Ball Bearing Mechanics

Bernard J. Hamrock
Lewis Research Center
Cleveland, Ohio

and

Duncan Dowson
The University of Leeds
Leeds, England

June 1981
CHAPTER 3
BALL BEARING MECHANICS

The loads carried by ball bearings are transmitted through the balls from one ring to the other. The magnitude of the load carried by an individual ball depends on the internal geometry of the bearing and the location of the ball at any instant. Having determined how a bearing carries load, we can determine how it is distributed among the balls. To do this, we must first develop load-deflection relationships for the ball-race contact. These relationships are developed in Section 3.1 for any type of elliptical contact, such as those found in a ball bearing. The deformation within the contact is, among other things, a function of the ellipticity parameter and the elliptic integrals of the first and second kinds. Simplified expressions that allow quick calculations of the deformation to be made simply from a knowledge of the applied load, the material properties, and the geometry of the contacting elements are presented in Section 3.2.

Most ball bearing applications involve steady-state rotation of either the inner or outer ring, or both. However, the rotational speeds are usually not so great as to cause centrifugal forces or gyroscopic moments of significant magnitude to act
on the ball. Consequently these effects are ignored in analyzing the distribution of radial, thrust, and combined ball loads in Section 3.3.

In high-speed ball bearings the centrifugal force acting on the individual balls can be significant compared with the applied forces acting on the bearing. In high-speed bearings ball gyroscopic moments can also be of significant magnitude, depending on the contact angles, such that the inner-race contact angles tend to increase and the outer-race contact angles tend to decrease. In bearings in which dry friction or boundary lubrication occurs in the conjunction between the balls and races, this can cause a shift of control between races and, in some cases, unstable bearing operation. This does not occur, however, if the conjunctions experience full elastohydrodynamic lubrication. Procedures for evaluating the performance of high-speed ball bearings are developed in Section 3.4. Elastohydrodynamic lubrication of ball-race contacts is not considered in this chapter but will be treated in Chapter 8.

No rolling-element bearing can give unlimited life because of the probability of fatigue. Any structural material subjected to an unlimited succession of repeated or reversed stresses will ultimately fail. Therefore all ball bearings eventually succumb to fatigue, which is manifested by surface distress in the form of flaking of metallic particles. In many cases flaking may begin as a crack below the surface that is propagated to
the surface, where it eventually forms a pit or spall. Fatigue is assumed to have occurred when the first crack or spall is observed on a load-carrying surface. A design criterion for the fatigue life of ball bearings is developed in Section 3.5.

Ball bearings can be lubricated satisfactorily with a small amount of lubricant supplied to the right area within the bearing. Section 3.6 considers the selection of a satisfactory lubricant, as well as describing systems that provide a constant flow of lubricant to the contact.

3.1 Elliptical-Contact Deformation

When an elastic solid is subjected to a load, stresses are produced that increase as the load is increased. These stresses are associated with deformations, which are defined by strains. Unique relationships exist between stresses and their corresponding strains. For elastic solids the stresses are linearly related to the strains, with the constant of proportionality being an elastic constant that adopts different values for different materials. Thus a simple tensile load applied to a bar produces a stress $\sigma_1$ and a strain $\varepsilon_1$, where

$$\sigma_1 = \frac{\text{Load}}{\text{Cross-sectional area}} = \text{Stress in axial direction} \quad (3.1)$$

$$\varepsilon_1 = \frac{\text{Change in length}}{\text{Original length}} = \text{Strain in axial direction} \quad (3.2)$$
Although no stress acts transversely to the axial direction, there will nevertheless be dimensional changes in that direction such that as a bar extends axially, it contracts transversely. The transverse strains \( \varepsilon_2 \) are related to the axial strains \( \varepsilon_1 \) by Poisson's ratio \( v \) such that

\[
v = -\frac{\varepsilon_2}{\varepsilon_1}
\]  

where the negative sign means that the transverse strain will be of the opposite sign to the axial strain. The modulus of elasticity and Poisson's ratio are two important parameters used to describe the material in the analysis of contacting solids.

As the stresses increase within the material, elastic behavior is replaced by plastic flow, in which the material is permanently deformed. The stress state at which the transition from elastic to plastic behavior occurs, known as the yield stress, has a definite value for a given material at a given temperature. In this book elastic behavior alone is considered.

3.1.1 Surface Stresses and Deformation

When two elastic solids are brought together under a load, a contact area develops, the shape and size of which depend on

\[
E = \frac{\sigma_1}{\varepsilon_1} = \text{Elastic constant or modulus of elasticity} \quad (3.3)
\]
the applied load, the elastic properties of the materials, and the curvatures of the surfaces. When the two solids shown in Figure 2.18 have a normal load applied to them, the shape of the contact area is elliptical, with \( a \) being the semimajor and \( b \) the semiminor axis. It has been common to refer to elliptical contacts as point contacts, but since this book deals mainly with loaded contacts, the term elliptical contact is adopted. For the special case where \( r_{ax} = r_{ay} \) and \( r_{bx} = r_{by} \), the resulting contact is a circle rather than an ellipse. Where \( r_{ay} \) and \( r_{by} \) are both infinite, the initial line contact develops into a rectangle when load is applied.

The contact ellipses obtained with either a radial or a thrust load for the ball - inner-race and ball - outer-race contacts in a ball bearing are shown in Figure 3.1. This book is concerned with the conjunctions between solids - with contact areas ranging from circular to rectangular - and with the analysis of contacts in a ball bearing. Inasmuch as the size and shape of these contact areas are highly significant to the successful operation of ball bearings, it is important to understand their characteristics.

Hertz (1881) considered the stresses and deformations in two perfectly smooth, ellipsoidal, contacting elastic solids much like those shown in Figure 2.18. His application of the classical theory of elasticity to this problem forms the basis of stress calculation for machine elements such as ball and rol-
ler bearings, gears, seals, and cams. The following assumptions were made by Hertz (1881):

1. The materials are homogeneous and the yield stress is not exceeded.

2. No tangential forces are induced between the solids.

3. Contact is limited to a small portion of the surface, such that the dimensions of the contact region are small compared with the radii of the ellipsoids.

4. The solids are at rest and in equilibrium (steady state).

Making use of these assumptions, Hertz (1881) was able to obtain the following expression for the pressure within the ellipsoidal contact shown in Figure 3.2:

$$ p = p_{\text{max}} \left[ 1 - \left( \frac{x}{b} \right)^2 - \left( \frac{y}{a} \right)^2 \right]^{1/2} \quad (3.5) $$

If the pressure is integrated over the contact area, it is found that

$$ p_{\text{max}} = \frac{3F}{2\pi ab} \quad (3.6) $$

Equation (3.5) determines the distribution of pressure or compressive stress on the common interface; it is clearly a maximum at the center of the contact and decreases to zero at the periphery.

The ellipticity parameter $k$ can be written in terms of the semimajor and semiminor axes of the contact ellipse as
\[ k = \frac{a}{b} \quad (3.7) \]

Harris (1966) has shown that the ellipticity parameter can be used to relate the curvature difference expressed in equation (2.25) and the elliptic integrals of the first \( \mathcal{F} \) and second \( \mathcal{G} \) kinds as

\[ J(k) = \left[ \frac{2\mathcal{F} - \mathcal{G}(1 + \Gamma)}{\mathcal{G}(1 - \Gamma)} \right]^{1/2} \quad (3.8) \]

where

\[ \mathcal{F} = \int_0^{\pi/2} \left[ 1 - \left(1 - \frac{1}{k^2}\right) \sin^2 \phi \right]^{-1/2} \, d\phi \quad (3.9) \]

\[ \mathcal{G} = \int_0^{\pi/2} \left[ 1 - \left(1 - \frac{1}{k^2}\right) \sin^2 \phi \right]^{1/2} \, d\phi \quad (3.10) \]

A one-point iteration method that was adopted by Hamrock and Anderson (1973) can be used to obtain the ellipticity parameter, where

\[ k_{n+1} = J(k_n) \quad (3.11) \]

The iteration process is normally continued until \( k_{n+1} \) differs from \( k_n \) by less than \( 1 \times 10^{-7} \). Note that the ellipticity parameter is a function of the radii of curvature of the solids only:

\[ k = f(r_{ax}, r_{bx}, r_{ay}, r_{by}) \quad (3.12) \]
That is, as the load increases, the semimajor and semiminor axes of the contact ellipse increase proportionately to each other, so the ellipticity parameter remains constant.

The ellipticity parameter and elliptic integrals of the first and second kinds are shown in Figure 3.3 for a range of the curvature ratio \( \frac{R_y}{R_x} \) usually encountered in concentrated contacts.

When the ellipticity parameter \( k \), the normal applied load \( F \), Poisson's ratio \( v \), and the modulus of elasticity \( E \) of the contacting solids are known, the semimajor and semiminor axes of the contact ellipse and the maximum deformation at the center of the contact can be written from the analysis of Hertz (1881) as

\[
a = \left( \frac{6k^2FR}{\pi E} \right)^{1/3} \tag{3.13}
\]

\[
b = \left( \frac{6kFR}{\pi KE} \right)^{1/3} \tag{3.14}
\]

\[
\delta = \varepsilon \left[ \frac{9FR}{2\pi E} \left( \frac{F}{\pi KE} \right)^2 \right]^{1/3} \tag{3.15}
\]

where

\[
E' = \frac{\frac{2}{1 - v_a^2}}{\frac{1 - v_a^2}{E_a} + \frac{1 - v_b^2}{E_b}} \tag{3.16}
\]

In these equations \( a \) and \( b \) are proportional to \( F^{1/3} \) and \( \delta \) is proportional to \( F^{2/3} \).

Knowing the deformation at the center of the contact and the natural geometrical separation between the solids, equation
(2.35), we can write the deformation at any point within the dry Hertzian contact as

\[ \delta = \delta - S \]  

(3.17)

This equation is used in later chapters to define the film thickness within the conjunction.

3.1.2 Subsurface Stresses

Fatigue cracks usually start at a certain depth below the surface in planes parallel to the direction of rolling. Because of this, special attention must be given to the shear stress amplitude occurring in this plane. Furthermore, a maximum shear stress is reached at a certain depth below the surface. The analysis used by Lundberg and Palmgren (1947) will be used to define this stress.

The stresses are referred to a rectangular coordinate system with its origin at the center of the contact, its z axis coinciding with the interior normal of the body considered, its x axis in the direction of rolling, and its y axis in the direction perpendicular to the rolling direction. In the analysis that follows it is assumed that y = 0.

From Lundberg and Palmgren (1947) the following equations can be written:

\[ \tau_{zx} = \frac{3F}{2\pi} \frac{\cos^2\phi \sin\phi \sin\gamma}{(a^2 \tan^2 \gamma + b^2 \cos^2 \phi)} \]  

(3.18)
The maximum shear stress amplitude is defined as

$$\tau_0 = \left| \tau_{zx} \right|_{\text{max}}$$

The amplitude of the shear stress $\tau_0$ is obtained from

$$\frac{\partial \tau_{zx}}{\partial \phi} = 0$$

$$\frac{\partial \tau_{zx}}{\partial \gamma} = 0$$

For the point of maximum shear stress

$$\tan^2 \phi = t_a$$

$$\tan^2 \gamma = t_a - 1$$

$$\frac{1}{k} = \sqrt{\frac{t_a^2 - 1}{2t_a - 1}}$$

The position of the maximum point is determined by

$$z = z_0 = \zeta^* b$$

$$x = \pm \eta^* b$$

where

$$\zeta^* = \frac{1}{(t_a + 1) \sqrt{2t_a - 1}}$$

$$\eta^* = \frac{t_a}{t_a + 1} \sqrt{\frac{2t_a + 1}{2t_a - 1}}$$
Furthermore the magnitude of the maximum shear stress is given by

\[ \tau_0 = \left( \frac{3F}{2\pi ab} \right) \frac{\sqrt{2t_a - 1}}{2t_a(t_a + 1)} \]  

(3.27)

3.2 Simplified Solution for Elliptical-Contact Deformation

The classical Hertzian solution presented in the previous section requires the calculation of the ellipticity parameter \( k \) and the complete elliptic integrals of the first and second kinds \( \mathcal{F} \) and \( \mathcal{G} \). This entails finding a solution to a transcendental equation relating \( k, \mathcal{F}, \) and \( \mathcal{G} \) to the geometry of the contacting solids, as expressed in equation (3.8). This is usually accomplished by some iterative numerical procedure, as described by Hamrock and Anderson (1973), or with the aid of charts, as shown by Jones (1946).

