and measured vibration response for the fan test rig with a bearing seeded with a line spall. The low and high frequency events that occur when a rolling element passes through the defect [39] occur in both the modeled and measured response. The time-frequency characteristics are investigated and the low frequency event is correlated to the dynamic properties of the bearing assembly. Concluding remarks are presented in Section 7.

2. Review of vibration models for bearings with raceway defects

2.1. Raceway waviness models

The vibrations produced by raceway waviness have been investigated by a number of authors [1–13, 25]. The term waviness normally refers to surface features with wavelengths that are greater than the width of the Hertzian contact patch between the rolling elements and raceways. Surface features with smaller wavelengths are referred to as roughness. The vibrations produced by surface roughness have been investigated [46] but the longer wavelength features of waviness generally have a more significant effect on low-frequency vibration levels [10]. The vibration models for raceway waviness assume that the rolling elements are always loaded when positioned in the load zone. This assumption is either included explicitly in the model through the derivation of a forcing function based on the waviness profile [1–10], or implicitly in the predicted vibrations by pre-loading the rolling elements in the model [11, 12] or considering waviness profiles of small amplitudes [13]. The waviness was also assumed to be distributed over the entire raceway in the developed models [1–13]. Therefore, the two significant differences between the waviness models and the modeling considered here are that the rolling elements may become unloaded when positioned in the defect, and the defect extends across a section of the raceway rather than the entire raceway.

2.2. Impulse train models for localized spalls

A number of impulse train vibration models [14–17] for single or multiple raceway line spalls have been developed. The excitation mechanism in these models is a series of force impulses exciting the resonant modes of the inner and outer raceways, which are modeled as circular rings [47]. The models can be used to predict the characteristic defect frequencies and amplitudes that occur in the measured vibration spectra, taking into account the applied loading, the transmission paths between the defect location and transducer, and the exponential decay of a resonant mode excited by an impulsive force. The effect of the shape of the force impulse has also been considered [17]. The impulse train models were extended [18–20, 30] to incorporate the slight random variations that occur between the impulses.
due to slippage of the rolling elements. These random variations cause a smearing of the harmonics of the defect components in the vibration spectrum. As opposed to the model presented here, the impulse train models do not model the varying compliance of the bearing assembly and cannot be used for simulating the time-frequency response of the nonlinear dynamic bearing system.

2.3. Multi-body nonlinear dynamic models

Multi-body nonlinear dynamic models of rolling element bearings have been developed for predicting the time-domain vibration response due to a line spall [21–29]. The inner raceway is usually modeled as a lumped mass while the outer raceway is either modeled as fixed and thus massless [25, 27, 28], a lumped mass [21–24, 26], or by finite elements to include its flexibility in the radial direction [29]. Most models are two-dimensional and only consider displacements in the radial direction [23, 24, 26–29] while some also consider axial displacements [25] and rotations [21, 22]. The nonlinearity in the models arise by modeling the Hertzian contacts between the rolling elements and the raceways as nonlinear springs. Damping is included either locally in the contacts between the rolling elements and raceways [25, 26], or globally by attaching a grounded damper to the inner raceway [21–24, 27, 28]. The mass of the rolling elements is included in some models [25, 26, 29] but excluded in others [21–24, 27, 28] because the effect of rolling element inertia is only significant at high speeds [45]. Some models consider the slippage of the rolling elements [23, 24, 29] which results in predicted vibration spectra that have a closer resemblance to typical measured vibration spectra for point defects. The modeling results presented in the previous research [21–29] generally include the acceleration time traces and the corresponding envelope, acceleration, and/or velocity spectra for a raceway line spall. The contact forces, which are inherently predicted by the previous models, and the varying stiffness of the bearing assembly have not been analyzed.

The extended spalls that occur when a line spills grows to cover a larger section of the raceway have only been considered by Sawalhi & Randall [24]. They further developed their two-dimensional multi-body dynamic model for line spalls [23] to enable modeling of extended spalls of varying circumferential length, depth, and surface roughness. The inner and outer raceways were modeled as rigid bodies connected by nonlinear contact springs representing the Hertzian contacts between the rolling elements and raceways. The mass of the rolling elements was not considered because previous research [45] showed that its effect was minimal at the low run speeds considered in the experiments. Slippage of the rolling elements was included to obtain a closer resemblance between the predicted and measured vibration spectra. Damping was included via a grounded damper attached to the inner raceway. A mass-spring-damper system representing a typical high-frequency bearing resonance was attached to the outer raceway. The general aim of the authors [23, 24, 48] was the differential diagnosis of gear and bearing defects, which was achieved by utilizing the difference in the cyclostationary [36] properties of the gear and bearing signals. The presented results included predicted acceleration signals for inner and outer raceway extended spills and their corresponding envelope, acceleration, and cyclic spectral densities. The predicted results were compared to experimental results measured on a gear test rig with a defective bearing, and reasonable agreement was achieved. The contact forces predicted by the model were not presented and analyzed.

The developed method for calculating and analyzing the stiffness variations of a defective bearing can be applied to the previously developed multi-body nonlinear dynamic models [21–29]. This will provide more insight into the time-frequency characteristics of the contact forces and vibration response predicted by these models.

3. Multi-body nonlinear dynamic model of a bearing with a raceway defect

3.1. Diagram of model

Figure 1 presents a diagram of the multi-body nonlinear dynamic model that is used to model the vibration response of a rolling element bearing with a raceway defect. The model is based on the model developed by Sawalhi & Randall [23, 24]. The model includes a mass \( m_i \) representing the inner raceway plus shaft, a mass \( m_o \) representing the outer raceway plus bearing support structure, and two masses \( m_r \) attached to the outer raceway via a spring \( k_r \) and damper \( c_r \) representing a typical measured high-frequency resonant response of the bearing. The six degrees-of-freedom included in the model are the inner raceway displacements \( x_i(t) \) and \( y_i(t) \), the outer raceway displacements \( x_o(t) \) and \( y_o(t) \), and the measured vibration response \( x_r(t) \) and \( y_r(t) \). The spring and damper constants \( k_{oi}, k_{oy}, c_{oi} \) and \( c_{oy} \) represent the stiffness and damping of the bearing support structure. These parameters are adjusted to describe a
low-frequency rigid-body mode of the bearing support structure. The Hertzian contacts between the rolling elements and raceways are modeled by time-varying nonlinear contact spring \( k_{c}(t) \). The dampers \( c_{p}(t) \) model the lubricant film damping in the contacts.

Four modifications were made to the model developed by Sawalhi & Randall [23, 24] to improve the analysis of the vibration response for extended spalls: (1) An additional mass-spring-damper system was attached to the outer raceway to enable prediction of the measured high-frequency resonant response in both radial directions; (2) the global damping that was included via a grounded viscous damper attached to the inner raceway is replaced by the dampers \( c_{p}(t) \) that model the lubricant film damping in the contacts [21, 22]; (3) a static radial load \( W \) is applied to the inner raceway; and (4) the grounded spring attached to the inner raceway is removed in order to obtain quasi-static load distributions for the defect-free case that agree with the well-known Stribeck results [49].

