Dynamics of Tapered and Spherical Roller Bearings: Modelling and Analysis

M. Lima
Instituto Superior Técnico
Universidade de Lisboa
Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
marisa.lima@tecnico.ulisboa.pt

ABSTRACT
The understanding of the dynamic performance of roller bearings, used in railway vehicles, is the fundamental objective of project MAXBE, which motivates the work presented here. The evaluation of the bearings performance via monitoring systems uses the vibration information to infer the health of their mechanical components. The vibration output of the axleboxes is the measurable outcome of the bearing dynamic response, under operating conditions, that is characterized in this work. The main goal of this work is to develop a dynamic analysis tool, referred to as BearDyn, in MATLAB®, able to handle models representative of actual railway axle bearings, by using a multibody formulation to describe the mechanical elements of the bearing and their interactions. Realistic bearing geometric data is obtained by precise measurements of spherical and tapered bearings, with the support of project partners. The interactions between the elements are described by continuous contact force models based on the Hertz elastic contact theory and modified according to experimental evidence. Tribological lubrication models are applied to describe the tangential forces in the presence of lubricant. Finally, BearDyn is tested being the inner raceway rotating with an angular velocity respective to realistic train operations and a load applied to its center, resulting from the wheel-rail contact force. The outer raceway is fixed for the analysis developed here. The bearing dynamic response is obtained in terms of forces, kinematic quantities and different interaction measures, in the time domain, being the dynamic response in the frequency domain, obtained by using Fast Fourier Transforms, which is ultimately used for direct correlation with the outcome of the monitoring stations.

Keywords: Roller Bearing, Multibody Dynamics, Railway Dynamics, Hertz Contact, Elastohydrodynamic Lubrication.

1 INTRODUCTION
In the railway industry, there is an important focus on the selection, understanding and maintenance of bearings, resulting from their direct impact on train safety and inactivity periods for maintenance. As the axle bearing damage process has a direct impact in safety and economics of railway systems, this has been the object of study of a collaborative project between several European organizations, known as MAXBE (Interoperable Monitoring, Diagnosis and Maintenance Strategies for Axle Bearings). This project appears with the objective to provide concepts and strategies for the interoperable monitoring and diagnosis of axle bearings. The work presented here is developed as part of this project, of which the dynamic performance of the roller bearings used in railway vehicles is the fundamental objective.

Axleboxes are the linking element between the rotating wheelset and the frame of the bogie via the primary suspension of a railway vehicle. All forces acting on the axleboxes are transmitted via springs, dampers and guiding elements, i.e., the primary suspension, to the bogie frame. An image of an axlebox mounted on a train is shown in Figure 1(a). The axlebox is an assembly normally comprising the axlebox housing, rolling bearings, sealing solution and a proper
The bearings predominantly used in these applications are spherical and tapered bearings [1]. Due to the integration of the work now presented in the MAXBE project, bearings with spherical and tapered rollers, presented in Figure 1(b), are the only types considered here.

Faults which typically occur in rolling element bearings caused by localized defects generate a series of impact vibrations every time a running roller passes over the surfaces of the defects. The vibrations of healthy roller bearings occur at Bearing Characteristic Frequencies (BCF), which are estimated based on the running speed of the shaft, the geometry of the bearing and number of rollers. In the case of defects, other than the healthy frequencies appear in the dynamic response of the roller bearings. The Fast Fourier Transforms is used to process the dynamic response of the roller bearing being its characteristic frequencies monitored to analyse the change of vibration amplitude and detect the existence of faults. There are, however, difficulties in this process, since the impact vibration is usually overwhelmed by noise and vibrations generated from other components such as the rail-wheel contact [2]. Due to the difficulties that arise from using data from experimental methods, to predict the life of a bearing, computational methods are seen as an advantageous alternative. The performance limitations of a mechanical system can often be determined by the dynamic response of its subsystems. In the cases of systems that include roller bearings, it is necessary to predict bearing dynamic performance as an integral part of machine design analysis, which can be done by studying the dynamic response of a bearing with prescribed load and speed [3].