Brewe and Hamrock (1977) used a linear regression by the method of least squares to obtain simplified equations for \( k, \mathcal{F}, \) and \( \mathcal{G} \). That is, for given sets of pairs of data, \( \{[k_j, (R_y/R_x)_j], j = 1, 2, \ldots, n\} \), a power fit using a linear regression by the method of least squares resulted in the following equation:

\[ k = 1.0339 \left( \frac{R_y}{R_x} \right)^{0.6360} \]  

(3.28)
The asymptotic behavior of \( \varepsilon \) and \( F \) was suggestive of the functional dependence that \( \varepsilon \) and \( F \) might exhibit. As a result a logarithmic and an inverse curve fit were tried for \( \varepsilon \) and \( F \), respectively. The following expressions from Brewe and Hamrock (1977) provide an excellent approximation to the relationships between \( \varepsilon \), \( F \) and \( R_y/R_x \):

\[
\bar{\varepsilon} = 1.0003 + \frac{0.5968}{R_y/R_x} \quad (3.29)
\]

\[
\bar{F} = 1.5277 + 0.6023 \ln \left( \frac{R_y}{R_x} \right) \quad (3.30)
\]

Values of \( \bar{k} \), \( \bar{\varepsilon} \), and \( \bar{F} \) are presented in Table 3.1 and compared with the Hamrock and Anderson (1973) numerically determined values of \( k \), \( \varepsilon \), and \( F \). The agreement is good.

Using these simplified expressions for \( \bar{k} \), \( \bar{\varepsilon} \), and \( \bar{F} \) and equation (3.15) gives the deformation at the center of the contact

\[
\bar{\delta} = \left( \frac{F}{K} \right)^{2/3} \quad (3.31)
\]

where

\[
K = \frac{\pi K E}{4.5 F^3} \left( \frac{R \bar{\delta}}{\bar{F}^3} \right)^{1/2} \quad (3.32)
\]

Note that the load-deflection constant \( K \) is a function of the ball-race geometry and the material properties.

The results of comparing \( \bar{\delta} \) with \( \delta \) are also shown in Table 3.1. The agreement is again quite good. Therefore the deformation at the center of the contact can be obtained di-
rectly from equations (3.28) to (3.32). This valuable approximation eliminates the need to use curve fitting, charts, or numerical methods.

Figure 3.4 shows three different degrees of ball-contact conformity: a ball on a ball, a ball on a plane, and a ball–outer ring contact. Table 3.2 uses this figure to show how the degree of conformity affects the contact parameters. The table shows that $k$ is not exactly equal to unity for the ball-on-ball and ball-on-plane situations because of the approximation represented by equation (3.28). The diameter of the balls is the same throughout, and the material of the solids is steel. The ball–outer-ring contact is representative of a 209 radial ball bearing. A 4.45-N (1-lbf) normal load has been considered for each situation. The maximum pressure decreases significantly as the curvature of the mating surface approaches that of the ball. Table 3.2 shows that the curvature of the mating surfaces is very important in relation to the magnitude of the maximum pressure or surface stress produced. A ball and ring of high conformity are thus desirable from the standpoint of minimizing the stress.

Table 3.2 also shows that the area of the contact $\pi ab$ increases with the conformity of the contacting solids. Although this effect minimizes contact stresses, it can have an undesirable effect on the force of friction, since friction force increases as the contact area and hence the area of the
sheared lubricant increase in a bearing operating under elastohydrodynamic conditions. The curvatures of the bearing races are therefore generally compromises that take into consideration the stress, load capacity, and friction characteristics of the bearing.

In equations (3.24) to (3.27) the location and magnitude of the maximum subsurface shear stress are written as functions of $t_a$, an auxiliary parameter. Furthermore in equation (3.23) the ellipticity parameter is written as a function of $t_a$. The range for $1/k$ is $0 < 1/k < 1$, and the corresponding range for $t_a$ is $1 < t_a < (1 + \sqrt{17})/4$. A linear regression by the method of least squares was used to obtain a simplified formula for $t_a$ in terms of $k$, the ellipticity parameter. That is, for given sets of pairs of data \{[(1 - t_a)_j, (1/k)_j], j = 1, 2, ..., n\}, a power fit using a linear regression by the method of least squares resulted in the following equation:

$$t_a - 1 = 0.3044 \left(\frac{1}{k}\right)^{1.8559}$$

(3.33)

The agreement between this approximate equation and the exact solution is within ±2 percent. The use of equation (3.33) greatly simplifies the determination of the values for the location and magnitude of the maximum subsurface shear stress expressed in equations (3.24) to (3.27).
3.3 Static Load Distribution

Now that a simple analytical expression for the deformation in terms of the load has been determined, it is possible to consider how the bearing load is distributed among the balls within a ball bearing. Most ball bearing applications involve steady-state rotation of either the inner or outer ring, or both. In analyzing the load distribution on the balls, it is usually satisfactory to ignore these effects in most applications. In this section the radial, thrust, and combined load distributions of statically loaded ball bearings are investigated.

For a given ball-race contact the load deflection relationship given in equation (3.31) can be rewritten as

\[ F = K_0 \delta^{3/2} \]  
(3.34)

The total normal approach between two races separated by a ball is the sum of the deformations under load between the ball and both races. Therefore

\[ \delta = \delta_0 + \delta_i \]  
(3.35)

where

\[ \delta_0 = \left( \frac{F}{K_0} \right)^{2/3} \]  
(3.36)

\[ \delta_i = \left( \frac{F}{K_1} \right)^{2/3} \]  
(3.37)
Substituting equations (3.35) to (3.37) into equation (3.34) and solving for \( K \) give

\[
K = \frac{1}{\left[\left(\frac{1}{K_1}\right)^{2/3} + \left(\frac{1}{K_0}\right)^{2/3}\right]^{3/2}} \tag{3.38}
\]

Recall that \( K_1 \) and \( K_0 \) are defined by equation (3.32) and that they are a function of ball-race geometry and material properties alone.

The analysis of deformation and load distribution presented in the following three sections is based on the work of Jones (1946).

3.3.1 Radial Load

A radially loaded ball bearing with radial clearance \( P_d \) is shown in Figure 3.5. In the concentric position shown in Figure 3.5(a) a uniform radial clearance between the balls and the rings of \( P_d/2 \) is evident. The application of a small radial load to the shaft causes the inner ring to move a distance \( P_d/2 \) before contact is made between a ball located on the load line and the inner and outer tracks. At any angle there will still be a small radial clearance \( c \) that, if \( P_d \) is small compared with the radius of the tracks, can be expressed with adequate accuracy by

\[
c = \frac{P_d}{2} (1 - \cos \psi)
\]
On the load line, where \( \psi = 0 \), the clearance is zero, but when \( \psi = 90^\circ \) the clearance retains its initial value of \( \frac{P_d}{2} \).

The application of further load will cause elastic deformation of some of the balls and the elimination of clearance around an arc \( 2\psi_2 \). If the interference or total elastic compression on the load line is \( \delta_{\text{max}} \), the corresponding elastic compression \( \delta_{\psi} \) along a radius at angle \( \psi \) to the load line will be given by

\[
\delta_{\psi} = (\delta_{\text{max}} \cos \psi - c) = \left( \delta_{\text{max}} + \frac{P_d}{2} \right) \cos \psi - \frac{P_d}{2}
\]

Now it is clear from Figure 3.5(c) that \( \delta_{\text{max}} + \frac{P_d}{2} \) represents the total radial displacement of the inner ring or shaft from the concentric position \( \delta \). Hence

\[
\delta_{\psi} = \left( \delta \cos \psi - \frac{P_d}{2} \right)
\] (3.39)

The relationship between load and the elastic compression along the radius at angle \( \psi \) to the load vector is given by equation (3.34) as

\[
F_{\psi} = K\delta_{\psi}^{3/2}
\] (3.40)

Substituting equation (3.39) into this equation gives

\[
F_{\psi} = K\left( \delta \cos \psi - \frac{P_d}{2} \right)^{3/2}
\] (3.41)
For static equilibrium the applied radial load must equal the sum of the components of the ball loads parallel to the direction of the applied load.

\[ F_r = \sum F_\psi \cos \psi \]

Therefore

\[ F_r = K \sum \left( \delta \cos \psi - \frac{p_d}{2} \right)^{3/2} \cos \psi \]

(3.42)

The angular extent of the bearing arc \( 2\psi_x \) in which the balls are loaded is obtained by setting the root expression in (3.42) equal to zero and solving for \( \psi \).

\[ \psi_x = \cos^{-1} \left( \frac{p_d}{2\delta} \right) \]

(3.43)

The summation in equation (3.42) applies only to the angular extent of the loaded region. This equation can be written in integral form as

\[ F_r = \frac{n}{\pi} K_0^{3/2} \int_0^{\psi_x} \left( \cos \psi - \frac{p_d}{2\delta} \right)^{3/2} \cos \psi \, d\psi \]

(3.44)

The integral in this equation can be reduced to a standard elliptic integral by the hypergeometric series and the beta function. If the integral is numerically evaluated directly, the following approximate expression is derived:
This approximate expression fits the exact numerical solution to within ±2 percent for a complete range of $P_d/2\delta$.

The load carried by the most heavily loaded ball is obtained by substituting $\psi = 0^\circ$ in equation (3.42) and dropping the summation sign.

$$F_{\text{max}} = K\delta^{3/2} \left(1 - \frac{P_d}{2\delta}\right)^{3/2}$$  \hspace{1cm} (3.46)

Dividing the maximum ball load (equation (3.46)) by the total applied radial load of the bearing (equation (3.44)), rearranging terms, and making use of equation (3.45) give

$$F_r = \frac{nF_{\text{max}}}{Z}$$  \hspace{1cm} (3.47)

where

$$Z = \frac{\pi \left(1 - \frac{P_d}{2\delta}\right)^{3/2}}{2.491 \left\{1 + \left(\frac{P_d}{2\delta}\right)^2 \right\}^{1/2}}$$  \hspace{1cm} (3.48)

When the diametral clearance $P_d$ is zero, the value of $Z$ becomes 4.37. This is the value derived by Stribeck (1901) for
bearings of zero diametral clearance. The approach used by Stribeck was to evaluate the finite summation $\sum \cos^{5/2} \psi$ for various numbers of balls. He then derived the celebrated Stribeck equation for static load-carrying capacity by writing the more conservative value of 5 for the theoretical value of 4.37:

$$F_r = \frac{nF_{\text{max}}}{5}$$

(3.49)

In using equation (3.49) it should be remembered that $Z$ is considered to be a constant and that the effects of clearance and applied load on load distribution are not taken into account. These effects are, however, considered in obtaining equation (3.47). Note also that the analytical expression for $Z$ in equation (3.48) enables a solution to be obtained without the aid of the charts used by Jones (1946) and Harris (1966).