### 3.2. Kinematics including slippage of rolling elements

The shaft in Figure 1 rotates at a run speed \( \omega_{r} = 2\pi f_{s} \). For a bearing pitch diameter \( D_{p} \), a loaded contact angle \( \alpha \), and a rolling element diameter \( D_{b} \), the resulting nominal cage speed \( \omega_{c} = 2\pi f_{s} \) is given by

\[
\omega_{c} = \omega_{r} \left( 1 - \frac{D_{b} \cos \alpha}{D_{p}} \right)
\]  

(1)

The position \( \phi_{p}(t) \) of rolling element \( j \) on row \( r \) shown in Figure 1 is defined as

\[
\phi_{p}(t) = \phi_{r}(t) + \frac{2\pi(j - 1)}{N_{b}} + \frac{(r - 1)\pi}{N_{b}}, \quad j = 1 \text{ to } N_{b}, \quad r = 1, 2
\]  

(2)

with \( \phi_{r}(t) \) the cage position given by

\[
\phi_{r}(t + dt) = \phi_{r}(t) + \omega_{c} dt + \nu(t)
\]  

(3)

where \( \nu(t) \) is a random process uniformly distributed between \([-\phi_{slip}, \phi_{slip}]\). This process accounts for slippage of the rolling elements, and typical values for the maximum phase variation \( \phi_{slip} \) are of the order 0.01 – 0.02 radians [23].
Equations (2) and (3) assume that the two rows share the same staggered cage, which is typically the case in double row self-aligning bearings, and defines the cage position to coincide with rolling element $j = 1$ on row $r = 1$. For double row bearings that do not share the same cage, cage positions $\phi_{cr}(t)$ can be generated for the individual rows using Equation (3).

### 3.3. Hertzian contact model

The contact deformation $\delta_{cr}(t)$ for rolling element $j$ on row $r$ is a function of the relative displacements $\delta_x(t)$ and $\delta_y(t)$ of the inner and outer raceways, the position $\phi_{cr}(t)$ of the rolling element, the radial clearance $c$, and the defect depth profile $C_r(\phi_{cr}(t))$ at the rolling element position, and is given by

$$
\delta_{cr}(t) = \delta_x(t) \cos \phi_{cr}(t) + \delta_y(t) \sin \phi_{cr}(t) - C_r(\phi_{cr}(t)) - c
$$

where the relative displacements of the inner and outer raceways are defined as

$$
\delta_x(t) = x_i(t) - x_o(t), \quad \delta_y(t) = y_i(t) - y_o(t)
$$

By allowing separate defect depth profiles $C_r(\phi)$ for the two rows, scenarios with defects on one or both rows can be modeled. The generation of the defect depth profiles is discussed in Section 3.6. The Hertzian contact force $F_{jr}(t)$ associated with the contact deformation $\delta_{jr}(t)$ acts on the inner and outer raceways in the radial direction, and is given by

$$
F_{jr}(t) = K_{jr}(t) \gamma_{jr}(t) = k_{jr}(t) \delta_{jr}(t) \gamma_{jr}(t)
$$

with $\gamma_{jr}(t) = \begin{cases} 1 & \text{if } \delta_{jr}(t) > 0 \\ 0 & \text{if } \delta_{jr}(t) \leq 0 \end{cases}$.

In Equations (6) and (7), the load-deflection factor $K$ (units of N/m$^3$) depends on the curvatures and material properties of the surfaces in contact, and the load deflection parameter $n$ equals 3/2 for ball bearings and 10/9 for roller bearings [49]. Summing the $x$ and $y$ components $F_{xjr}(t)$ and $F_{yjr}(t)$ of the radial contact forces $F_{jr}(t)$ over all rolling elements on both rows results in the net contact forces $F_x(t)$ and $F_y(t)$ acting on the inner and outer raceways, such that

$$
\begin{bmatrix}
F_x(t) \\ F_y(t)
\end{bmatrix} = \sum_{r=1}^{N_r} \sum_{jr=1}^{N_{jr}} \begin{bmatrix}
F_{xjr}(t) \\ F_{yjr}(t)
\end{bmatrix} = \sum_{r=1}^{N_r} \sum_{jr=1}^{N_{jr}} F_{jr}(t) \begin{bmatrix}
\cos \phi_{cr}(t) \\ \sin \phi_{cr}(t)
\end{bmatrix} = \sum_{r=1}^{N_r} \sum_{jr=1}^{N_{jr}} k_{jr}(t) \delta_{jr}(t) \gamma_{jr}(t) \begin{bmatrix}
\cos \phi_{cr}(t) \\ \sin \phi_{cr}(t)
\end{bmatrix}.
$$

The contact forces defined in Equation (8) are parametric excitations caused by the time-varying nature of the nonlinear contact spring stiffness $k_{jr}(t)$.

### 3.4. Damping in rolling element contacts

To account for lubricant film damping in the contacts between the rolling elements and raceways [21], the dampers $c_{jr}(t)$ are included in the model as illustrated in Figure 1. The contact damping force $F_{djr}(t)$ associated with a rolling element located at $\phi_{cr}(t)$ acts on the inner and outer raceways in the radial direction and is given by

$$
F_{djr}(t) = c \dot{\delta}_{jr}(t) \gamma_{jr}(t) = c_{jr}(t) \dot{\delta}_{jr}(t)
$$

with $\gamma_{jr}(t) = \begin{cases} 1 & \text{if } \delta_{jr}(t) > 0 \\ 0 & \text{if } \delta_{jr}(t) \leq 0 \end{cases}$.

where $c$ is the viscous contact damping constant. The total contact damping forces $F_{dx}(t)$ and $F_{dy}(t)$ acting on the inner and outer raceways in the $x$ and $y$ direction are the sum of the radial damping forces over the $N_r$ rolling elements, such that

$$
\begin{bmatrix}
F_{dx}(t) \\ F_{dy}(t)
\end{bmatrix} = \sum_{r=1}^{N_r} \sum_{jr=1}^{N_{jr}} \begin{bmatrix}
F_{dxjr}(t) \\ F_{dyjr}(t)
\end{bmatrix} = \sum_{r=1}^{N_r} \sum_{jr=1}^{N_{jr}} F_{djr}(t) \begin{bmatrix}
\cos \phi_{cr}(t) \\ \sin \phi_{cr}(t)
\end{bmatrix} = c \sum_{r=1}^{N_r} \sum_{jr=1}^{N_{jr}} \dot{\delta}_{jr}(t) \gamma_{jr}(t) \begin{bmatrix}
\cos \phi_{cr}(t) \\ \sin \phi_{cr}(t)
\end{bmatrix}.
$$

Damping in the bearing assembly is typically in the order of $0.25-2.5 \times 10^{-5}$ times the linearized stiffness of the bearing assembly [21]. The linearized stiffness and damping of the bearing assembly can be calculated as described in Section 4, and the viscous contact damping constant $c$ is adjusted to achieve damping within the specified range.
3.5. Nonlinear equations of motion

Using Equations (4) to (10) and the diagram of the bearing model illustrated in Figure 1, the nonlinear equations of motion for the inner raceway \( m_i \), outer raceway \( m_o \), and high-frequency resonant mode \( m_r \) are now given by

\[
M \ddot{x}(t) + C \dot{x}(t) + K x(t) + \sum_{j=1}^{2} \sum_{i=1}^{N} \left[ k_{ij}(t) \delta_{ij}(t) + c \dot{\delta}_{ij}(t) \right] y_{ij}(t) R_{ij}(t) = 0
\]  