With the progress of computer technology, Computer Aided Engineering (CAE) becomes popular and essential to design better bearings to match user requirements with optimized performance. In this context, some dynamic bearing analysis programs were created, such as ADORE, which is a commercially available software package that allows its users to analyse the performance of roller and ball bearings by using multibody formulation to describe the elements of the bearing and their interactions [4] [5]. IBDAS [6], BRAIN [4] and BEAST [7] are some other examples of similar tools that also use this approach but that are proprietary of roller bearing manufacturers. The approach used here includes the development of a dynamic analysis tool, referred to as BearDyn which serves as the acronym for Bearing Dynamic Analysis Program, able to handle bearing models representative of the actual bearings used in railway vehicles, developed in MATLAB®. This code uses a multibody formulation to describe the elements of the bearing and their interactions. The interactions between the rollers, cage, and raceways are described by continuous contact force models based on the Hertz elastic contact theory [8] and modified according to experimental evidence. Tribological lubrication models are applied to describe the tangential forces in the presence of lubricant [9] [10]. This work constitutes the description of the dynamic analysis tool, BearDyn, its fundamental formulations and its application to realistic models of spherical and tapered roller bearings. The objectives of this work require the understanding of the performance of the axle bearings in actual operation conditions and include the development of a dynamic analysis tool able to handle bearing models representative of the
actual bearings used in railway operations. The final goal of the work described in this document is to obtain a dynamic analysis tool BearDyn which is able to perform a study on tapered roller bearings, whose model is described by data supplied in dedicated files. The dynamic response of the forces is to be post-processed to obtain the Frequency Response Functions (FRF), which serve as the basis for the evaluation of the bearing health. The development of the necessary methods for the dynamic analysis of the bearings is presented, being the methods for contact between elements and tribology contact models also included. The demonstration of the developed program is done by performing dynamic analysis of realistic roller bearing models.

2 MULTIBODY DYNAMIC EQUATIONS

Due to its simplicity and computational easiness, Cartesian coordinates and Newton-Euler’s method are used to formulate the equations of motion of the spatial multibody systems [11]. When Cartesian coordinates are used, the position and orientation of a rigid body \( i \) must be defined by a set of translational and rotational coordinates. The position of the body with respect to global coordinate system XYZ is defined by the coordinate vector \( \mathbf{r}_i = [x \ y \ z]^T \) that represents the location of the local reference frame \((\xi \eta \zeta)\). The orientation of the body is described by the rotational coordinate’s vector \( \mathbf{p}_i = [e_x \ e_y \ e_z]^T \), which is built with the Euler parameters for the rigid body [11]. Therefore, the vector of coordinates that describes the rigid body \( i \) is,

\[
\mathbf{q}_i = [\mathbf{r}_i^T \ \mathbf{p}_i^T]^T
\]

A spatial multibody system with \( nb \) bodies is described by a set of coordinates \( \mathbf{q}^* \) in the form,

\[
\mathbf{q}^* = [\mathbf{q}_1^T \ \mathbf{q}_2^T \ \ldots \ \mathbf{q}_n^T]^T
\]

The velocities and accelerations of body \( i \) are given by vectors,

\[
\dot{\mathbf{q}}_i = [\mathbf{r}_i^T \ \dot{\mathbf{p}}_i^T]^T
\]

\[
\ddot{\mathbf{q}}_i = [\mathbf{r}_i^T \ \ddot{\mathbf{p}}_i^T]^T
\]

and the relation between the angular velocities, \( \omega_i^* \), and the time derivatives of the Euler parameters, \( \dot{\mathbf{p}}_i \), is,

\[
\dot{\mathbf{p}}_i = \frac{1}{2} \mathbf{L}_i^T \ddot{\mathbf{p}}_i
\]

being matrix \( \mathbf{L} \), defined in Nikravesh [11].