3.3.2 Thrust Load

The static thrust-load capacity of a ball bearing may be defined as the maximum thrust load that the bearing can endure before the contact ellipse approaches a race shoulder, as shown in Figure 3.6, or the load at which the allowable mean compressive stress is reached, whichever is smaller. Both the limiting shoulder height and the mean compressive stress must be calculated to find the static thrust-load capacity.
The contact ellipse in a bearing race under a thrust load is shown in Figure 3.6. Each ball is subjected to an identical thrust component \( F_t/n \), where \( F_t \) is the total thrust load. The initial contact angle before the application of a thrust load is denoted by \( \beta_f \). Because of the applied thrust, the contact angle becomes \( \beta \). The normal ball thrust load \( F_t \) acts at this contact angle and is written as

\[
F_t = \frac{F}{n \sin \beta} \tag{3.50}
\]

A cross section through an angular-contact bearing under a thrust load \( F_t \) is shown in Figure 3.7. Both races are assumed to be rigidly mounted, that is, incapable of radial deformation. From this figure the contact angle after the thrust load has been applied can be written as

\[
\beta = \cos^{-1} \left( \frac{D - P_d/2}{D + \delta} \right) \tag{3.51}
\]

The initial contact angle was given in equation (2.9). Using that equation and rearranging terms in equation (3.51) give

\[
\delta = D \left( \frac{\cos \beta_f}{\cos \beta} - 1 \right) \tag{3.52}
\]

From equation (3.34) we can write

\[
F_t = KD^{3/2} \left( \frac{\cos \beta_f}{\cos \beta} - 1 \right)^{3/2} \tag{3.53}
\]

where

\[
K = \frac{\eta kE'}{4.5 \bar{R}^3} \tag{3.54}
\]

21
and \( \varepsilon, \xi, \) and \( \eta \) are given by equations (3.28), (3.29), and (3.30), respectively.

From equations (3.50) and (3.53)

\[
\frac{F}{nKD^{3/2}} = \sin \beta \left( \frac{\cos \beta_f}{\cos \beta} - 1 \right)^{3/2}
\] (3.55)

Equation (3.55) can be solved numerically by the Newton-Raphson method. The iterative equation to be satisfied is

\[
b' = b + \frac{F}{nKD^{3/2}} - \sin \beta \left( \frac{\cos \beta_f}{\cos \beta} - 1 \right)^{3/2} \frac{\cos \beta \left( \frac{\cos \beta_f}{\cos \beta} - 1 \right)^{3/2} + \frac{3}{2} \cos \beta_f \tan^2 \beta \left( \frac{\cos \beta_f}{\cos \beta} - 1 \right)}{\cos \beta \left( \frac{\cos \beta_f}{\cos \beta} - 1 \right)^{1/2}}
\] (3.56)

This equation is satisfied when \( b' - b \) is essentially zero.

When a thrust load is applied, the shoulder height is limited to the distance by which the pressure-contact ellipse can approach the shoulder. As long as the following inequality is satisfied, the pressure-contact ellipse will not exceed the shoulder height limit:

\[
\theta > b + \sin^{-1} \left( \frac{a}{r_d} \right)
\] (3.57)

From Figure 2.17 and equation (2.15) the angle used to define the shoulder height \( \theta \) can be written as

\[
\theta = \cos^{-1} \left( 1 - \frac{s}{r_d} \right)
\] (3.58)

22
From Figure 3.7 the axial deflection $\delta_t$ corresponding to a thrust load can be written as

$$\delta_t = (D + \delta)\sin \beta - D \sin \beta_f$$  \hspace{1cm} (3.59)

Substituting equation (3.52) into equation (3.59) gives

$$\delta_t = \frac{D \sin(\beta - \beta_f)}{\cos \beta}$$  \hspace{1cm} (3.60)

Having determined $\beta$ in equation (3.56) and $\beta_f$ in equation (2.9), we can easily evaluate the relationship for $\delta_t$.

3.3.3 Combined Load

For a combined radial and axial load on a ball bearing we consider the relative displacements of the inner and outer rings. We assume that negligible misalignment of the bearing can occur. The displacements are therefore limited to an axial displacement $\delta_t$ and a radial displacement $\delta_r$. The races are therefore constrained to relative movement in parallel planes. The end result of this combined loading is shown in Figure 3.8. Note the difference between this figure and Figure 3.7, which represents axial loading alone. As was found when dealing with a purely radial load, the radial displacement is a function of the ball position relative to the applied load.

From Figure 3.8

$$(D + \delta)^2 = (D \cos \beta_f + \delta_r \cos \psi)^2 + (D \sin \beta_f + \delta_t)^2$$  \hspace{1cm} (3.61)
or
\[
\delta = D \left\{ \left[ \left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2 \right]^{1/2} \right\} - 1 \quad (3.62)
\]
Substituting this equation into equation (3.34) gives
\[
F = KD^{3/2} \left[ \left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2 \right]^{1/2} \left( \frac{\sin \beta_f + \frac{\delta_t}{D}}{\left[ \left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2 \right]^{1/2}} \right)^{3/2}
\]
(3.63)
where \( K \) is defined in equation (3.54). Also from Figure 3.8
\[
\sin \beta = \frac{\sin \beta_f + \frac{\delta_t}{D}}{\left[ \left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2 \right]^{1/2}}
\]
(3.64)
\[
\cos \beta = \frac{\cos \beta_f + \frac{\delta_r}{D} \cos \psi}{\left[ \left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2 \right]^{1/2}}
\]
(3.65)
The normal ball load \( F \), which acts at the contact angle \( \beta \) (along the \( D + \delta \) line in Figure 3.8), can be resolved into two components. One is the thrust force \( F_t \) parallel to the bearing axis, and the other is the radial force \( F_r \). The thrust component \( F_t \) can be written as
\[
F_t = F \sin \beta
\]
(3.66)
By using equations (3.63) and (3.64) this relationship becomes

\[
\frac{K D^{3/2} \left( \sin \beta_f + \frac{\delta_t}{D} \right) \left\{ \left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2 \right\} - 1}{\left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2}^{1/2} \left( \frac{\cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2 \right\}^{1/2}^{3/2}.
\]

(3.67)

The radial component of load \( F_r \) can be written as

\[
F_r = F \cos \beta \cos \psi
\]

(3.68)

From equations (3.63) and (3.65) this expression can be written as

\[
F_r = \frac{K D^{3/2} \left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right) \cos \psi \left\{ \left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2 \right\} - 1}{\left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2}^{1/2} \left( \frac{\cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2 \right\}^{1/2}^{3/2}.
\]

(3.69)

For the bearing to be in equilibrium after displacement, the following conditions must be satisfied:

Applied axial load - \( \sum F_t \) -

\[
K D^{3/2} \left( \sin \beta_f + \frac{\delta_t}{D} \right) \left\{ \left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2 \right\} - 1}{\left( \cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2}^{1/2} \left( \frac{\cos \beta_f + \frac{\delta_r}{D} \cos \psi \right)^2 + \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2 \right\}^{1/2}^{3/2}.
\]

(3.70)
Applied radial load = \( \sum F_r \) –

\[
\sum_{r=1}^{n} F_r = \frac{nKU^{3/2}}{\pi} \frac{1}{1_t}
\]

\[
F_r = \frac{nKU^{3/2}}{\pi} \frac{1}{1_r}
\]

The extent of the load zone \( \psi_r \) is obtained by setting the numerator in these equations to zero or

\[
\psi_r = \cos^{-1} \left\{ \frac{1 - \left( \sin \beta_f + \frac{\delta_t}{D} \right)^2}{\delta_r/D} - \cos \beta_f \right\}
\]

(3.72)

Under certain conditions of axial preload and radial displacement the value of \( \cos \psi_r \) as determined by equation (3.72) will be less than -1. This indicates that the loaded zone extends completely around the pitch circle. In such cases the limiting value \( \psi_r \) is taken as \( \pi \).

Equations (3.71) and (3.72) can be generalized to include any number of balls by the following:

\[
\sum_{r=1}^{n} F_r = \frac{nKU^{3/2}}{\pi} \frac{1}{1_t}
\]

(3.73)

\[
F_r = \frac{nKU^{3/2}}{\pi} \frac{1}{1_r}
\]

(3.74)
where

\[ I_t = \int_0^{\psi_t} \frac{(\sin \theta_f + \frac{\delta_t}{D}) \left[ \left( \cos \theta_f + \frac{\delta_r}{D} \cos \phi \right)^2 + \left( \sin \theta_f + \frac{\delta_t}{D} \right)^2 \right]^{1/2} - 1}{\left[ \left( \cos \theta_f + \frac{\delta_r}{D} \cos \phi \right)^2 + \left( \sin \theta_f + \frac{\delta_t}{D} \right)^2 \right]^{1/2}} \, d\phi \]

(3.75)

\[ I_r = \int_0^{\psi_r} \frac{(\cos \theta_f + \frac{\delta_r}{D} \cos \phi) \cos + \left[ \left( \cos \theta_f + \frac{\delta_r}{D} \cos \phi \right)^2 + \left( \sin \theta_f + \frac{\delta_t}{D} \right)^2 \right]^{1/2} - 1}{\left[ \left( \cos \theta_f + \frac{\delta_r}{D} \cos \phi \right)^2 + \left( \sin \theta_f + \frac{\delta_t}{D} \right)^2 \right]^{1/2}} \, d\phi \]

(3.76)

Note that these integrals are functions of the three parameters \( \theta_f, \delta_t/D, \) and \( \delta_r/D. \) These integrals are hyperelliptic integrals that cannot be reduced to standard form to permit solution in terms of elliptic functions and must therefore be evaluated numerically on a digital computer. Having determined \( \delta_t/D \) and \( \delta_r/D \) from equations (3.73) and (3.74), we can obtain the normal ball load and operating contact angle at any ball position \( \psi \) from equations (3.63) and (3.64).
3.4 High-Speed Bearing Load

For ball bearings that operate at modest speeds, as considered in the preceding section, the centrifugal force on the ball is so negligible that the only forces that keep the ball in equilibrium are the two contact forces resulting from the externally applied load. For such conditions the contact forces are equal and opposite, and the inner- and outer-race contact angles are approximately equal. The present section deals with high-speed bearings, where the centrifugal force developed on the balls becomes significant and the inner- and outer-race contact angles are no longer equal. An angular-contact bearing is analyzed since the equations developed can be applied to other types of ball bearings. A combined radial and axial load is considered, but misalignment of the inner and outer rings is excluded. The material in this section was first developed by Jones (1956).

When a ball bearing operates at high speed, the body forces resulting from the ball's motion become significant and must be considered in any analysis. Figure 3.9 shows the forces and moments acting on a ball in a high-speed ball bearing. The operating contact angle at the outer contact is less than that at the inner contact because of appreciable centrifugal force and gyroscopic moment. In this figure, as with the rest of the
book, subscript $i$ refers to the inner race and subscript $o$ to the outer race.

An exaggerated view in Figure 3.10 shows the ball fixed in the plane of the paper and rotating about its own center with an angular velocity $\omega_B$ directed at an angle $\zeta$ to the bearing centerline. The inner and outer races rotate about the bearing axis with the angular velocities $\omega_i$ and $\omega_o$ relative to the separator. For the linear velocity of the races to be equal to the ball velocity at the contact, the following relationships must be satisfied:

$$\omega_B = \frac{(d_i' + d \cos \beta_o)\omega_o}{d \cos(\beta_o - \zeta)}$$  (3.77)

$$\omega_B = -\frac{(d_i' - d \cos \beta_i)\omega_i}{d \cos(\beta_i - \zeta)}$$  (3.78)

If the outer race is stationary, the ball will orbit the bearing axis with an angular velocity $\omega_c$, where

$$\omega_c = -\omega_o$$  (3.79)

Then the absolute angular velocity of the inner race is $\Omega_i$, where

$$\Omega_i = \omega_i + \omega_c = \omega_i - \omega_o$$  (3.80)

Therefore for a stationary outer race and a rotating inner race the following can be written:

$$\omega_B = \frac{-\Omega_i}{d \left[ \frac{\cos(\beta_i - \zeta)}{d_i' - d \cos \beta_i} + \frac{\cos(\beta_o - \zeta)}{d_i' + d \cos \beta_o} \right]}$$  (3.81)
Similarly for a stationary inner race and a rotating outer race

$$\omega_c = \frac{\Omega_i}{1 + \left(\frac{d' + d \cos \beta_0}{d' - d \cos \beta_1}\right) \cos(\beta_i - \xi) \cos(\beta_0 - \xi)}$$  \hspace{1cm} (3.82)

For simultaneous rotation of the outer and inner races

$$\omega_B = \frac{\Omega_0}{d \left[\frac{\cos(\beta_i - \xi)}{d' - d \cos \beta_1} + \frac{\cos(\beta_0 - \xi)}{d' + d \cos \beta_1}\right]}$$  \hspace{1cm} (3.83)

$$\omega_c = \frac{\Omega_0}{1 + \left(\frac{d' - d \cos \beta_0}{d' + d \cos \beta_1}\right) \cos(\beta_0 - \xi) \cos(\beta_i - \xi)}$$  \hspace{1cm} (3.84)

For an arbitrary choice of $\xi$, the ball will spin relative to the race about the normal at the center of the contact area.