(11)

where the state vector \( x(t) \) is defined as

\[
x(t) = \begin{bmatrix} x_i(t) & y_i(t) & x_o(t) & y_o(t) & x_r(t) & y_r(t) \end{bmatrix}^T.
\]  

(12)

The mass, stiffness and damping matrices in Equation (11) are formulated as

\[
M = \begin{bmatrix} M_i & 0 & 0 \\ 0 & M_o & 0 \\ 0 & 0 & M_r \end{bmatrix}, \quad K = \begin{bmatrix} 0 & 0 & 0 \\ 0 & K_o + K_r & -K_r \\ 0 & -K_r & K_r \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & C_o + C_r & -C_r \\ 0 & -C_r & C_r \end{bmatrix}
\]  

(13)

with

\[
M_i = \begin{bmatrix} m_i & 0 \\ 0 & m_i \end{bmatrix}, \quad M_o = \begin{bmatrix} m_o & 0 \\ 0 & m_o \end{bmatrix}, \quad M_r = \begin{bmatrix} m_r & 0 \\ 0 & m_r \end{bmatrix}
\]  

(14)

and

\[
K_o = \begin{bmatrix} k_{ox} & 0 \\ 0 & k_{oy} \end{bmatrix}, \quad C_o = \begin{bmatrix} c_{ox} & 0 \\ 0 & c_{oy} \end{bmatrix}, \quad K_r = \begin{bmatrix} k_r & 0 \\ 0 & k_r \end{bmatrix}, \quad C_r = \begin{bmatrix} c_r & 0 \\ 0 & c_r \end{bmatrix}
\]  

(15)

The matrix \( R_{ij}(t) \) in Equation (11) defines a transformation from orthogonal to radial coordinates and is given by

\[
R_{ij}(t) = \begin{bmatrix} \cos \phi_{ij}(t) & -\cos \phi_{ij}(t) & \sin \phi_{ij}(t) & -\sin \phi_{ij}(t) & 0 & 0 \end{bmatrix}^T.
\]  

(16)

The dynamic system in Equation (11) is excited parametrically by the time-varying nonlinear contact stiffnesses \( k_{ij}(t) \) defined by Equation (7), with the stiffness increasing and decreasing as the contact deformations \( \delta_{ij}(t) \) increase and decrease.

3.6. Generation of the defect depth profile

In the diagram of the bearing model shown in Figure 1, the blue and red lines indicate the extent of an inner or outer raceway defect of circumferential length \( \Delta \phi \) and centered at an angle \( \phi \). The defect location \( \phi \) is constant for an outer raceway defect but rotates at the shaft speed \( \omega \), for an inner raceway defect, such that \( \phi(t) = \phi(0) + \omega t \).

The defect depth profiles representative for the line and extended spalls considered in the gearbox test rig simulations and experiments presented in Section 5 are illustrated in Figure 6. Generally, in the absence of an accurately measured defect depth profile provided by a laser surface scanner, the defect depth profile \( C_r(\phi) \) can be generated as [24]

\[
C_r(\phi) = C_{rw}(\phi) + C_{rv}(\phi)
\]  

(17)

where \( C_{rw}(\phi) \) defines the average defect depth \( d \) and the sharpness of the defect entrance and exit, and \( C_{rv}(\phi) \) defines the surface waviness in the defect. For the component \( C_{rw}(\phi) \), the section outside the defect is set to zero and the section within the defect is generated using a Tukey window with a height equal to the defect depth \( d \). The angular extent \( \Delta \phi_{ew} \) and \( \Delta \phi_{ew} \) of the cosine tapered sections of the window determine the sharpness of the defect entrance and exit. A sharp exit results in a rapidly changing contact stiffness \( k_{ij}(t) \) and the associated parametric excitation produces and impulsive vibration response.

The surface roughness component \( C_{rv}(\phi) \) in Equation (17) is generated such that it only contains waviness features with wavelengths that are greater than the width of the Hertzian contact patch between the rolling elements and raceways. This is achieved by low-pass filtering (in the spatial domain) a uniformly distributed random noise sequence. The cut-off wave number is set to correspond to the approximate width of the Hertzian contact patch such that short wavelength roughness features are removed. In the simulations and experimental results presented in Section 5, the applied loading and geometrical properties of the ball bearing are such that the width of the Hertzian contact patch is estimated to be in the order of 0.2 mm [49].
4. Quasi-static load distributions and stiffness variations in a defective bearing

This section presents the method for calculating the quasi-static load distribution and varying stiffness of a radially loaded double row bearing in the presence of a raceway defect. The term quasi-static is used to indicate that a variable is calculated as a function of cage position while assuming the cage is not rotating. Once the quasi-static solutions are known, the multi-body nonlinear dynamic model introduced in Section 3 is linearized at these solutions and the dynamic properties of the resulting linearized dynamic system are derived. This analysis improves the understanding of the time-frequency characteristics of the vibration response predicted by the nonlinear dynamic model.

4.1. Quasi-static load distributions in a defective bearing

For a defective bearing subjected to a static load \( W \) in the \( y \) direction, static equilibrium is achieved when the net contact force \( F_r(t) \) defined in Equation (8) equals the applied load, and the net contact force \( F_x(t) \) equals zero. Using Equations (4) to (8), the quasi-static loads on the rolling elements are thus found by solving the following set of nonlinear algebraic equations as a function of the cage position \( \phi_c \)

\[
K \sum_{r=1}^{2N_c} \sum_{j=1}^{N_r} \begin{bmatrix} \delta_x \cos \phi_{jr} + \delta_y \sin \phi_{jr} - C_r(\phi_{jr}) - c \sin \gamma_{jr} \cos \phi_{jr} \\ \delta_x \cos \phi_{jr} + \delta_y \sin \phi_{jr} - C_r(\phi_{jr}) - c \sin \gamma_{jr} \sin \phi_{jr} \end{bmatrix} - \begin{bmatrix} 0 \\ W \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},
\]

where the rolling element positions \( \phi_{jr} \) depend on the cage position \( \phi_c \) as defined by Equation (2), and the quasi-static relative displacements of the inner and outer raceways that solve Equation (18) are defined as

\[
\delta_x = \bar{x}_i - \bar{x}_o, \quad \delta_y = \bar{y}_i - \bar{y}_o,
\]

where \( \bar{x}_i, \bar{y}_i, \bar{x}_o, \) and \( \bar{y}_o \) are the quasi-static displacements of the inner and outer raceways. A Newton-Raphson method can be used to solve Equation (18) at each considered cage position. Once the quasi-static relative displacements are solved as a function of the cage position, the quasi-static contact deformations are calculated using Equation (4) as

\[
\delta_{jr} = \delta_x \cos \phi_{jr} + \delta_y \sin \phi_{jr} - C_r(\phi_{jr}) - c, \quad r = 1, 2, \quad j = 1, 2, \ldots, N_b.
\]