The knowledge of the system response includes the evaluation of the position and orientation, velocity and acceleration of each of the bodies of the system. For this purpose, the equations of motion of the system must be established and their solution evaluated and integrated in time. In the computational tool BearDyn developed for this project, kinematic constraints were not introduced as its range of application includes isolated roller bearings only. Consequently, the system is defined as a system of unconstrained bodies. In terms of Cartesian coordinates, the equations of motion of an unconstrained multibody mechanical system are written as,

\[
\mathbf{M} \ddot{\mathbf{q}} = \mathbf{g}
\]

where \( \mathbf{M} \) is the global mass matrix, containing the mass and the inertia tensor of all bodies and \( \mathbf{g} \) is a force vector that contains the external and internal forces, acting on the bodies of the system, resulting from gravity, gyroscopic forces, contact, friction and elastohydrodynamic forces. The system of Eqs. (5) is solved for \( \ddot{\mathbf{q}} \), after the vector of forces \( \mathbf{g} \) is evaluated. In each integration time step, the accelerations vector, \( \ddot{\mathbf{q}} \), together with velocities, \( \dot{\mathbf{q}} \), are integrated in order to obtain the new velocities \( \dot{\mathbf{q}} \) and positions \( \mathbf{q}^* \) for the next time step. This process is repeated until the complete description of the system motion is obtained for a selected time interval. The commonly used numerical integration algorithms are useful in solving first-order differential equations that take the form [12],

\[
\dot{\mathbf{y}} = f(\mathbf{y}, t)
\]
The \( nc \) second-order differential equations, are converted to \( 2nc \) first-order equations by defining the \( y \) and \( \dot{y} \) vectors, which contains, respectively, the system positions and velocities and the system velocities and accelerations, as follows,

\[
y = \begin{bmatrix} q^+ \\ q^- \end{bmatrix} \quad \text{and} \quad \dot{y} = \begin{bmatrix} q^+ \\ q^- \end{bmatrix}
\]

being \( q^+ \) obtained from \( \dot{q} \) using the relation described by Eq. (4). The reason for introducing the vectors \( y \) and \( \dot{y} \) is that most numerical integration algorithms deal with first-order differential equations [13]. The integration method used in BearDyn is a MATLAB® ordinary differential equation solver of medium order accuracy, ODE45, or any other suitable solver.

### 3 ROLLER BEARING INITIAL POSITION AND VELOCITY

The dynamic analysis of the roller bearings requires that the positions and velocities of each component of the bearing is known and properly initialized. Both types of roller bearings considered have 2 rows of rollers, an outer raceway, an inner raceway, two cages and a given number of rollers, as illustrated in Figure 2. Due to the precision with which the bearings are mounted, with tight tolerances, and the complexity of the geometries it is necessary to build a pre-processor that, based on a limited set of data, is able to evaluate the position, orientation and the velocities of all rolling elements of the bearing, in such a way that no interference between the mechanical elements exist as the dynamic analysis starts.

![Figure 2: Typical tapered rolling bearing elements. (a) Assembled roller bearing; (b) Individual elements of the rolling bearing, from left to right: outer raceway, cages, rollers and inner raceway.](image)

The data necessary to describe the geometry of either the spherical or the tapered bearing is provided to BearDyn by the user as an input text file with a specified format. The acquisition of the necessary information to fulfil the input data tables relative to spherical and roller bearings, was achieved by the tribology group at the Faculty of Engineering of University of Porto (FEUP), also partner in the MAXBE project. At this point it is assumed that the bearings are in full functional conditions and that they have no defects. The initial positions, orientations and velocities are initialized for each element of the spherical and tapered bearings, based on their geometric properties. The initializations are developed for different possible configurations of the bearings, such as a single or double row spherical bearing or tapered bearing.

The initial conditions derived in this section ensure that the analysis starts with consistent velocities between all rolling elements and that there is no interference between the bodies at the start of the analysis. It is only during the evolution of the system that contact may develop.