It is clear from this analysis that the spin of the ball may be different relative to each race, and this prompted Jones (1956) to introduce the concept of race control. If Coulomb friction or boundary lubrication prevails in the conjunctions between the ball and the inner and outer races, the conjunction subjected to the least torque will be prevented from spinning by
friction while the other conjunction experiences spin. The contact at which no spin occurs is called the controlling race.

If a lubricating film exists between the ball and each of the races, each of the conjunctions can experience spin, and the relative motion between the ball and the races is determined by the equilibrium of the torques resulting from viscous tractions within the lubricant. The problem of predicting viscous tractions in elastohydrodynamic films still requires further work, but the recent development of the understanding of lubricant rheology in EHL conjunctions outlined in Chapter 10, together with the ability to predict film thickness outlined in this text, indicates that a complete solution to the problem may not be far away.

Spin in the conjunctions between a ball and the races of a bearing is important from the point of view of energy losses and heat generation. For this reason the race-control theory originated by Jones will be outlined here, although it must be recalled that it was developed for dry friction or boundary lubrication conditions before solutions to the elastohydrodynamic lubrication problem became available. The elastohydrodynamic lubrication of ball bearings will be considered in Chapter 8, Section 8.9.

From Figure 3.10 the ball spin rotational velocities at the inner and outer races can be written as

\[ \omega_{s_i} = \omega_i \sin \beta_i - \omega_B \sin(\beta_i - \zeta) \]  

(3.87)
\[ \omega_{so} = -\omega_0 \sin \beta_0 + \omega_B \sin(\beta_0 - \zeta) \quad (3.88) \]

The race-control concept of Jones (1956) assumes that all the spin occurs at one contact and none at the other. The contact at which no spin occurs is called the controlling race. Lightly loaded bearings may depart somewhat from this situation.

If \( \omega_{si} \) and \( \omega_{so} \) are made zero in equations (3.87) and (3.88), respectively, the following will result:

Inner-race control:

\[
\zeta = \tan^{-1} \left( \frac{d' \sin \beta_i}{d' \cos \beta_i - d} \right) \quad (3.89)
\]

Outer-race control:

\[
\zeta = \tan^{-1} \left( \frac{d' \sin \beta_0}{d' \cos \beta_0 + d} \right) \quad (3.90)
\]

The existence of a particular type of control depends on the relative torques required to produce spin at the two contacts.

The frictional heat generated at the ball-race contacts, where slip takes place, is

\[ H_f = \omega_s M_s \quad (3.91) \]

where \( M_s \) is the torque required to produce spin. Poritsky, et al. (1947) integrated the friction force over the contact ellipse to obtain \( M_s \) as

\[ M_s = \frac{3}{8} \mu F a \overline{e} \quad (3.92) \]
where
\[ \mu = \text{coefficient of sliding friction} \]
\[ F = \text{contact load} \]
\[ a = \text{semimajor axis of contact ellipse obtained from equation (3.13)} \]
\[ \bar{e} = \text{elliptic integral of second kind obtained from equation (3.29)} \]

Equation (3.92) can be written for both the outer- and inner-race contacts. Outer-race control will exist if \( M_{so} \geq M_{si} \). Inner-race control will exist if \( M_{so} < M_{si} \). In a given ball bearing that operates under a given speed and load, rolling will take place at one race and spinning at the other. Rolling will therefore take place where \( M_s \) is greater because of the greater gripping action.

The positions of the ball center and the race curvature centers at angular position \( \psi \) are shown in Figure 3.11 with and without an applied combined load. In this figure the outer-race curvature is fixed. When speeds are high and the centrifugal force is appreciable, the inner and outer-race contact angles become dissimilar. This results in the outer-race contact angle \( \beta_o \) being less than the initial contact angle \( \beta_f \), as shown in Figure 3.11.

In accordance with the relative axial displacement of the inner and outer rings \( \delta_t \), the axial distance between the loci of inner- and outer-race curvature centers is
Furthermore in accordance with a relative radial displacement of the ring centers \( \delta_r \), the radial displacement between the loci of the race curvature center at each ball location is

\[
L_2 = D \cos \beta_f + \delta_r \cos \psi
\]  

where

\[
\psi = \frac{2\pi(j - 1)}{n} \quad j = 1, 2, ..., n
\]

and \( n \) is the number of balls. From Figure 3.11 the following equations can be written:

\[
\cos \beta_0 = \frac{L_4}{d(f_0 - 0.5) + \delta_0}
\]  

\[
\sin \beta_0 = \frac{L_3}{d(f_0 - 0.5) + \delta_0}
\]  

\[
\cos \beta_i = \frac{L_2 - L_4}{d(f_i - 0.5) + \delta_i}
\]  

\[
\sin \beta_i = \frac{L_1 - L_3}{d(f_i - 0.5) + \delta_i}
\]

The following relationships can thus be written with reference to Figure 3.11:

\[
L_4^2 + L_3^2 - [d(f_0 - 0.5) + \delta_0]^2 = 0
\]  

\[
(D \cos \beta_f + \delta_r \cos \psi - L_4)^2 + (D \sin \beta_f + \delta_t - L_3)^2
\]

\[
- [(f_i - 0.5)d + \delta_i]^2 = 0
\]
The forces and moments acting on the ball are shown in Figure 3.9. The normal forces shown in this figure can be written from equation (3.34) as

\[ F_0 = K_0 \delta_0^{3/2} \quad (3.102) \]

\[ F_i = K_i \delta_i^{3/2} \quad (3.103) \]

Equilibrium of forces in the horizontal and vertical directions requires that

\[ F_0 \sin \beta_0 - F_i \sin \beta_i - \frac{2M}{d} \left[ \lambda \cos \beta_0 - (1 - \lambda) \cos \beta_i \right] = 0 \quad (3.104) \]

\[ F_0 \cos \beta_0 - F_i \cos \beta_i + \frac{2M}{d} \left[ \lambda \sin \beta_0 - (1 - \lambda) \sin \beta_i \right] - F_c = 0 \quad (3.105) \]

where

\[ \lambda = 1 \quad \text{for outer-race control} \]
\[ \lambda = 0 \quad \text{for inner-race control} \]

The centrifugal force in equation (3.105) can be written as

\[ F_c = \frac{1}{2} m d_e^2 \omega_c^2 \quad (3.106) \]

where

\[ d'_e = d_e + 2L_4 - 2d(f_0 - 0.5) \cos \beta_f \quad (3.107) \]

and \( m \) is the mass of the ball. Also the gyroscopic moment in equations (3.104) and (3.105) can be written as

\[ M_g = I_p \omega_p \omega_c \sin \zeta \quad (3.108) \]
where $I_p$ is mass moment of inertia of the ball. From these relationships, equations (3.104) and (3.105) can be written as

$$0 = \frac{K_0 \delta_0^{3/2} L_3}{d(f_0 - 0.5) + \delta_0} - \frac{K_i \delta_i^{3/2} (D \sin \theta_f + \delta_t - L_3)}{d(f_i - 0.5) + \delta_i}$$

$$- \frac{2}{d} I_p \omega_o \omega_c \sin \xi \left[ \frac{\lambda L_4}{d(f_0 - 0.5) + \delta_0} \right]$$

$$+ \frac{(1 - \lambda)(D \cos \theta_f + \delta_r \cos \psi - L_4)}{d(f_i - 0.5) + \delta_i} \right]$$

Equations (3.100), (3.101), (3.109), and (3.110) can be solved simultaneously for $L_3$, $L_4$, $\delta_o$, and $\delta_i$ at each ball location once the values of $\delta_t$ and $\delta_r$ are assumed. The Newton-Raphson method is generally used to solve these simultaneous nonlinear equations.

To find how good the initial guess of the values of $\delta_r$ and $\delta_t$ is, a condition of equilibrium applied to the entire bearing is used

$$F_t = \sum_{j=1}^{n} \left[ F_{ij} \sin \beta_{ij} - \frac{2(1 - \lambda_j) M_{ij}}{d} \cos \beta_{ij} \right] = 0 \quad (3.111)$$
\[ F_r = \sum_{j=1}^{n} \left[ F_{ij} \cos \beta_{ij} - \frac{2(1 - \lambda_i)M_{ij}}{d} g_j \sin \beta_{ij} \right] \cos \psi_j = 0 \] (3.112)

Having computed values for \( L_3, L_4, \delta_i, \) and \( \delta_0 \) at each ball position and knowing \( F_t \) and \( F_r \) as input conditions, we can obtain the values of \( \delta_t \) and \( \delta_r \) from equations (3.111) and (3.112). After obtaining these values for \( \delta_t \) and \( \delta_r \) it is necessary to repeat the calculations for \( L_3, L_4, \delta_i, \) and \( \delta_0 \) at each ball position until the assumed values of \( \delta_t \) and \( \delta_r \) agree with these values found from equations (3.111) and (3.112).

3.5 Fatigue Life

Ball bearings can fail from numerous causes, including faulty handling and fitting, wear associated with dirt, damage to the races or separators, and fatigue. However, if they survive all the other hazards, ball bearings eventually fail because of fatigue of the bearing material. For this reason the subject of fatigue calls for special consideration. Fatigue is caused by the repeated stresses developed in the contact areas between the ball and the races and manifests itself as a fatigue crack starting at or below the surface. The fatigue crack propagates until a piece of the race or ball material spalls out.
and produces the failure. A typical fatigue spall is shown in Figure 2.24. On a microscale we can surmise that there will be a wide dispersion in material strength, or resistance to fatigue, because of inhomogeneities in the material. Bearing materials are complex alloys and are thus neither homogeneous nor equally resistant to failure at all points. Therefore the fatigue process can be expected to be one in which a group of apparently identical ball bearings subjected to identical loads, speeds, lubrication, and environmental conditions exhibit wide variations in failure times. For this reason the fatigue process must be treated statistically. That is, the fatigue life of a bearing is normally defined in terms of its statistical ability to survive for a certain period of time.

3.5.1 Load Factor

The predominant factor in determining the fatigue life of a ball bearing is the load factor. The relationship between life and load developed here is based on a well-lubricated system and a bearing made of air-melted materials. To predict how long a particular bearing will run under a specific load, two essential pieces of information are required:

(1) An accurate, quantitative estimate of the life dispersion or scatter
(2) An expression for the dynamic load capacity or ability of the bearing to endure a given load for a stipulated number of stress cycles or revolutions.

A typical distribution of the fatigue life of identical ball bearings operating under nominally identical conditions is presented in Figure 3.12. This figure shows that the number of revolutions that a bearing can complete with 100 percent probability of survival, $\xi_a = 1$, is zero. Alternatively the probability of any bearing in the population having infinite endurance is zero. Failure is normally assumed to have occurred when the first spall is observed on a load-carrying surface.