Using the force-displacement relationship for the contact springs defined by Equation (6), the quasi-static contact forces are given by

\[
\bar{F}_{jr} = K \delta_{jr} \bar{y}_{jr} \quad \text{with} \quad \bar{y}_{jr} = \begin{cases} 1 & \text{if} \ \delta_{jr} > 0, \\ 0 & \text{if} \ \delta_{jr} \leq 0, \end{cases}, \quad r = 1, 2, \quad j = 1, 2, \ldots, N_b,
\]

and the net quasi-static contact forces in the \( x \) and \( y \) directions are defined as

\[
\begin{bmatrix} \bar{F}_x \\ \bar{F}_y \end{bmatrix} = \sum_{r=1}^{2N_c} \sum_{j=1}^{N_r} \begin{bmatrix} \bar{F}_{xr,j} \\ \bar{F}_{yr,j} \end{bmatrix} = \sum_{r=1}^{2N_c} \sum_{j=1}^{N_r} \bar{F}_{jr} \begin{bmatrix} \cos \phi_{jr} \\ \sin \phi_{jr} \end{bmatrix}.
\]

For a defect-free bearing, the quasi-static loads are found by solving Equation (18) while setting \( C_r(\phi) = 0 \). This results in the well-known load distribution defined by [49]

\[
\bar{F}_{jr} = \begin{cases} F_{max} \left[ 1 - \frac{1 + \sin \phi_{jr}}{2\epsilon} \right]^n & \text{if} \ \sin \phi_{jr} \leq 2\epsilon - 1, \\ 0 & \text{if} \ \sin \phi_{jr} > 2\epsilon - 1, \end{cases}
\]

where the load distribution factor \( \epsilon = 0.5 \) for zero clearance, \( 0 < \epsilon < 0.5 \) for positive clearance, and \( 0.5 < \epsilon < 1 \) for negative clearance. The maximum load \( F_{max} \) for a double row ball bearing with zero clearance is given by [49]

\[
F_{max} = \frac{4.37W}{2N_b \cos \alpha}
\]

where each row carries half the load \( W/2 \). For ball bearings with negative or positive clearance, the constant 4.37 in Equation (24) has to be decreased or increased, respectively.
4.2. Formulation of bearing stiffness and damping matrices

For double row bearings, the translational stiffnesses of each row act in parallel such that the quasi-static bearing stiffness matrix $K_b$ can be calculated by superposing the stiffness matrices $K_{br}$ for each row [50]. This means that each row is effectively modeled as a single row bearing on the same shaft and housing. The quasi-static bearing stiffness matrix can therefore be formulated as

$$K_b = \begin{bmatrix} k_{xx} & k_{xy} \\ k_{yx} & k_{yy} \end{bmatrix} = \sum_{r=1}^{2} K_{br} = \sum_{r=1}^{2} \begin{bmatrix} k_{xxr} & k_{xyr} \\ k_{yxr} & k_{yyr} \end{bmatrix}$$

(25)

where $k_{xx} = k_{xx1} + k_{xx2}$ and $k_{yy} = k_{yy1} + k_{yy2}$ define the radial bearing stiffness, and $k_{xy} = k_{xy1} + k_{xy2}$ defines the degree of cross-coupling. The quasi-static bearing stiffness matrix is now calculated by linearizing the force-displacement relationship defined by Equations (19) to (22) around the quasi-static relative displacements of the raceways calculated in the previous section, such that

$$K_b = \sum_{r=1}^{2} \sum_{j=1}^{N_r} \begin{bmatrix} \frac{\partial F_{jr}}{\partial x_{rr}} & \frac{\partial F_{jr}}{\partial y_{rr}} \\ \frac{\partial F_{jr}}{\partial x_{yr}} & \frac{\partial F_{jr}}{\partial y_{yr}} \end{bmatrix} = nK \sum_{r=1}^{2} \sum_{j=1}^{N_r} \ddot{y}_{jr} \begin{bmatrix} \cos^2 \phi_{jr} & \cos \phi_{jr} \sin \phi_{jr} \\ \cos \phi_{jr} \sin \phi_{jr} & \sin^2 \phi_{jr} \end{bmatrix}$$

(26)

The quasi-static stiffness matrix $K_b$ varies with cage position even for the case of a defect-free bearing with $C_r(\phi) = 0$, leading to the so-called varying compliance vibrations [40]. However, a raceway defect typically causes much larger variations in the quasi-static stiffness than the damping terms $c_{xx}$ and $c_{yy}$, respectively [21].

4.3. Linearized dynamic system

When analyzing the vibration response of the bearing assembly governed by the nonlinear equations of motion in Equation (11), it will prove insightful to linearize the equations of motion around the quasi-static displacements $\bar{x}$ that result when the bearing is subjected to a static load. These displacements are calculated as a function of the cage position as described in Section 4.1. The resulting linearized equations of motion are formulated as

$$M\ddot{x}(t) + C\dot{x}(t) + Kx(t) = 0$$

(28)

where $\ddot{x}(t) = x(t) - \bar{x}$ are small displacements around the quasi-static solution $\bar{x}$. The linearized stiffness and damping matrices in Equation (28) are given by

$$\tilde{K} = \begin{bmatrix} K_b & -K_b & 0 \\ -K_b & K_b + K_r & -K_r \\ 0 & -K_r & K_r \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} C_b & -C_b & 0 \\ -C_b & C_b + C_r & -C_r \\ 0 & -C_r & C_r \end{bmatrix}$$

(29)

where $K_b$ and $C_b$ are the quasi-static bearing stiffness and damping matrices defined in Equations (26) and (27) which vary with the cage position $\phi_c$. The undamped natural frequencies $\omega_n$ and damping ratios $ζ_n$ of the modes of the
linearized dynamic system defined by Equation (28) are calculated by solving the following generalized eigenvalue problem [51]

\[
\begin{bmatrix}
0 & -\bar{K} \\
-M & 0
\end{bmatrix}
\begin{bmatrix}
v \\
-\lambda
\end{bmatrix}
= \begin{bmatrix}
M & C \\
0 & M
\end{bmatrix}
\begin{bmatrix}
v \\
\lambda
\end{bmatrix}.
\]

(30)

This results in twelve eigenvalues \(\lambda_n\) (six complex conjugate pairs) and eigenvectors \(v_n\) from which the undamped natural frequencies and damping ratios are calculated as

\[
\omega_n = |\lambda_n|, \quad \zeta_n = \frac{|\text{Re}(\lambda_n)|}{|\lambda_n|}, \quad n = 1, 2, \ldots, 12
\]

(31)

The time-frequency characteristics of the modeled vibration response of the nonlinear dynamic system is better understood by analyzing the natural frequencies of the quasi-static linearized dynamic system defined in Equation (28). Note that these natural frequencies vary periodically at the ball pass frequency.

In the simulations and experiments presented in Sections 5 and 6, the static load \(W\) acts in the \(y\) direction which results in a cross-coupling stiffness term \(k_{yx}\) that is one to two orders of magnitude lower than the radial stiffness terms \(k_{xx}\) and \(k_{yy}\). In this instance, the undamped natural frequencies \(\omega_n\) of the linearized system are reasonably approximated by the natural frequencies of the uncoupled mass-spring-damper systems illustrated in Figure 2. The resonance frequency \(\omega_r = \sqrt{k_r/m_r}\) of the two single-degree-of-freedom systems in Figure 2 correspond to the high-frequency bearing resonant mode included in the model. Analytical solutions for the natural frequencies of the two-degrees-of-freedom systems can be found in reference [52]. The presented analysis is used in Section 6 to gain more insight into the low frequency event that occurs as a rolling element enters the line spall on the fan test rig [39].