### 4 CONTACT DETECTION

For the formulation of the contact between the elements of the tapered bearing system, the positions and geometries of each body are evaluated to identify the points of proximity and the eventual existence of contact. The contacts considered in this work are mainly listed as: contact between the roller and the inner or outer raceways, contact between the roller and the right or left flanges and contact between the roller and cage pocket tops or sides.
4.1 Contact between two generic surfaces

To evaluate the existence of contact, assume two bodies as generic surfaces that approach each other in space. The search for contact between the two surfaces requires the identification of two parameters of each surface, associated to the location of the points that are either in contact or in closer proximity, here identified as points $P$ and $Q$.

![Diagram of contact between two surfaces](image)

*Figure 3: Candidates to contact points between two parametric surfaces. Point $P$ belongs to a surface on body $i$ and point $Q$ to a surface on body $j$."

With reference to Figure 3, the position of point $P$ on body $i$ is defined by the sum of vectors $\mathbf{r}_i$, which is the position of the roller center of mass with respect to the inertial frame (XYZ), and vector $\mathbf{s}_P$ representing the position of point $P$ with respect to the body $i$ fixed frame. The position of point $Q$ is analogous to point $P$. The coordinates of points $P$ and $Q$, expressed in body $i$ and $j$ reference frames and defined as $\mathbf{s}_P$ and $\mathbf{s}_Q$, respectively, depend on the parametric description of each surface. With reference to Figure 3, on point $P$ the vector normal to the surface is $\mathbf{n}_P$, while $\mathbf{t}_P$ and $\mathbf{b}_P$ are the tangent and binormal vectors to the surface, respectively, forming an orthogonal basis. The same applies to $\mathbf{n}_Q$, $\mathbf{t}_Q$ and $\mathbf{b}_Q$, respective to point $Q$. The conditions for minimal distance between the two surfaces are that points $P$ and $Q$ must be such that,

\[
\begin{align*}
\mathbf{d}^T \mathbf{t}_Q &= 0 ; \\
\mathbf{d}^T \mathbf{b}_Q &= 0
\end{align*}
\]  

(8)

Effective contact occurs if, besides the fulfilment of Eq.(8), interference between the surfaces also exists, which is expressed by

\[
\delta = \mathbf{d}^T \mathbf{n}_Q \leq 0
\]  

(9)

Otherwise, the points are of close proximity, but not in contact.

4.2 Contact between roller and other surfaces

When contact between two surfaces is detected a penalty force model is used to relate the interference with the normal contact law. The normal force contact models used here are based in the Hertz elastic contact [8] and must be either point or line contacts. Due to the complex spatial geometry involving the roller contact, instead of considering the roller geometry as it is, the roller is divided into a given number of slices, $N_s$, as shown in Figure 4, and the contact of each slice, represented by its middle circle, with a surface is evaluated instead. This procedure, known as the slice method, is proposed by Gupta [9] and by Harada [14] and used here.
Each contact problem is described by the contact of a circle, the central cross-section of the slice. The conditions for minimal distance between the circle and a surface are described by,

\[
\begin{align*}
\mathbf{d}' \mathbf{t}_p &= 0 \\
\mathbf{d}' \mathbf{t}_0 &= 0 \\
\mathbf{d}' \mathbf{b}_0 &= 0
\end{align*}
\]

where point \( P \) belongs to the circle and \( Q \) to the surface. For each roller, \( N_{dl} \) contact problems must be solved to fully represent the contact between each roller and a single surface.

### 4.3 Contact between roller and cage

In the case of contact between the roller and the top of the cage pocket, the end circle of the roller is the responsible for such contact, while the central cross-section of each slice is responsible for contact with the side of the cage. In both cases the procedure is to detect the contact between a circle and a line. The equations to be fulfilled in order to find the closest points are given as,

\[
\begin{align*}
\mathbf{d}' \mathbf{b}_0 &= 0 \\
\mathbf{d}' \mathbf{t}_p &= 0
\end{align*}
\]

It should be noted that in case of circle to line contact two different potential contact points must be monitored at all time.