Bearing manufacturers have chosen to use one or two points on the curve in Figure 3.12 to describe bearing endurance:

1. The fatigue life that 90 percent of the bearing population will endure ($L_{10}$)
2. The median life, that is, the life that 50 percent of the population will endure ($L_{50}$)

Bearing manufacturers almost universally refer to a "rating life" as a measure of the fatigue endurance of a given bearing operating under given load conditions. This "rating life" is the estimated $L_{10}$ fatigue life of a large population of such bearings operating under the specified loading.
Fatigue life is generally stated in millions of revolutions. As an alternative it may be and frequently is given in hours of successful operation at a given speed.

Weibull (1949) has postulated that the fatigue lives of a homogeneous group of ball bearings are dispersed according to the following relation:

\[
\ln \ln \frac{1}{\xi_a} = e \ln \frac{L}{A} \quad (3.113)
\]

where

- \( L \) = life, millions of revolutions
- \( e \) = dispersion exponent (slope of Weibull plot) or measure of scatter in bearing lives
- \( A \) = constant, such that \( e \ln A \) is vertical intercept on Weibull plot when \( L = 1 \)

The fatigue life \( L \) in equation (3.113) is the \( L_{10} \) life, but it is simply referred to here and throughout the remainder of the book as fatigue life \( L \).

The so-called Weibull distribution given in equation (3.113) results from a statistical theory of strength based on the theory of probability, where the dependence of strength on volume is explained by the dispersion in material strength. This is the "weakest link" theory. Equation (3.113) is used for plotting fatigue failures to determine the \( L_{10} \) lives. A typical Weibull plot of bearing fatigue failures is given in Figure 3.13. The experimental results shown as circular points
in this figure confirm that bearing lives conform well with the Weibull distribution and that the bearing fatigue data will plot as a straight line.

With a technique for treating life dispersion now available, an expression for the dynamic load capacity that a bearing can carry for a given number of stress cycles with a given probability of survival must be derived. From the weakest-link theory we get the relationship between the life of an assembly (the bearing) and its components (the inner and outer rings):

\[
\frac{1}{L_e} = \frac{1}{L_i} + \frac{1}{L_o} \quad (3.114)
\]

For ball bearings \( e = 10/9 \). The following expression can be written for the fatigue life of elliptical contacts

\[
L = \left( \frac{C}{F} \right)^3 \quad (3.115)
\]

where

- \( F \) = static load capacity
- \( C \) = dynamic load capacity

Using this equation and changing the fatigue life from millions of revolutions to hours of successful operation at a given speed, we can write equation (3.114) as

\[
L = \frac{1 \times 10^6}{60 \cdot N} \frac{1}{\left[ \left( F_i \frac{10}{3} \right) + \left( F_o \frac{10}{3} \right)^{0.9} \right]} \quad (3.116)
\]
The static loads $F_i$ and $F_0$ can be obtained from either Section 3.3 or 3.4 for the appropriate load and speed condition. In equation (3.116) $N$ is expressed in revolutions per minute, and the fatigue life is expressed in hours of successful operation at the given speed $N$.

From Lundberg and Palmgren (1947) the dynamic load capacity of the inner ring can be written as

$$C_i = 84,000 \left( \frac{T_i}{T_i^0} \right)^{3.1} \left( \frac{\phi_i}{\phi_i^0} \right)^{0.4} \left( \frac{2\varepsilon_i R_i}{R_i} \right)^{2.1} \frac{k_i}{d_i^{0.3}} \left( \overline{u_i} \right)^{-1/3}$$

where

$$T_i = \left( \frac{\tau_0}{p_{\text{max}}^i} \right)$$

$$T_i^0 = \left( \frac{\tau_0}{p_{\text{max}}^i} \right)_{k=1}$$

$\tau_0$ = maximum orthogonal subsurface shear stress

$\phi_i$ = ratio of depth of maximum shear stress of inner ring to semiminor axis of contact ellipse, $z_0/b$

$\phi_i^0 = (\phi)_{k=1}$

$\overline{u_i}$ = number of stress cycles per revolution of inner ring

With proper changing of subscripts from $i$ to $o$, equation (3.117) can represent the dynamic load capacity of the outer ring $C_o$.

The number of stress cycles per revolution $\overline{u}$ denotes the number of balls that pass a given point (under load) on the race.
of one ring while the other ring has turned through one complete revolution. Therefore the number of balls passing a point on the inner ring per unit of time is

\[ \bar{u}_i = \frac{n}{2d_e} (d'_e + d \cos \beta_i) \quad (3.118) \]

\[ \bar{u}_o = \frac{n}{2d_e} (d'_e - d \cos \beta_o) \quad (3.119) \]

In equation (3.117) the diameters of the inner and outer races are written as

\[ d_i = d'_e - d \cos \beta_i \quad (3.120) \]

\[ d_o = d'_e + d \cos \beta_o \quad (3.121) \]

Hamrock and Anderson (1973) found that for most ball bearing configurations the variation of \( T \) and \( \phi \) is such that the following approximation can be made:

\[ \left( \frac{T}{T^1} \right)^{3.1 \left( \frac{\phi}{\phi_1} \right)^{0.4}} \right) = 0.718 \quad (3.122) \]

Table 3.3 presents corresponding values for \( 1/k \), \( T \), and \( \phi \), as well as values of \( (T/T^1)^{3.1 \left( \phi/\phi_1 \right)^{0.4}} \) for corresponding values of \( 1/k \). From these values the following simple formula can be written:

\[ \left( \frac{T}{T^1} \right)^{3.1 \left( \frac{\phi}{\phi_1} \right)^{0.4}} = 0.706 + 0.3 \left( \frac{1}{k} \right)^{1.92} \quad (3.123) \]
Table 3.3 also shows the good accuracy of this approximate formula.

In equation (3.117) the curvature sum $R$ can be obtained from equation (2.24), and the elliptic integral of the second kind $\varpi$ and the ellipticity parameter $k$ can be obtained from equations (3.29) and (3.28), respectively. By making use of the static loads $F_i$ and $F_o$ obtained from either Section 3.3 or 3.4 and equations (3.117) to (3.123), the fatigue life in operating hours of the bearing can be obtained from equation (3.116).

The dynamic load capacity $C$ just developed can be used to determine the relative importance of centrifugal effects in ball bearings of different sizes. This was done by Hamrock and Anderson (1973) by comparing the ratio of $d^3N^2$ to the dynamic load capacity $C$. In the previous chapter it was noted that $d_b$ is the bore diameter in millimeters and $N$ is the rotational speed in revolutions per minute. The factor $d^3N^2$ is proportional to the centrifugal force, and the dynamic capacity is a measure of the load capacity of the bearing. For extra-light series angular-contact ball bearings operating at a value of $d_bN$ of 3 million, Table 3.4 shows the ratio of $d^3N^2$ to dynamic capacity $C$ for four bore diameters $d_b$. Centrifugal effects are shown to be relatively more severe in small bearings when $d_bN$ is kept constant.
The effect of race conformity ratio $f$ on fatigue life at high operating speeds is shown in Figure 3.14. This figure was obtained from Winn, et al. (1974) for a 20-mm-bore ball bearing operating at 120,000 rpm. Note that an increase in outer-race curvature brings about a substantial decrease in fatigue life. On the other hand an increase in inner-race curvature does not affect the life to any appreciable degree. The reason for this is that at high speeds the centrifugal force acts against the outer race. It is thus important in optimizing the bearing life in high-speed applications that the outer-race conformity ratio should remain as low as possible. Conformity expressed by a curvature ratio $f$ of 0.515 to 0.520 represents the lowest threshold of present manufacturing practices.

The contact angle in ball bearings is extremely important inasmuch as it critically affects the bearing stiffness and life. Typical variations of fatigue life with initial contact angle $\theta_f$ for a medium-size bearing operating at a value of $d_bN$ of 1.5 million are shown in Figure 3.15. The contact-angle range suggested in Figure 3.15 is typical of bearings operating at high speeds.

In recent years better understanding of ball bearing design, materials, processing, and lubrication has permitted an improvement in bearing performance. This is reflected in either higher bearing reliability or longer expected lives than those obtained from equation (3.116) or ball bearing catalogs. As a
result Bamberger, et al. (1971) arrived at an expression for the adjusted bearing fatigue life

\[ L_a = \tilde{D}\tilde{E}\tilde{F}\tilde{G} \]  

(3.124)

where

\[ \tilde{D} = \text{material factor} \]

\[ \tilde{E} = \text{processing factor} \]

\[ \tilde{F} = \text{lubrication factor} \]

\[ \tilde{G} = \text{hardness factor} \]

The next three sections deal with these factors.

3.5.2 Lubrication Factor

If a ball bearing is adequately designed and lubricated, the rolling surfaces can be separated by a lubricant film. Endurance testing of bearings, as reported by Tallian, et al. (1965), has demonstrated that when the lubricant film is thick enough to separate the two contacting bodies, fatigue life of the bearing is greatly extended. Conversely, when the film is not thick enough to provide full separation between the asperities in the contact zone, the life of the bearing is adversely affected by the high shear resulting from direct metal-to-metal contact. An expression for the film thickness in ball bearings is developed later, but it is convenient to illustrate its effect on fatigue life in this section.
To establish the effect of film thickness on the life of any given bearing, we first calculate the film parameter \( A \). The relationship between \( A \) and the film thickness \( h \) is

\[
A = \frac{h}{\left(f_r^2 + f_b^2\right)^{1/2}} \quad (3.125)
\]

where:

\( f_r \) = rms surface finish of race
\( f_b \) = rms surface finish of ball

A more detailed discussion of surface topography is given in Section 4.1, and the rms is defined by equation (4.2).

With the film parameter \( A \) known, Figure 3.16 can be used to determine the lubrication factor \( \tilde{F} \). Note from this figure that when the film parameter values fall below approximately 1.2, the bearing fatigue life is adversely affected since \( \tilde{F} \) is less than 1. Conversely, when the values of \( A \) are between 1.2 and 3, bearing fatigue life is appreciably extended. Film parameters higher than 3 do not yield any further improvement in the lubrication factor \( \tilde{F} \) mainly because at these values of \( A \) the lubricant film is thick enough to separate the extreme peaks of the interacting surfaces.

3.5.3 Material Factor

Bamberger, et al. (1971) have shown that bearing materials can significantly affect the ultimate performance of a bearing.
As mentioned in Chapter 2 the most frequently used steel for ball bearings is AISI 52100. The dynamic load capacity, as calculated from equation (3.117) or any bearing manufacturer's catalog, is based on air-melted 52100 steel that has been hardened to 58 Rockwell C ($R_c$). Because of improvements in the quality of air-melted steels, Bamberger, et al. (1971) suggested the value of the material factor $\tilde{D}$ shown in Table 3.5. Factors taking into account vacuum remelting, hardness, and other processing variables are considered separately. Many of the materials in this table were discussed in Section 2.4, and the chemical compositions of many of these steels are given in Table 2.1.

3.5.4 Processing Factors

Improvements in processing techniques have also extended fatigue life. The various melting practices have been discussed in Section 2.4.1. Zaretsky, et al. (1969) found that consumable-electrode vacuum remelting (CVM) gave up to 13 times longer life than air melting. However, Bamberger, et al. (1971) recommended that a processing factor $\tilde{E}$ of 3 be used for all CVM bearing steels. This value may be somewhat conservative, but the confidence factor for achieving this level of improvement is high.

Another processing factor that seriously affects bearing fatigue life is material hardness. The minimum recommended
hardness for ball bearing steels is $58 R_c$. A drop in hardness from that value because of either poor heat treatment or high operating temperatures will appreciably shorten the bearing fatigue life, as pointed out by Bamberger, et al. (1971). To enable an estimate to be made of the effects of hardness change on bearing life, a hardness factor $G$ is defined as

$$G = \left(\frac{R_c}{58}\right)^{10.8}$$  \hspace{1cm} (3.126)

where $R_c$ is the operational hardness of the bearing material. Note that the relationship presented by equation (3.126) indicates that bearing life is highly sensitive to changes in hardness. Thus, for example, a two-point drop in hardness to $56 R_c$ will cause a 32 percent drop in bearing fatigue life.