5. Gearbox test rig simulations and experiments

5.1. Test rig and defect description

Figure 3 shows a photo of the gearbox test rig which has been used in a number of studies into the modeling and diagnostics of gear and bearing defects and their interaction [23, 24, 53–58]. The single stage gearbox includes a spur gear set with a 49:32 gear ratio. The input shaft is driven by a three-phase electric motor, and flywheels are included to minimize speed fluctuations. Loading is applied by a hydraulic motor and the load applied on the input shaft is measured by a torque transducer. Flexible couplings are used to minimize torsional vibrations of the shafts. The input and output shafts are supported by two Koyo 1205 double row self-aligning ball bearings, which have \(N_b = 12\) balls on each row, a pitch diameter \(D_p = 38.75\) mm, a ball diameter \(D_b = 7.12\) mm, and a contact angle \(\alpha = 0^\circ\). The
balls are retained in a staggered arrangement by a single cage, with the angular offset between the two rows equal to $\pi/N_b = 15^\circ$, i.e. half the $2\pi/N_b = 30^\circ$ angular ball spacing on the individual rows.

The experimental data used here was collected by Ho [58] who seeded defects into the drive-end bearing on the output shaft. An accelerometer was placed on the gearbox directly above the defective bearing to measure the vibration response in the vertical direction under various load and speed conditions. A sample frequency $F_s = 24$ kHz was used and 1.4 seconds of data was collected. Photos of the line and extended spalls inserted on the outer raceway are shown in Figure 4. The line spall was inserted using electric spark erosion and the extended rough surface defect was created using an electric etching pencil. Both defects extended across half of the outer raceway width such that only one row of balls (row $r = 1$) interacted with the defect. The line spall had a depth of $d = 0.3$ mm and an angular extent of $\Delta \phi_f = 2^\circ$. The extended spall had a angular extent of $\Delta \phi_f = 40^\circ$ and a roughness of $8 \mu$m although it was not specified if this was a standard deviation or another descriptor. A torque of 16 Nm was applied to the gears which
results in a static load of \( W = 346 \) N on the defective bearings acting along the gear pressure angle of 20°. The \( y \) direction in the model is assumed to be along this pressure angle. The outer raceway defects were centered in the gear load zone such that \( \phi_y = 270° \) in Figure 1. The running speed is \( f_s = 7 \) Hz which results in a nominal outer raceway ball pass frequency \( f_{bpo} = N_h f_s = 34.2 \) Hz, and a gear toothmesh frequency of 224 Hz.

5.2. Model parameters

The nonlinear dynamic model introduced in Section 3 was implemented to simulate the vibration response measured on the gear box test rig, and to predict and analyze the contact forces and bearing stiffness variations that occur for the defect-free and line and extended spall cases. The model was implemented in Simulink® and the equations of motion were solved using the ordinary differential equation solver (ode45) which is based on a Runge-Kutta method. Initial conditions were set to the quasi-static solution corresponding to the cage position at time \( t = 0 \). The radial clearance was assumed to be \( c = 0 \) μm. The parameter values used in the model are included in Table 1. The natural frequency and damping ratio of the high-frequency resonant mode included in the model were set to 10.3 kHz and 1%, respectively. This corresponds to one of the bearing resonances that is excited when a ball exits the spall. The cage position was generated in the model such that approximately 1% of variation in the ball pass frequency was obtained. The continuous time-domain results were discretized using a sample frequency of \( F_s = 65,536 \) Hz to enable calculation of velocity and envelope spectra. Power spectral densities were calculated from 1.4 seconds of simulated data using a hanning window and a 1 Hz frequency resolution.

The modeled defect depth profiles \( C_i(\phi) \) for the line and extended spalls are illustrated in Figures 5(a) and (b), respectively. The defects are included on row \( r = 1 \) (blue line) and row \( r = 2 \) is modeled as defect-free (red line). The depth of the line spall is adjusted to the maximum depth reached by the ball as it passes through the defect, and the rectangular geometry is adjusted to the approximate path followed by the ball center. Previous research [23] has shown that making these adjustments results in better agreement between the modeled and measured response.

5.3. Quasi-static load distributions

Figures 5(c)-(f) present the quasi-static load distributions for the defect-free bearing and the defective bearings with a line or extended spall, which were calculated as described in Section 4.1. The blue lines show the load distribution for the row with the defect, the red lines for the defect-free row, and the green lines for the defect-free bearing. The grey rectangles indicate the angular extent of the defects. For the defect-free bearing, the load zone extends over 180° as defined by Equation (23), and the load is shared equally between the rows resulting in the same load distribution on the two rows. The spacing between peaks of the same color in the line spall results included in Figures 5(c) and (e) equals the 30° ball spacing on a single row. The spacing between peaks of different color equals the 15° angular offset between the two rows enforced by the staggered cage arrangement.

The results in Figures 5(c) and (e) show that the load distribution for the line spall only deviates from the defect-free bearing case when a ball is positioned in the spall. In this instance, this ball loses its load carrying capacity, and the load it carried is redistributed over the adjacent loaded balls on both the defective and defect-free rows. For the extended spall, the load distribution in Figures 5(d) and (f) deviates from the defect-free bearing case over the entire load zone because there are always two balls in the defect. These balls lose some or all of their load carrying capacity depending on the defect depth at the ball position. The load is again redistributed over the balls on both rows. The results indicate that when a raceway defect is present, defect-free sections of the raceway are subjected to increased loading. The angular separation between these defect-free raceway sections is equal to the ball spacing for the case of a line spall on the outer raceway. For both the line and extended spall cases, most of the load that needs to be redistributed is taken on by the balls closest to the load zone center. The above observations will also apply to single row bearings.

<table>
<thead>
<tr>
<th>Hertzian contacts</th>
<th>Mass</th>
<th>Stiffness</th>
<th>Damping</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K = 5157 ) MN/m^1/3</td>
<td>( m_i = 0.5134 ) kg</td>
<td>( k_{ox} = 17.7 ) MN/m</td>
<td>( c_{ox} = 673.8 ) Ns/m</td>
</tr>
<tr>
<td>( c = 100 ) Ns/m</td>
<td>( m_o = 10.268 ) kg</td>
<td>( k_{oy} = 17.7 ) MN/m</td>
<td>( c_{oy} = 673.8 ) Ns/m</td>
</tr>
<tr>
<td>( m_p = 0.5 ) kg</td>
<td>( k_i = 2094.0 ) MN/m</td>
<td>( c_i = 317.3 ) Ns/m</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Parameter values used in the nonlinear dynamic model of the defective bearing in the gear test rig.
Figure 5: (a,b) Defect depth profiles for line and extended spalls on outer raceway with circumferential lengths $\Delta \phi_f = 2^\circ$ and $40^\circ$ and a width extending across one row of balls. (c-f) Load distributions in $x$ and $y$ directions with a line (left column) or extended (right column) spall on row $r = 1$ (blue) and no defect on row $r = 2$ (red) compared to defect-free bearing (green).
For the considered line spall case, a ball loses all of its load carrying capacity when it is positioned in the defect. More generally, a ball will lose either a part or all of its load carrying capacity depending on the defect geometry, the applied load, the clearance, and the bearing macro-geometry. In either case, the reduction in the load carried by the ball in the defect is compensated for by increased loading on the balls outside the defect.