### 4.4 Solution of the nonlinear system of equations

For a single roller bearing, at any particular evaluation of the equations of motion represented by Eq. (5), more than 1200 contact pairs need to be evaluated by solving systems of nonlinear equations, such as those expressed by Eqs. (10) or (11). Therefore the computational efficiency and the robustness of the numerical methods selected for the solution of the nonlinear system of equations is of critical importance.

For each of the contact pairs, the equations of minimum distance, represented by Eqs. (8), (10) or (11) are solved for the parameters that describe the surfaces using the Newton-Raphson method [15]. In case of difficulty of convergence, an optimization methodology, coded in the MATLAB® solver function `fsolve`, is used. Note that the use of Newton-Raphson method is much faster, computationally, than the optimization method and, therefore, should be preferred. Note also that the search for the different 1200 plus contacts does not need to be sequential and, consequently, there is room for the implementation of a parallel computational approach.
5 CONTACT FORCE MODEL

5.1 Normal contact forces

Once contact is detected, the contact forces that develop between the surfaces are evaluated. The contact forces are applied over each body in each pair of contacting points \( P \) and \( Q \) with a penalty force \( f_n \), which is dependent of the indentation developed between the two bodies, \( \delta \), calculated by Eq. (9). The normal contact forces developed during contact are distributed over a small area when compared with the dimensions of the contacting surfaces. Under these conditions, the stress distribution over the contact area is described by Hertz elastic contact theory [8]. Typically point and line contacts are considered in the contact of rolling elements.

The normal contact force in the case of point contact is,

\[
f_n = K_{pt} \delta^{3/2}
\]  

where proportionality factor \( K_{pt} \), or contact stiffness, is given by [16],

\[
K_{pt} = \pi \kappa E'(\frac{\rho}{4.5\kappa^3})^{1/2}
\]

where \( \kappa = a/b \), with \( a \) and \( b \) being the semi-axis of the contact ellipse, and \( E, \rho \) complete elliptical integrals of first and second kind [17]. Using a least-square relation, Brew and Hamrock [18] present an approximate expression for the elliptical integrals and for \( \kappa \) as dependent on the characteristic radii of the contacting surfaces, \( r_{ax}, r_{ay}, r_{bx}, r_{by} \). The equivalent modulus \( E' \) is dependent on the Young’s modulus and Poisson’s ratio of each body in contact.

Since on rollers of the bearings the geometry is not perfectly cylindrical, models for ideal line contact do not represent correctly the relation between indentation and normal force. In order to apply the normal contact force model for line contact, the roller is discretized in \( N_{sl} \) strips, by considering the slice method. The normal contact force of a strip of the roller is given by Palmgren’s simplified equation [19]

\[
f_{ns} = 0.356 E' N_{sl} \delta^{9/8} L_{eff}^{1/8} \delta^{3/2} s = 1, \ldots, N_{sl}
\]

where the counter \( s \) refers to the slice number in the roller and \( L_{eff} \) is the effective contact length.

5.2 Tangential forces

Besides the normal forces that develop during contact, also tangential forces due to friction or to the lubrication fluid develop between the contacting bodies. Depending on the lubricant film thickness and on the roughness of the contacting surfaces, different lubrication regime may occur and the equivalent friction coefficient has to be evaluated differently. The different contact modes range between dry contact to full fluid film lubricated mode [14]. Regardless of the type of contact, the relation between the tangential forces and the normal contact forces is given by

\[
f_t = \mu f_n
\]

where \( \mu \) is the equivalent friction coefficient. The tangential force is applied in the opposite direction of the relative velocity between the contacting surfaces. The relation between the fluid film thickness and the roughness of the contacting surfaces defines the type of contact mode that is taking place. The use of the different methodologies to describe the lubrication mode actually taking place for the calculation of the equivalent friction coefficient is not implemented in this work, being a test value of \( \mu = 0.1 \) used for simulations with BearDyn.