Once the various factors in equation (3.124) have been defined, the adjusted fatigue life $L_a$ can be calculated from that equation. This equation enables the designer to arrive at a more realistic estimate of bearing fatigue life.

3.6 Bearing Lubrication

Without adequate lubrication of the ball-race conjunction, various degrees of damage will result to the rolling elements or the races, or both. These include the development of scuffing, plastic flow, and pitting. The fatigue life of the ball-race
contact therefore depends on this conjunction having an adequate lubricant film, as pointed out in the previous section.

For many years the opinion prevailed that the maximum contact pressure in the ball-race contact precluded the possibility of a lubricant film existing in the conjunction. However, it is now generally accepted not only that a lubricant film is present, but also that the nature of the lubricant film has an important influence on the fatigue life of the bearing. Besides providing a film, the lubricant in a ball bearing must provide corrosion protection and act as a coolant.

Not only the ball-race contact but also all the interfaces between moving elements must be properly lubricated. The ball-separator and race-separator contacts experience mostly impact loading and therefore have greater possibilities of metal-to-metal contact, even when the bearing has an adequate supply of lubricant. For this reason the separator surfaces are generally coated with a low-friction material.

The ball-race contacts in ball bearings can generally be satisfactorily lubricated with a small amount of appropriate lubricant supplied to the right area within the bearing. The major considerations in proper ball bearing lubrication are

(1) Selection of a suitable lubricant

(2) Selection of a system that will provide an adequate and constant flow of this lubricant to the contact

These two topics are considered in the following sections.
3.6.1 Lubricants

Both oils and greases are extensively used as lubricants for all types of ball bearings over a wide range of speeds and operating temperatures. The choice is frequently determined by considerations other than lubrication requirements alone.

Because of its fluidity oil has a number of advantages over grease: It can enter the loaded conjunction most readily to flush away contaminants, such as water and dirt, and particularly to transfer heat from heavily loaded bearings. It is also frequently advantageous to lubricate bearings from a central oil system used for other machine parts.

Grease, however, is extensively used because it permits simplified designs of housings and bearing enclosures, which require less maintenance, and because it is more effective in sealing against dirt and contaminants. It also reduces possible damage to the process or product from oil leakage.

Oil Lubrication

Except for a few special requirements petroleum oils satisfy most operating conditions. High-quality products, free from adulterants that can have an abrasive or lapping action, are recommended. Animal or vegetable oils or petroleum oils of poor quality tend to oxidize, to develop acids, and to form
sludge or resinlike deposits on the bearing surfaces. They thus penalize bearing performance or endurance.

A composite of recommended lubricant viscosities at 38° C (100° F) is shown as Figure 3.17. In many ball bearing applications an oil equivalent to an SAE-10 motor oil (40x10^{-6} \text{ m}^2/\text{s}, or 40 \text{ cS}, at 38° C (100° F)) or a light turbine oil is the most frequent choice.

For a number of military applications where the operational requirements span the temperature range -54° to 204° C (-65° to 400° F), synthetic oils are used. Ester lubricants are most frequently employed in this temperature range. In applications where temperatures exceed 260° C (500° F), most synthetics will quickly break down, and either a solid lubricant (e.g., MoS_2) or a polyphenyl ether is recommended. A more detailed discussion of synthetic lubricants can be found in Bisson and Anderson (1964).

Grease Lubrication

The simplest method of lubricating a bearing is to apply grease, because of its relatively nonfluid characteristics. Danger of leakage is reduced, and the housing and enclosure design can be simpler and less costly than those used with oil. Grease can be packed into bearings and retained with inexpensive
closures, but packing should not be excessive and the manufacturer's recommendations should be closely adhered to.

The major limitation of grease lubrication is that it is not particularly useful in high-speed applications. In general it is not employed for speed factors (d_bN, bore in millimeters times speed in revolutions per minute) over 200,000 although selected greases have been used successfully for higher speed factors with specially designed ball bearings.

Greases vary widely in properties, depending on the type and grade or consistency. For this reason few specific recommendations can be made. Greases used for most bearing operating conditions consist of petroleum, diester, polyester, or silicone oils thickened with sodium or lithium soaps or with more recently developed nonsoap thickeners. General characteristics of greases are as follows:

(1) Petroleum oil greases are best for general-purpose operation from -34° to 149° C (-30° to 300° F).

(2) Diester oil greases are designed for low-temperature service down to -54° C (-65° F).

(3) Ester-based greases are similar to diester oil greases but have better high-temperature characteristics, covering a range from -73° to 177° C (-100° to 350° F).

(4) Silicone oil greases are used for both high- and low-temperature operation, over the widest temperature range of all
greases (−73° to 232° C; −100° to 450° F), but have the disadvantage of low load-carrying capacity.

(5) Fluorosilicone oil greases have all the desirable features of silicone oil greases plus good load-carrying capacity and resistance to fuels, solvents, and corrosive substances. They have a very low volatility in vacuums down to 10⁻⁷ torr, which makes them useful in aerospace applications.

(6) Perfluorinated oil greases have a high degree of chemical inertness and are completely nonflammable. They have good load-carrying capacity and can operate at temperatures as high as 288° C (550° F) for long periods, which makes them useful in the chemical process and aerospace industries, where high reliability justifies the additional cost.

Grease consistency is important since grease will slump badly and churn excessively when too soft and fail to lubricate when too hard. Either condition causes improper lubrication, excessive temperature rise, and poor performance and can shorten bearing life.

A valuable guide to the estimation of the useful life of grease in rolling-element bearings has been published by the Engineering Sciences Data Unit (1978).

It has been demonstrated recently by Aihara and Dowson (1979) and by Wilson (1979) that the film thickness in grease-lubricated components can be calculated with adequate accuracy by using the viscosity of the base oil in the elastohydrodynamic
equations (see Chapter 8). Aihara and Dowson compared film thickness measurements made by capacitance techniques on a grease-lubricated, two-disc machine with the predictions of elastohydrodynamic theory. Wilson reported an extensive and impressive range of experiments on a grease-lubricated roller bearing. This work enables the elastohydrodynamic theory developed in this text to be applied with confidence to grease-lubricated ball bearings.

3.6.2 Lubrication Systems

The quantity of lubricant required to maintain adequate lubrication of ball bearings is small. Data presented by Wilcock and Booser (1957) show that for medium-size, deep-groove ball bearings operating at moderate loads and speeds ($2.16 \times 10^6 \text{dbN}$), the quantity of oil required is about $0.5 \text{ mg/hr}$. The oil requirement is determined by the severity of the operating conditions.

Some of the techniques most frequently used to lubricate a ball bearing are described in the following paragraphs.

**Forced Lubrication**

Although the quantity of oil required to provide adequate lubrication is small, it is frequently desirable in heavily
loaded, high-speed bearings to use the oil to transfer heat away from the bearing. In such cases a circulating lubrication system is employed in which a pump delivers the lubricant oil to jets directed into the bearing and suitable drains return it to the reservoir. Gravity discharge can be used in systems where the pressure drop across the oil return lines is small. Care must be taken to ensure that no oil accumulates directly within the bearing area since submergence of the bearing in oil will cause excessive churning and result in high temperatures within the lubricating system. To avoid oil accumulation within the bearing cavities, scavenge pumps are frequently used. The scavenge pump provides the pressure differential necessary to evacuate the oil from the bearing cavity.

Mist Lubrication

At extremely high speeds, where any appreciable quantity of oil present in the bearings would cause an intolerable power loss due to oil churning, oil-mist lubrication is necessary. Small-bore bearings that operate at speeds of 50,000 to about 100,000 rpm are usually lubricated by means of an oil mist. In oil-mist systems the stream of atomized oil is generally directed against the inner race of the bearing, which is the most difficult to lubricate because of the effect of centrifugal force on the oil. Air pressures of 100 to 200 N/m² and nozzle
diameters of $0.8 \times 10^{-3}$ to $1.3 \times 10^{-3}$ m are used. Because the compressed air must be free from moisture, it is customary to install a drier in the air line.

Oil-mist systems are noncirculating: The oil is passed through the bearing once and then discarded. Extremely low oil flow rates are required for lubrication, exclusive of cooling. Since the required rates are so low, a nominal-capacity sump is adequate for supplying a multiple-bearing system for an extended time.

Splash Lubrication

Splash-feed lubrication systems are employed in low-speed machinery such as some gearboxes. In these systems oil is permitted to accumulate within the bearing cavities and is delivered directly to the bearing either by the immersion of the bearing separator and balls in the oil or through the use of an oil ring. Oil rings are mainly used in horizontally mounted machines. The loose-fitting ring, whose diameter is considerably larger than that of the shaft, dips into the lubricant reservoir beneath the bearing and carries oil to the top of the shaft by a viscous lifting process.
Wick Lubrication

Wick-feed systems are used either in low-speed machinery or in gyro and momentum wheel applications. In both cases the bearing cavity is packed with either felt or a porous plastic capable of providing a lubricant reservoir. Both the felt and the porous plastic hold oil by capillary forces. A wick bridging the oil reservoir supplies the required quantity of lubricant to the bearing.

In some designs the reservoir is located above the bearing to make use of gravity flow in the wick. Care must be taken to adjust the wick so that a very light contact is made with the reliever, or the wick may be charred.

Separator Lubrication

Separator lubrication systems are used mainly in instrument bearings, where the bearings are not readily accessible for re-lubrication and the system is sensitive to torque perturbations. Cotton-base phenolic separators have been most frequently used (as pointed out in Section 2.5). The separators are impregnated with either mineral or synthetic oils. During operation the lubricant is transferred from the separator onto the ball and hence reaches the ball-race contact.
Separator lubrication systems offer the advantages of compactness and relatively low torque variation. Their disadvantages include difficulties encountered in lubricant distribution and retention that result in premature failure. Because separator-impregnated bearings also have low heat-removal capability, they severely limit the maximum permissible operating speeds.

Dry Lubrication

Dry lubrication is mainly employed for high-temperature or cryogenic operations where known liquid lubricants lose their liquidity and/or are subject to chemical changes such as excessive oxidation. The most common dry lubricants have a base of molybdenum disulfide (MoS₂), polytetrafluoroethylene (PTFE), or carbon powder. Materials combining all three dry lubricants are also available. The lubricant is applied to metallic separators as a coating or is incorporated into the cage material. Lubricant is transferred from the cage onto the ball-race contacts as in the case of separator-impregnated lubrication, by the ball sliding against the cage pocket wall and transferring the lubricant into the ball-race contact area (e.g., Brewe, et al., 1969).

Dry-lubricated bearings often have relatively high wear rates, which limit their useful life. Because of this limi-
tion such bearings are generally employed only in applications where conventional lubricants become incapable of meeting the harsh operating requirements.

3.7 Closure

In this chapter load-deflection relationships have been developed for any type of elliptical contact. The deformation within the contact is a function, among other things, of the ellipticity parameter and the elliptic integrals of the first and second kinds. Simplified expressions have been written in Section 3.2 that allow deformation to be calculated quickly, and generally with adequate accuracy.

The methods developed in Section 3.3 to calculate the distribution of load among the balls in a complete bearing can be used in most applications because rotational speeds are usually slow to moderate. Under these conditions the effects of ball centrifugal forces and gyroscopic moments are negligible. At high rotational speeds these body forces become significant and tend to alter contact angles and clearances. They can thus greatly affect the static load distribution. The effect of these parameters on high-speed-bearing load distribution has been discussed in Section 3.4.

Ball bearings can fail from a number of causes; but if all other hazards are avoided, they will eventually fail because of
material fatigue. The fatigue life of a bearing is normally defined in terms of its statistical ability to survive for a certain time. The predominant factor in determining the fatigue life of a ball bearing is the dynamic load factor. In Section 3.5 this factor has been discussed and combined with expressions for the static load capacity from previous sections to enable the fatigue life of a ball bearing to be evaluated. In that section material, processing, and lubrication factors were also introduced in an adjusted fatigue life expression that greatly improves prediction of ball bearing performance.