5.4. Quasi-static bearing stiffness variations

The redistribution of the static load that occurs when a defect is present on a raceway described in Section 5.3 has a direct effect on the stiffness of the bearing assembly. This is shown in Figure 6 which presents the variations in the bearing stiffness matrix $K$ with cage position for both the line spall (left column), extended spall (right column), and defect-free bearing cases (both columns). The bearing stiffness matrix was calculated as a function of cage position as described in Section 4.2. For the defective bearings, the blue lines indicate the stiffness provided by the defective row $(r = 1)$, the red lines the stiffness provided by the defect-free row $(r = 2)$, and the magenta lines the total stiffness. For the defect-free bearing, the green and cyan lines indicate the stiffness provided by each row, with the total stiffness indicated by the black lines.

The defect-free bearing results in Figure 6 illustrate the varying compliance of the bearing assembly investigated in previous research [40–43]. Figure 6(a) shows that the bearing stiffness terms $k_{xx1}$ and $k_{xx2}$ vary periodically at the 30° spacing between balls on individual rows. The staggered arrangement of the two rows results in a total bearing stiffness term $k_{xx}$ that varies periodically at the 15° spacing between balls on both rows. The same observations apply to the stiffness terms $k_{yy}$ and $k_{xy}$ but the amount of variation is much smaller in the loaded direction, i.e. about 0.1-0.2 MN/m, and therefore not visible in Figure 6(b). The variations would increase if the clearance was increased but would remain small in comparison to typical variations caused by a defect. The bearing stiffness variations produce parametric excitations of the bearing assembly leading to the so-called varying compliance vibrations [40]. The results demonstrate that for a double row bearing with a staggered cage arrangement, the varying compliance vibrations occur at a frequency that is equal to twice the characteristic frequency $f_{bpo}$ for an outer raceway defect extending across one row only. This is different from a single row bearing where these frequencies coincide [40–43].

The results in Figure 6 show that for the line spall (left column), the bearing stiffness only deviates from the defect-free bearing case when a ball is positioned in the defect. For the extended spall, there are always two balls in the defect at once resulting in a bearing stiffness in Figure 6 (right column) that deviates from the defect-free bearing case for all cage positions. When a ball is in the defect, the stiffness in the loaded direction reduces for the defective row ($k_{yy1}$) and increases for the defect-free row ($k_{yy2}$). The stiffness reduction for the defective row is smaller than the increase for the defect-free row such that the net result is a reduction in the bearing stiffness $k_{yy}$ in the loaded direction. The stiffness $k_{xx}$ in the unloaded direction increases when a ball is in the defect, with the stiffness increase provided by both rows. This increase occurs because, as illustrated in Figure 5(c), the load on balls that are loaded in the $x$-direction increases when a ball is in the defect, which in turn increases the stiffness of the nonlinear contact springs. The stiffness $k_{yy}$ decreases because, although the load reductions and increases on the balls observed in Figure 5(e) balance to zero to achieve static equilibrium, this is not the case for the contact spring stiffness reductions and increases due to the nonlinear nature of the Hertzian contacts.

Comparing the results in Figures 6(e) and (f) indicates that the cross-coupling stiffness $k_{xy}$ is larger for the extended spall case than for the line spall case. For the defect-free and both defective bearing cases, the cross-coupling stiffness is one to two orders of magnitude lower than the radial stiffnesses $k_{xx}$ and $k_{yy}$. This means that the natural frequencies of the linearized dynamic system are reasonably approximated by the natural frequencies of the mass-spring-damper systems illustrated in Figure 2.

Similar to the varying compliance of the defect-free bearing case [40], the defective bearing stiffness terms for the individual rows vary periodically at the 30° spacing between balls on each row, and the total bearing stiffness varies periodically at the 15° spacing between balls on both rows. However, the rate of change and the magnitude of the variations are much larger in comparison to the defect-free bearing case, especially in the loaded direction. The bearing stiffness variations produce parametric excitations of the bearing assembly at the characteristic frequency $f_{bpo}$ for an outer raceway defect that extends across one row only. More precisely, the excitation is pseudo-periodic due to the slippage of the balls. The character of the resulting vibration response can be directly correlated to the character of the stiffness variations, and also to the average stiffness over the aforementioned 15° period. For example, the rate of change in the bearing stiffness determines the degree of impulsiveness.
Figure 6: Bearing stiffness variations in the presence of a line spall (left column) and extended spall (right column) on row \( r = 1 \) compared to the variations for a defect-free bearing. For the defective bearing case, the blue, red and magenta lines indicate the stiffness provided by the defective and defect-free rows and the net stiffness. The green, cyan and black lines show the stiffness for both rows and their sum for the defect-free bearing case.
5.5. Dynamic contact force analysis

Figures 7(a-d) present the dynamic contact forces $F_{jr}(t)$ between the raceways and the balls on both rows for the line and extended spall cases. The quasi-static loads $F_y$ on the balls which were illustrated in Figure 5 are also included, with the thick and thin lines of the same color indicating the dynamic and quasi-static loads for the same ball. Note that in the presented results, ball $j = 6$ enters the defect after ball $j = 7$. The net dynamic contact forces $F_x(t)$ and $F_y(t)$ acting on the inner and outer raceways are illustrated in Figures 7(e-f) for the line and extended spalls, respectively. The static load $W$ is subtracted from the net contact force in the loaded direction $F_y(t)$ to illustrate the dynamic component only.

The line spall results presented in Figures 7(a) and (c) show that as ball $j = 6$ enters the defect, it destresses and loses its load carrying capacity. The static load it carried needs to be redistributed between the loaded balls on both rows that are positioned outside the defect. The required redistribution of the static load causes dynamic contact forces which fluctuate around the quasi-static load solution for each individual ball. Upon exiting the defect, ball $j = 6$ restresses between the raceways and the other balls that took on additional load destress back to the normal load carried. This event again causes dynamic contact forces which decay to the quasi-static load solution. The fluctuations around the quasi-static load solutions result in the net dynamic contact forces presented in Figure 7(e).

For the extended spall results presented in Figures 7(b) and (d), ball $j = 6$ is in the defect with ball $j = 7$ when it enters the defect and ball $j = 5$ when it exits. When the two balls in the defect (partly or entirely) destress and/or restress at the defect entrance and exit and also within the defect zone, the required redistribution of the static load again causes dynamic contact forces which fluctuate around the quasi-static load solution for each individual ball. The fluctuations, which are smaller in comparison to the line spall results and are not visible in Figures 7(b) and (d), produce the net dynamic contact forces presented in Figure 7(f). The differences in character between the net contact forces for the line and extended spall cases are directly related to the differences in the quasi-static bearing stiffness variations illustrated in Figure 6. Because the quasi-static bearing stiffness varies more rapidly for the line spall in comparison to the extended spall, the dynamic fluctuations around the quasi-static load solutions, and therefore the resulting net contact forces, are more impulsive in character and of a higher magnitude for the line spall.