6 APPLICATION TO RAILWAY BEARINGS

6.1 Railway operation scenario

The load resulting from forces external to the roller bearing considered here is caused by the weight of the train car, while travelling at the operational velocity of 50 km h\(^{-1}\). This loading,
which represents the average vertical force on each axle bearing, must in future studies, be replaced by a more realistic force resulting from the wheel-rail interaction and transferred from the wheel to the axle. Knowing that each railway vehicle is supported by two bogie, which distribute the vehicle weight over its components, the resulting load is considered as being applied equally and directly over the axles. The inner raceways of the roller bearings are rigidly fixed to the shaft of the wheelsets, being assumed that the proper fraction of load resulting from the weight of the train is applied directly to the center of mass of the roller bearing inner raceway. For a common passenger train car, weighing approximately 40 tons, with 8 axle bearings, a force of 50 kN is applied downwards in the vertical direction, Z, in the inner raceway of the bearing. Since the inner raceway of the roller bearing is fixed to the wheel shaft, while the outer raceway to the axlebox, a constant angular velocity is applied to the inner raceway and the outer raceway is fixed to the inertia frame. Considering the operational velocity of 50 km h$^{-1}$ of the train, as the wheel shaft is part of the wheelset and considering this as a rigid body where the wheel has a mean radius of 0.45 meters, the inner raceway angular velocity is initialized with a value of 30 rad s$^{-1}$.

6.2 Dynamic response of a tapered roller bearing

BearDyn is prepared to perform a dynamic analysis of either tapered or spherical roller bearings. In this work, only the results of the dynamic response of the tapered roller bearing are illustrated. For a double row tapered bearing, such as that illustrated in Figure 2, its kinematics is illustrates in Figure 5 by using the animation program SAGA [20]. The visualization of the roller bearing kinematics allows appraising for the correct contact detection and expected kinematics.

![Figure 5: Frames from simulation of a tapered roller bearing with SAGA.](image)

The preliminary results of the dynamic analysis of the tapered roller bearing are depicted in Figure 6, by the total forces on the inner raceway by the rollers and by the forces on a typical roller, depicted in Figure 7.

![Figure 6: Total forces and moments applied on the inner raceway resulting from simulation of a tapered roller bearing in operating conditions.](image)
The dynamic response of the roller bearing is processed by a Fast Fourier Transform to obtain the frequency response, depicted in Figure 8. The frequency response is used to evaluate the basic frequencies of the roller.

In the forthcoming developments of the roller model, by including defects, the frequency response is used to identify such defects via the frequency contents of the roller bearing when compared to healthy bearings.

7 CONCLUSIONS

The work here presented, as part of the project MAXBE, addresses the development of methods and implementation of an efficient bearing dynamic analysis tool, called BearDyn, which allows the analysis of either spherical or tapered roller bearings. The initial position, orientation and velocities of each roller bearing components are first defined ensuring kinematic consistency. The contact problem for a wide number of geometries is addressed to evaluate the normal and tangential forces, due to normal contact and to lubrication. The contact forces are detected and applied between the roller and the raceways, flanges and cage. The necessary models to calculate the tangential forces caused by lubrication are approximated here by the equivalent friction coefficient, which is not validated but simply provides the background for the tangent force models. A preliminary dynamic analysis of a tapered roller bearing model allows extracting the kinematic quantities of importance to its characterization. The frequency response is a fundamental piece to support the development of any monitoring systems in particular. When performing dynamic simulations of the models with BearDyn considering approximate railway conditions, it is identified the need to address the extremely high computational effort. This difficulty clearly identifies the need for the development of robust and efficient computational algorithms, eventually based on the use of parallel computation strategies. In order to allow results to be obtained with BearDyn for a tapered roller bearing in complete working conditions, a
constant time-step solver is used to integrate the equations of motion of the system over time. With this solver, the models used in BearDyn are able to be verified, even for a short integration span, since the interactions between bodies are perceived as occurring as expected, by using verification methods based on the analysis of the dynamic response and visualization tools. In any case, the length of the time responses obtained are not yet enough to draw conclusions about the general dynamic performance of the roller bearing models, but are enough to verify the proper functioning of the dynamic analysis tool. Future work must first address the computational efficiency and after the complete tribological characterization of the contact.

REFERENCES