Ball bearings can be satisfactorily lubricated with a small amount of an appropriate lubricant supplied to the right area within the bearing. The selection of an appropriate lubricant and the lubrication systems that provide a constant flow of lubricant to the contact have been discussed in Section 3.6.
### SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, A*, B*, C*, D*, L*, M*</td>
<td>Relaxation coefficients</td>
</tr>
<tr>
<td>A_v</td>
<td>Drag area of ball, m^2</td>
</tr>
<tr>
<td>a</td>
<td>Semimajor axis of contact ellipse, m</td>
</tr>
<tr>
<td>( a/2m )</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Total conformity of bearing</td>
</tr>
<tr>
<td>b</td>
<td>Semiminor axis of contact ellipse, m</td>
</tr>
<tr>
<td>( b/2m )</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Dynamic load capacity, N</td>
</tr>
<tr>
<td>( C_V )</td>
<td>Drag coefficient</td>
</tr>
<tr>
<td>( C_1, ..., C_8 )</td>
<td>Constants</td>
</tr>
<tr>
<td>c</td>
<td>19,609 N/cm^2 (28,440 lbf/in^2)</td>
</tr>
<tr>
<td>( c )</td>
<td>Number of equal divisions of semimajor axis</td>
</tr>
<tr>
<td>D</td>
<td>Distance between race curvature centers, m</td>
</tr>
<tr>
<td>( D )</td>
<td>Material factor</td>
</tr>
<tr>
<td>( D_e )</td>
<td>Defined by equation (5.63)</td>
</tr>
<tr>
<td>d</td>
<td>Ball diameter, m</td>
</tr>
<tr>
<td>( d )</td>
<td>Number of divisions in semiminor axis</td>
</tr>
<tr>
<td>( d_a )</td>
<td>Overall diameter of bearing (Figure 2.13), m</td>
</tr>
<tr>
<td>( d_b )</td>
<td>Bore diameter, m</td>
</tr>
<tr>
<td>( d_e )</td>
<td>Pitch diameter, m</td>
</tr>
<tr>
<td>( d_e' )</td>
<td>Pitch diameter after dynamic effects have acted on ball, m</td>
</tr>
<tr>
<td>( d_i )</td>
<td>Inner-race diameter, m</td>
</tr>
<tr>
<td>( d_o )</td>
<td>Outer-race diameter, m</td>
</tr>
</tbody>
</table>
modulus of elasticity, N/m$^2$

effective elastic modulus, $2 \sqrt{\left(\frac{1 - \nu_a^2}{E_a} + \frac{1 - \nu_b^2}{E_b}\right)}$, N/m$^2$

internal energy, m$^2$/s$^2$

processing factor

$E_1 = \left[\left(\frac{H_{\text{min}} - H_{\text{min}}}{H_{\text{min}}}\right) \times 100\right]$ elliptic integral of second kind with modulus $(1 - 1/k^2)^{1/2}$

approximate elliptic integral of second kind

dispersion exponent

normal applied load, N

normal applied load per unit length, N/m

lubrication factor

integrated normal applied load, N

centrifugal force, N

maximum normal applied load (at $\psi = 0$), N

applied radial load, N

applied thrust load, N

normal applied load at angle $\psi$, N

elliptic integral of first kind with modulus $(1 - 1/k^2)^{1/2}$

approximate elliptic integral of first kind

race conformity ratio

rms surface finish of ball, m

rms surface finish of race, m

dimensionless materials parameter, $aE$

fluid shear modulus, N/m$^2$

hardness factor

gravitational constant, m/s$^2$
\( g_E \)
dimensionless elasticity parameter, \( W^{8/3}/U^2 \)

\( g_V \)
dimensionless viscosity parameter, \( GW^3/U^2 \)

\( H \)
dimensionless film thickness, \( h/R_x \)

\( \bar{H} \)
dimensionless film thickness, \( H(W/U)^2 = F^2h/U^2n_0^2R_x^3 \)

\( H_c \)
dimensionless central film thickness, \( h_c/R_x \)

\( H_{c,s} \)
dimensionless central film thickness for starved lubrication condition

\( H_f \)
frictional heat, \( N \text{ m/s} \)

\( H_{\text{min}} \)
dimensionless minimum film thickness obtained from EHL elliptical-contact theory

\( H_{\text{min},r} \)
dimensionless minimum film thickness for a rectangular contact

\( H_{\text{min},s} \)
dimensionless minimum film thickness for starved lubrication condition

\( \bar{H}_c \)
dimensionless central film thickness obtained from least-squares fit of data

\( \bar{H}_{\text{min}} \)
dimensionless minimum film thickness obtained from least-squares fit of data

\( \overline{H}_c \)
dimensionless central-film-thickness – speed parameter, \( H_cU^{-0.5} \)

\( \overline{H}_{\text{min}} \)
dimensionless minimum-film-thickness – speed parameter, \( H_{\text{min}}U^{-0.5} \)

\( \overline{H}_0 \)
new estimate of constant in film thickness equation

\( h \)
film thickness, m

\( h_c \)
central film thickness, m

\( h_i \)
inlet film thickness, m
film thickness at point of maximum pressure, where
\[ \frac{dp}{dx} = 0, \text{ m} \]

minimum film thickness, m

constant, m

diametral interference, m

ball mass moment of inertia, m N s^2

integral defined by equation (3.76)

integral defined by equation (3.75)

function of k defined by equation (3.8)

mechanical equivalent of heat

polar moment of inertia, m N s^2

load-deflection constant

ellipticity parameter, a/b

approximate ellipticity parameter

thermal conductivity, N/s °C

lubricant thermal conductivity, N/s °C

fatigue life

adjusted fatigue life

reduced hydrodynamic lift, from equation (6.21)

lengths defined in Figure 3.11, m

fatigue life where 90 percent of bearing population will endure

fatigue life where 50 percent of bearing population will endure

bearing length, m

constant used to determine width of side-leakage region

moment, Nm

65
\( M_g \)  
gyroscopic moment, Nm

\( M_p \)  
dimensionless load-speed parameter, \( WU^{-0.75} \)

\( M_s \)  
torque required to produce spin, N m

\( m \)  
mass of ball, N s\(^2\)/m

\( m^* \)  
dimensionless inlet distance at boundary between fully flooded and starved conditions

\( \tilde{m} \)  
dimensionless inlet distance (Figures 7.1 and 9.1)

\( \bar{m} \)  
number of divisions of semimajor or semiminor axis

\( m_w \)  
dimensionless inlet distance boundary as obtained from Wedeven, et al. (1971)

\( N \)  
rotational speed, rpm

\( n \)  
number of balls

\( n^* \)  
refractive index

\( \bar{n} \)  
constant used to determine length of outlet region

\( P \)  
dimensionless pressure

\( P_D \)  
dimensionless pressure difference

\( P_d \)  
diametral clearance, m

\( P_e \)  
free endplay, m

\( P_{Hz} \)  
dimensionless Hertzian pressure, N/m\(^2\)

\( p \)  
pressure, N/m\(^2\)

\( p_{max} \)  
maximum pressure within contact, \( 3F/2\pi ab \), N/m\(^2\)

\( P_{iv,as} \)  
isoviscous asymptotic pressure, N/m\(^2\)

\( Q \)  
solution to homogeneous Reynolds equation

\( Q_m \)  
thermal loading parameter

\( \bar{Q} \)  
dimensionless mass flow rate per unit width, \( q_0\rho_0E'R^2 \)

\( q_f \)  
reduced pressure parameter

\( q_x \)  
volume flow rate per unit width in \( x \) direction, m\(^2\)/s
$q_y$  volume flow rate per unit width in $y$ direction, m$^2$/s

$R$  curvature sum, m

$R_a$  arithmetical mean deviation defined in equation (4.1), m

$R_c$  operational hardness of bearing material

$R_x$  effective radius in $x$ direction, m

$R_y$  effective radius in $y$ direction, m

$r$  race curvature radius, m

$r_{ax}, r_{bx}$  radii of curvature, m

$r_{ay}, r_{by}$  cylindrical polar coordinates

$r_c, \phi_c, z$  spherical polar coordinates

$r_s, \theta_s, \phi_s$  defined in Figure 5.4

$\overline{r}$  geometric separation, m

$S^*$  geometric separation for line contact, m

$S_0$  empirical constant

$s$  shoulder height, m

$T$  $\tau_0/P_{max}$

$\overline{T}$  tangential (traction) force, N

$T_m$  temperature, °C

$T^*_b$  ball surface temperature, °C

$T^*_f$  average lubricant temperature, °C

$\Delta T^*$  ball surface temperature rise, °C

$T_1$  $(\tau_0/P_{max})_{k=1}$

$T_v$  viscous drag force, N

$t$  time, s

$t_a$  auxiliary parameter

$u_B$  velocity of ball-race contact, m/s
velocity of ball center, m/s

dimensionless speed parameter, $n_0U/E'x$

surface velocity in direction of motion, $(u_a + u_b)/2$, m/s

number of stress cycles per revolution

sliding velocity, $u_a - u_b$, m/s

surface velocity in transverse direction, m/s

dimensionless load parameter, $F/E'R^2$

surface velocity in direction of film, m/s

dimensionless coordinate, $x/R_x$

dimensionless coordinate, $y/R_x$

dimensionless grouping from equation (6.14)

external forces, N

canstant defined by equation (3.48)

viscosity pressure index, a dimensionless constant

coordinate system

pressure-viscosity coefficient of lubrication, m$^2$/N

radius ratio, $R_y/R_x$

contact angle, rad

free or initial contact angle, rad

iterated value of contact angle, rad

curvature difference

viscous dissipation, $N/m^2 \cdot s$

total strain rate, $s^{-1}$

elastic strain rate, $s^{-1}$

viscous strain rate, $s^{-1}$
\( \gamma_a \)  
flow angle, deg

\( \delta \)  
total elastic deformation, m

\( \delta^* \)  
lubricant viscosity temperature coefficient, \( \text{°C}^{-1} \)

\( \delta_D \)  
elastic deformation due to pressure difference, m

\( \delta_r \)  
radial displacement, m

\( \delta_t \)  
axial displacement, m

\( \delta_x \)  
displacement at some location \( x \), m

\( \overline{\delta} \)  
approximate elastic deformation, m

\( \bar{\delta} \)  
elastic deformation of rectangular area, m

\( \varepsilon \)  
coefficient of determination

\( \varepsilon_1 \)  
strain in axial direction

\( \varepsilon_2 \)  
strain in transverse direction

\( \zeta \)  
angle between ball rotational axis and bearing centerline (Figure 3.10)

\( \zeta_a \)  
probability of survival

\( n \)  
absolute viscosity at gauge pressure, \( \text{N s/m}^2 \)

\( \overline{n} \)  
dimensionless viscosity, \( n/n_0 \)

\( n_0 \)  
viscosity at atmospheric pressure, \( \text{N s/m}^2 \)

\( n_\infty \)  
6.31x10\(^{-5}\) \( \text{N s/m}^2 \) (0.0631 cP)

\( \theta \)  
angle used to define shoulder height

\( \Lambda \)  
film parameter (ratio of film thickness to composite surface roughness)

\( \lambda \)  
equals 1 for outer-race control and 0 for inner-race control

\( \lambda_a \)  
second coefficient of viscosity

\( \lambda_b \)  
Archard-Cowking side-leakage factor, \( (1 + 2/3 \alpha_a)^{-1} \)

\( \lambda_c \)  
relaxation factor
\( \mu \)  
coefficient of sliding friction

\( \mu^* \)  
\( \rho/\rho_0 \)

\( \nu \)
Poisson's ratio

\( \rho \)  
lubricant density, \( N s^2/m^4 \)

\( \bar{\rho} \)  
dimensionless density, \( \rho/\rho_0 \)