5.6. Modeled and measured vibration spectra

Figures 8 and 9 compare the modeled and measured velocity and acceleration envelope spectra in the $y$ direction for the good bearing and defective bearings with a line or extended spall. The markers indicate defect frequency components (red circles), cage frequency sidebands of defect components (red diamonds), run speed components (magenta circles), gear toothmesh components (green circles), and run speed sidebands of gear toothmesh components (green squares). The defect and gear toothmesh components occur at frequencies of 34.2 Hz and 224 Hz and harmonics. The envelope spectra were calculated after band-pass filtering the acceleration response between 10.2 and 10.4 kHz. This demodulation band was chosen as it includes one of the bearing resonances that is excited when a balls exits the spall. Figure 10 presents the measured and modeled acceleration spectra within the demodulation band. The defect frequency components cannot be observed in the acceleration spectra at these high frequencies due to the slippage of the rolling elements which is accounted for in the modeling.

For the good bearing, Figure 8(a) shows that the measured velocity spectrum is dominated by gear toothmesh components with run speed sidebands, and run speed components. The gear components do not occur in the modeled results since the forces generated by the gears are not included in the model. The run speed components, which are of a much lower level than the gear components, are typically caused by small misalignments or unbalances which are also not considered in the model. Figure 8(b) shows that the modeled velocity spectrum for the good bearing contains very low amplitude components at the defect frequency which are caused by the varying bearing stiffness [43]. The predicted low amplitude of these components explains their absence in the measured velocity spectrum for the good bearing, where they are completely masked by the gear and run speed components.

The measured velocity spectra for the defective bearings shown in Figures 8(c) and (e) also contain the gear and run speed components. These components mask the defect components in the measured velocity spectrum for the line spall case as shown in Figure 8(c) but not for the extended spall case as shown in Figure 8(e). For the latter case, the measured velocity spectrum contains a strong defect component at 205.2 Hz which is of a similar level to the gear components. Figures 8(d) and (f) show that the difference between the measured velocity spectra for the line and extended spall cases is predicted by the model, with the modeled defect components having a much higher amplitude.
Figure 7: (a-d) Dynamic contact forces $F_{jr}(t)$ between the raceways and the balls on the defective ($r = 1$) and defect-free rows ($r = 2$) for the line and extended spall cases. The thick and thin lines of the same color indicate the dynamic and quasi-static contact forces $\bar{F}_{jr}$ for the same ball. (e-f) Net dynamic contact forces for the line and extended spalls. The static load $W$ is subtracted from the net contact force $F_j(t)$ in the loaded direction to illustrate the dynamic component only.
for the extended spall case compared to the line spall case. The measured and modeled difference in the velocity spectra is explained by comparing the variations in the modeled quasi-static stiffness \( k_e \) for the direction of greatest excitation which were illustrated in Figure 6. Compared to the line spall, the quasi-static stiffness variations for the extended spall have more low-frequency energy. In other words, the parametric excitations caused by the extended spall have more low-frequency energy which results in the stronger defect components in the velocity spectrum.

The measured velocity spectrum for the extended spall shown in Figure 8(e) contains cage frequency sidebands around the defect components which the current model cannot predict as indicated by the modeled velocity spectrum shown in Figure 8(f). These sidebands could be caused by non-uniformity between the ball diameters. However, the authors and other researchers [59] have observed cage frequency sidebands in other bearings with outer raceway defects and the exact cause of these sidebands remains unknown at this stage.

Figures 9(c–f) show that for both the measured and modeled results, the line spall produces higher amplitude defect components in the envelope spectra than the extended spall. This occurs because the rate of change in the varying bearing stiffness is greater in Figure 6 for the line spall than for the extended spall. The more rapid stiffness changes cause increased parametric excitation of the high-frequency bearing resonant mode and therefore a more impulsive vibration response. It is well-known that this results in higher amplitude defect components in the envelope spectrum. For the extended spall, the modeled envelope spectrum in Figure 9(f) includes low amplitude defect components which are not observed in the measured envelope spectrum in Figure 9(e). This occurs because the background vibrations that mask the defect components in the measured envelope spectrum were not considered in the model. This also explains the difference between the modeled and measured envelope spectra for the good bearing in Figures 9(a) and (b). Modeling of the background vibration environment is typically achieved by simply adding pink or random noise to the simulated results [23, 24, 26]. However, this is not necessary here since the model accuracy is determined by the accuracy with which it predicts the amplitudes of the defect components in the velocity and envelope spectra.

6. Fan test rig simulations and experiments

6.1. Fan test rig and model description

The fan test rig data considered in this section was collected as part of a previous study [39] into the development of spall size estimation techniques for defective bearings. The test rig includes a fan disk with 19 blades fitted on a shaft which is supported by two Nachi 21465GK double row self-aligning ball bearings. These bearings have \( N_b = 14 \) balls on each row, a pitch diameter \( D_p = 45 \text{ mm} \), a ball diameter \( D_b = 8 \text{ mm} \), and a contact angle \( \alpha = 0^\circ \). The balls are retained in a staggered arrangement by a single cage, with the angular offset between the two rows equal to half the angular ball spacing on the individual rows. The bearings are mounted on sleeves and contained within plunger blocks. The shaft is driven by a motor via a V-belt and the rotational speed is controlled using a variable voltage and frequency drive. Line spalls of varying angular extent were inserted on the inner or outer raceway, and the vibration response was measured at various speeds using an accelerometer mounted on the defective bearing. A detailed description of the test rig and the defects that were inserted can be found in reference [39].

Analysis of the measurements in the previous research [39] showed that when a ball enters the defect, a low frequency event occurs. Upon exiting the defect, a high frequency event occurs with a higher amplitude and more impulsiveness than the low frequency event. This high frequency impulsive event is the event normally used in envelope analysis to diagnose a defective bearing. The spill size was estimated by detecting the time separation between the low and high frequency events, and by assuming that the spatial separation between the events corresponds to half the defect size. It is generally accepted that the high frequency impulsive event corresponds to the excitation of a high-frequency bearing resonant mode which occurs when a ball exits the defect. However, the reported low frequency event [39] has not been related to a dynamic property of the bearing assembly. This is done here using the methods developed in Section 4.

The multi-body nonlinear dynamic model introduced in Section 3 was implemented to model and analyze the vibration response measured on the fan test rig. The implementation was similar to the implementation for the gearbox test rig described in Section 5.2. Table 2 presents the parameter values used in the model. The measured response considered here was sampled at \( F_s = 65,536 \text{ Hz} \) and is for a running speed of 800 rpm and a spall size of 1.2 mm which is equivalent to \( \Delta \phi_f = 2.6^\circ \). The line spall was included in the defect depth profile for row \( r = 1 \) and row \( r = 2 \) was modeled as defect-free. The depth was again adjusted to the maximum depth reached by the ball as it passes.
through the defect. The defect frequency at the considered run speed equals $f_{bpo} = 76.7$ Hz. The natural frequency and damping ratio of the high-frequency resonant mode included in the model were set to $9.7$ kHz and $3\%$, respectively. This corresponds to a high-frequency bearing resonant mode observed in the measured vibration response. A static load $W = 100$ N and a clearance of $c = 3\mu m$ were used in the model. The measured vibration response was low-pass filtered below $11$ kHz using a butterworth filter of order 6 because the response measured at higher frequencies would have been affected by using beeswax to mount the accelerometer [52]. Modeled and measured envelope spectra were calculated using a demodulation band of $9.2-10.2$ kHz centred at the bearing resonance mode at $9.7$ kHz.

6.2. Vibration response analysis

Figures 10(a-d) present the modeled and measured acceleration responses and their spectrograms. Figures 10(e-f) compare the modeled and measured acceleration spectra in the demodulation band and the resulting envelope spectra. The magenta lines included in both spectrograms indicate the natural frequencies of the linearized dynamic model of the bearing assembly, which were calculated as described in Section 4.3. Because the coupling between the $x$ and $y$ directions is weak, these natural frequencies are closely approximated by the natural frequencies of the mass-spring-damper systems illustrated in Figure 2. In addition, the modes in the $x$ direction are hardly excited because excitation occurs predominantly in the $y$-direction and cross-coupling is weak. Therefore, only the natural frequencies of the three excited modes in the $y$ direction are included in Figures 10(c) and (d). The mode with a natural frequency of $9.7$ kHz corresponds to the high-frequency bearing resonance mode included in the model. Modes with natural frequencies of $0.90$ kHz and $1.99$ kHz correspond to the rigid body modes of the bearing assembly.

The results in Figure 11 show that the model reasonably predicts the time-frequency characteristics of the measured vibration response. The spectrograms clearly visualize the low and high frequency events that occur when a ball passes through the defect, and good agreement is achieved between the measured and modeled envelope spectra. Moreover, the results indicate that the low frequency event at the defect entrance corresponds to a parametric excitation of the rigid body modes of the linearized and decoupled dynamic model of the bearing assembly. At the defect entrance, the stiffness varies gradually such that only the low frequency rigid body modes of the bearing assembly are excited, resulting in the low frequency event with characteristic frequencies of $0.90$ kHz and $1.99$ kHz. At the defect exit, the stiffness varies more rapidly which excites both the low and high frequency modes of the bearing assembly. The measured spectrogram in Figure 11(d) indicates the presence of additional high frequency bearing resonant modes around $6$ kHz and $11.5$ kHz which were not included in the model.

7. Conclusion

A method was presented for calculating and analyzing the quasi-static load distribution and stiffness variations for a radially loaded double row rolling element bearing with a raceway defect. The method was applied to defective bearings on gearbox and fan test rigs. Analysis of the load distributions showed that when balls pass through the defect and lose all or part of their load carrying capacity, the load is redistributed between the other loaded balls. This includes the balls that are positioned outside the defect such that good raceway sections are subjected to increased static loading when a raceway defect is present. Analysis of the bearing stiffness variations showed that when balls are positioned in the defect, the stiffness decreases in the loaded direction and increases in the unloaded direction. For an extended spall, which always has one or more balls positioned in the defect, this results in a reduced average stiffness in the loaded direction compared to both the line spall and good bearing cases. The bearing stiffness variations due to the defect result in parametric excitations of the bearing assembly, which play an important role in the resulting vibration response. The qualitative character of the resulting vibration response correlates strongly to the character of the stiffness variations. Rapid stiffness changes, which typically occur at a sharp defect exit, produce high frequency

<table>
<thead>
<tr>
<th>Hertzian contacts</th>
<th>Mass</th>
<th>Stiffness</th>
<th>Damping</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 10.1$ GN/m$^3$</td>
<td>$m_l = 2$ kg</td>
<td>$k_{ox} = 262.8$ MN/m</td>
<td>$c_{ox} = 2050.4$ Ns/m</td>
</tr>
<tr>
<td>$c = 150$ Ns/m</td>
<td>$m_o = 2.5$ kg</td>
<td>$k_{oy} = 262.8$ MN/m</td>
<td>$c_{oy} = 2050.4$ Ns/m</td>
</tr>
<tr>
<td></td>
<td>$m_r = 0.1$ kg</td>
<td>$k_r = 371.5$ MN/m</td>
<td>$c_r = 365.7$ Ns/m</td>
</tr>
</tbody>
</table>
impulsive events in the vibration response. Slower stiffness variations due to large wavelength waviness features in an extended spall produce low frequency parametric excitations which results in defect components in the velocity spectra. This was demonstrated based on the modeled and measured vibration response for a gearbox test rig with defective bearings. The modeled response was generated using a previously developed multi-body nonlinear dynamic model of a defective bearing [23, 24], which was enhanced and extended to a double row bearing. The predicted contact forces were shown to fluctuate around the quasi-static loads on the balls, with rapid stiffness changes producing high magnitude impulsive force fluctuations. Furthermore, the low frequency event that occurs when a ball enters a line spall, which was reported in previous research on the fan test rig [39], was linked to the dynamic properties of the bearing assembly.

The presented methods and results contribute to the understanding of varying stiffness excitations in defective bearings and their effect on the measured vibration signature. They can be applied to other multi-body nonlinear dynamic models of defective bearings [21–29] to improve understanding of the time-frequency characteristics of the predicted vibration response. Besides the varying stiffness excitations considered here, the other important excitation mechanism is the impact force caused by a rolling element mass striking the bottom or edge of a defect. Since previous work found that rolling element inertia could be neglected at the lower speeds considered here [21–24, 45], this excitation mechanism was not considered in the modelling. Nevertheless, to advance our understanding of the vibration response of defective bearings, future work should investigate the relative importance of the two excitation mechanisms as a function of speed, load, clearance, defect geometry, etc. The analysis method developed here will be of great benefit to such future research.

8. Acknowledgement

The work presented in the paper was funded by Australian Research Council (ARC) Linkage Project LP110100529.

References


Figure 8: Comparison of the measured and modeled velocity spectra for the spur gear test rig for the case of a good bearing and a defective bearing with a line or extended spall inserted on one of its outer raceways. The markers indicate defect frequency components (red circles), cage frequency sidebands of defect components (red squares), run speed components (magenta circles), gear toothmesh components (green circles), and run speed sidebands of gear toothmesh components (green squares). The defect frequency components occur at $f_{bpo} = 34.2$ Hz and its harmonics.
Figure 9: Comparison of the measured and modeled envelope spectra for the spur gear test rig for the case of a good bearing and a defective bearing with a line or extended spall inserted on one of its outer raceways. The markers indicate defect frequency components (red circles) which occur at $f_{bpo} = 34.2$ Hz and its harmonics.
Figure 10: Comparison of the measured and modeled acceleration spectra within the 10.2–10.4 kHz demodulation band used for envelope analysis of the spur gear test rig data.
Figure 11: (a) Modeled vibration response. (b) Measured vibration response. (c) Spectrogram of modeled vibration response with the magenta lines indicating the natural frequencies of the linearized dynamic model. (d) Spectrogram of measured vibration response with the magenta lines indicating the natural frequencies of the linearized dynamic model. (e) Modeled and measured acceleration spectra within 9.2-10.2 kHz demodulation band used in envelope analysis. (f) Modeled and measured envelope spectra.