\( \rho_0 \)  
density at atmospheric pressure, \( N s^2/m^4 \)

\( \sigma \)  
normal stress, \( N/m^2 \)

\( \sigma_1 \)  
stress in axial direction, \( N/m^2 \)

\( \tau \)  
shear stress, \( N/m^2 \)

\( \tau_0 \)  
maximum subsurface shear stress, \( N/m^2 \)

\( \bar{\tau} \)  
shear stress, \( N/m^2 \)

\( \bar{\tau}_e \)  
equivalent stress, \( N/m^2 \)

\( \bar{\tau}_L \)  
limiting shear stress, \( N/m^2 \)

\( \phi \)  
ratio of depth of maximum shear stress to semiminor axis of contact ellipse

\( \phi^* \)  
\( \rho_0 h^{3/2} \)

\( \phi_1 \)  
(\( \phi \))_{k=1}

\( \phi \)  
 auxiliary angle

\( \phi_T \)  
thermal reduction factor

\( \psi \)  
angular location

\( \psi_L \)  
limiting value of \( \psi \)

\( \Omega_i \)  
absolute angular velocity of inner race, rad/s

\( \Omega_0 \)  
absolute angular velocity of outer race, rad/s

\( \omega \)  
angular velocity, rad/s

\( \omega_B \)  
angular velocity of ball-race contact, rad/s

\( \omega_b \)  
angular velocity of ball about its own center, rad/s
\( \omega_c \)  
angular velocity of ball around shaft center, rad/s

\( \omega_s \)  
ball spin rotational velocity, rad/s

Subscripts:

- \( a \) solid a
- \( b \) solid b
- \( c \) central
- \( bc \) ball center
- \( IE \) isoviscous-elastic regime
- \( IR \) isoviscous-rigid regime
- \( i \) inner race
- \( K \) Kapitza
- \( min \) minimum
- \( n \) iteration
- \( o \) outer race
- \( PVE \) piezoviscous-elastic regime
- \( PVR \) piezoviscous-rigid regime
- \( r \) for rectangular area
- \( s \) for starved conditions
- \( x,y,z \) coordinate system

Superscript:

- \((-)\) approximate
REFERENCES


Agricola, G. (1556) De Re Metallica, Basel.


Clark, R. H. (1938) "Earliest Known Ball Thrust Bearing Used in Windmill," English Mechanic, 30 (Dec.) 223.


Goodman, J. (1912) "(1) Roller and Ball Bearings; 


Jacobson, B. (1972) "Elasto-Solidifying Lubrication of Spherical Surfaces."
American Society of Mechanical Engineers Paper No. 72-LUB-7.


Parker, R. J. and Kannel, J. W. (1971) "Elastohydrodynamic Film Thickness Between Rolling Disks with a Synthetic Paraffinic Oil to 589 K (600° F); NASA TN D-6411.

Parker, R. J. and Zaretsky, E. V. (1978) "Rolling-Element Fatigue Life of AISI M-50 and 18-4-1 Balls." NASA TP-1202.


Rowe, J. (1734) "All Sorts of Wheel-Carriage Improved," printed for Alexander Lyon under Tom's Coffee House in Russell Street, Covent Garden, London.


Varlo, C. (1772) "Reflections Upon Friction with a Plan of the New Machine for Taking It Off in Wheel-Carriages, Windlases of Ships, etc., Together with Metal Proper for the Machine, the Full Directions for Making It."


Table 3.1 Comparison of Approximate and Exact Formulas

<table>
<thead>
<tr>
<th>Radius Ratio, $R_y/R_x$</th>
<th>Ellipticity</th>
<th>Elliptic Integrals</th>
<th>Deformation at Center of Contact</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k$</td>
<td>$k$</td>
<td>Percent Error</td>
</tr>
<tr>
<td></td>
<td>$\delta$</td>
<td>$\delta$</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>1.000</td>
<td>1.00</td>
<td>1.03</td>
<td>3.00</td>
</tr>
<tr>
<td>2.820</td>
<td>1.99</td>
<td>2.00</td>
<td>.50</td>
</tr>
<tr>
<td>5.314</td>
<td>3.01</td>
<td>3.00</td>
<td>-.33</td>
</tr>
<tr>
<td>8.330</td>
<td>4.01</td>
<td>4.00</td>
<td>-.25</td>
</tr>
<tr>
<td>11.805</td>
<td>4.99</td>
<td>5.00</td>
<td>.20</td>
</tr>
<tr>
<td>15.697</td>
<td>5.97</td>
<td>6.00</td>
<td>.50</td>
</tr>
<tr>
<td>19.971</td>
<td>6.92</td>
<td>7.00</td>
<td>1.16</td>
</tr>
<tr>
<td>24.605</td>
<td>7.87</td>
<td>8.00</td>
<td>1.65</td>
</tr>
<tr>
<td>29.576</td>
<td>8.80</td>
<td>9.00</td>
<td>2.27</td>
</tr>
<tr>
<td>34.869</td>
<td>9.72</td>
<td>10.00</td>
<td>2.88</td>
</tr>
</tbody>
</table>

Table 3.2 Effect of Degree of Conformity on Contact Parameters

[For all three cases, $E' = 2.28 \times 10^{11}$ N/m² (steel on steel), $F = 4.45$ N, and $r_{ax} = r_{ay} = 6.35$ mm. For the ball - outer ring contact, $d_e = 65$ mm, $\beta = 0^\circ$, $f_o = 0.52$ (assume 209 radial ball bearing).]

<table>
<thead>
<tr>
<th>Contact Parameters</th>
<th>Ball on Ball</th>
<th>Ball on Plane</th>
<th>Ball - Outer Ring Contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{bx}$</td>
<td>6.35 mm</td>
<td></td>
<td>-38.9 mm</td>
</tr>
<tr>
<td>$r_{by}$</td>
<td>6.35 mm</td>
<td></td>
<td>-6.60 mm</td>
</tr>
<tr>
<td>$R$</td>
<td>1.59 mm</td>
<td>3.18 mm</td>
<td>7.26 mm</td>
</tr>
<tr>
<td>$R_y/R_x$</td>
<td>1</td>
<td>1</td>
<td>22.1</td>
</tr>
<tr>
<td>$k$</td>
<td>1.03</td>
<td>1.03</td>
<td>7.330</td>
</tr>
<tr>
<td>$\overline{\varepsilon}$</td>
<td>1.60</td>
<td>1.60</td>
<td>1.03</td>
</tr>
<tr>
<td>$\overline{\gamma}$</td>
<td>1.53</td>
<td>1.53</td>
<td>3.38</td>
</tr>
<tr>
<td>$\bar{a}$</td>
<td>0.0465 mm</td>
<td>0.0586 mm</td>
<td>0.247 mm</td>
</tr>
<tr>
<td>$\bar{b}$</td>
<td>0.0451 mm</td>
<td>0.0569 mm</td>
<td>0.0336 mm</td>
</tr>
<tr>
<td>$\tau_{ab}$</td>
<td>$0.659 \times 10^{-2}$ mm²</td>
<td>$1.04 \times 10^{-2}$ mm²</td>
<td>$2.60 \times 10^{-2}$ mm²</td>
</tr>
<tr>
<td>$\overline{\delta}$</td>
<td>$6.31 \times 10^{-4}$ mm</td>
<td>$4.87 \times 10^{-4}$ mm</td>
<td>$2.56 \times 10^{-4}$ mm</td>
</tr>
<tr>
<td>$P_{max}$</td>
<td>$1.01 \times 10^9$ N/m²</td>
<td>$0.657 \times 10^9$ N/m²</td>
<td>$0.256 \times 10^9$ N/m²</td>
</tr>
</tbody>
</table>
Table 3.3 Shear Stress Parameters as a Function of Inverse Ellipticity Parameter

<table>
<thead>
<tr>
<th>1/k</th>
<th>0</th>
<th>0.2050</th>
<th>0.3358</th>
<th>0.5020</th>
<th>0.7849</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.2500</td>
<td>0.2475</td>
<td>0.2436</td>
<td>0.2371</td>
<td>0.2241</td>
<td>0.2139</td>
</tr>
<tr>
<td>(\Phi)</td>
<td>0.5</td>
<td>0.4854</td>
<td>0.4651</td>
<td>0.4347</td>
<td>0.3842</td>
<td>0.3509</td>
</tr>
<tr>
<td>((T_1/T)^{3.1}(\Phi/\Phi_1)^{0.4})</td>
<td>0.7105</td>
<td>0.7243</td>
<td>0.7480</td>
<td>0.7918</td>
<td>0.8975</td>
<td>1</td>
</tr>
<tr>
<td>(0.706 + 0.3(1/k)^{1.92})</td>
<td>0.7060</td>
<td>0.7203</td>
<td>0.7429</td>
<td>0.7859</td>
<td>0.8944</td>
<td>1.006</td>
</tr>
<tr>
<td>Percent error</td>
<td>-0.63</td>
<td>-0.55</td>
<td>-0.68</td>
<td>-0.75</td>
<td>-0.36</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Table 3.5 Material Factor for Throughhardened Bearing Materials

<table>
<thead>
<tr>
<th>Material</th>
<th>Material Factor, D</th>
</tr>
</thead>
<tbody>
<tr>
<td>52100</td>
<td>2.0</td>
</tr>
<tr>
<td>M-1</td>
<td>.6</td>
</tr>
<tr>
<td>M-2</td>
<td>.6</td>
</tr>
<tr>
<td>M-10</td>
<td>2.0</td>
</tr>
<tr>
<td>M-50</td>
<td>2.0</td>
</tr>
<tr>
<td>T-1</td>
<td>.6</td>
</tr>
<tr>
<td>Halmo</td>
<td>2.0</td>
</tr>
<tr>
<td>M-42</td>
<td>.2</td>
</tr>
<tr>
<td>WB-49</td>
<td>.6</td>
</tr>
</tbody>
</table>
Figure 3.1 - Contact areas in a ball bearing.

Figure 3.2 - Pressure distribution in an ellipsoidal contact.
Figure 3.3 - Ellipticity parameter and elliptic integrals of first and second kinds as a function of curvature ratio.

Figure 3.4 - Three degrees of conformity.

(a) Ball on ball.

(b) Ball on plane.

(c) Ball - outer-ring contact.
Figure 3.5 - Radially loaded ball bearing.

Figure 3.6 - Contact ellipse in bearing race.
Figure 3.7 - Angular-contact ball bearing under thrust load.

Figure 3.8 - Internal fit resulting from combined loading of ball bearing.

Figure 3.9 - Ball forces and moments at high-speed conditions.
Figure 3.10 - Ball motion vectors.

Figure 3.11 - Race-curvature center deflection.
Figure 3.12 - Ball bearing fatigue life distribution.

Figure 3.13 - Typical Weibull plot of bearing fatigue failures.
Suggested range of operation

Figure 3.14 - Bearing fatigue life as a function of race curvature. From Winn, et al. (1974). Ball diameter, \(d_c\), 4.76 mm (0.1875 in.); number of balls, \(n\), 10; pitch diameter, \(d_p\), 28.5 mm (1.122 in.); initial contact angle, \(\beta_i\), 20°; radial load, \(F_r\), 0; thrust load, \(F_t\), 445 N (100 lb).

Figure 3.15 - Typical variation of fatigue life as a function of initial contact angle.

Figure 3.16 - Lubrication factor as a function of film parameter.
Figure 3.17 - Recommended lubricant viscosities for ball bearings.
Load-deflection relationships for different types of elliptical contacts, such as those found in a ball bearing are developed. Simplified expressions that allow quick calculations of deformation to be made simply from a knowledge of the applied load, the material properties, and the geometry of the contacting elements are presented. Ball bearings subjected to radial, thrust, and combined ball loads are analyzed. A design criterion for fatigue life of ball bearings is developed. The selection of a satisfactory lubricant, as well as describing systems that provide a constant flow of lubricant to the contact, is considered.