Introduction to Ball Bearings

Bernard J. Hamrock
Lewis Research Center
Cleveland, Ohio

and

Duncan Dowson
The University of Leeds
Leeds, England

February 1981
The purpose of a ball bearing is to provide a relative positioning and rotational freedom while transmitting a load between two structures, usually a shaft and a housing. For high rotational speeds—as found, for example, in gyroscope ball bearings—the purpose can be expanded to include rotational freedom with practically no wear in the bearing. This condition can be achieved by separating the bearing parts with a coherent film of fluid known as an elastohydrodynamic film. Denhard (1966) pointed out that the elastohydrodynamic film can be maintained not only when the bearing carries the load on a shaft, but also when the bearing is preloaded to position the shaft to within micro- or nano-inch accuracy and stability. This chapter provides background information on ball bearings; elastohydrodynamic lubrication theory is discussed in Chapter 7.

Ball bearings are used in many kinds of machines and devices with rotating parts, as pointed out in the preceding chapter. The designer is often confronted with decisions on whether a ball or fluid-film bearing should be used in a particular application. The following characteristics make ball bearings more desirable than fluid-film bearings in many situations:
(1) Low starting and good operating friction
(2) The ability to support combined radial and thrust loads
(3) Less sensitivity to interruptions in lubrication
(4) No self-excited instabilities
(5) Good low-temperature starting

Within reasonable limits, changes in load, speed, and operating temperature have but little effect on the satisfactory performance of ball bearings.

The following characteristics make ball bearings less desirable than fluid-film bearings:

(1) Finite fatigue life subject to wide fluctuations
(2) Larger space required in the radial direction
(3) Low damping capacity
(4) Higher noise level
(5) More severe alignment requirements
(6) Higher cost

In the light of these characteristics, piston engines normally use fluid-film bearings, whereas jet engines use ball bearings almost exclusively. Each type of bearing has its particular strong points, and care should be taken in choosing the most appropriate type of bearing for a given application. Useful guidance on the important issue of bearing selection has been presented by the Engineering Sciences Data Unit (ESDU 1965, 1967).

The Engineering Sciences Data Unit documents (1965, 1967) provide an excellent guide to the selection of the type of journal or thrust bearing that is most likely to give the required performance.
when considering the load, speed, and geometry of the bearing. The types of bearing considered were

(1) Rubbing bearings, where the two bearing surfaces rub together (e.g., unlubricated bushings made from materials based on nylon, polytetrafluoroethylene (PTFE), and carbon)

(2) Oil-impregnated porous metal bearings, where a porous metal bushing is impregnated with lubricant and thus gives a self-lubricating effect (as in sintered-iron and sintered-bronze bearings)

(3) Rolling bearings, where relative motion is facilitated by interposing rolling elements between stationary and moving components (as in ball, roller, and needle bearings)

(4) Hydrodynamic film bearings, where the surfaces in relative motion are kept apart by pressures generated hydrodynamically in the lubricant film

Figure 2.1, reproduced from the Engineering Sciences Data Unit publication (1965), gives a guide to the typical load that can be carried at various speeds, for a nominal life of 10,000 hours at room temperature, by journal bearings of various types on shafts of the diameters quoted. The heavy curves indicate the preferred type of journal bearing for a particular load, speed, and diameter and thus divide the graph into distinctive regions. The applied load and speed are usually known, and this enables a preliminary assessment to be made of the type of journal bearing most likely to be suitable for a particular application.
In many cases the shaft diameter will already have been determined by other considerations, and Figure 2.1 can be used to find the type of journal bearing that will give adequate load capacity at the required speed.

These curves are based on good engineering practice and commercially available parts. Higher loads and speeds or smaller shaft diameters are possible with exceptionally high engineering standards or specially produced materials. Except for rolling bearings the curves are drawn for bearings with a width equal to the diameter. A medium mineral oil lubricant is assumed for the hydrodynamic bearings.

Similarly Figure 2.2 reproduced from the Engineering Sciences Data Unit publication (1967), gives a guide to the typical maximum load that can be carried at various speeds, for a nominal life of 10,000 hours at room temperature, by thrust bearings of various types on shafts of the diameters quoted. The heavy curves again indicate the preferred type of bearing for a particular load, speed, and diameter and thus divide the graph into major regions.

The remainder of this chapter is concerned with the definition of different types of ball bearings and their geometry and kinematics. Bearing materials and manufacturing processes are also discussed in this chapter, as well as separators (sometimes called cages or retainers). The factor that remains constant in these important introductory sections is that it is assumed, for the purposes of analysis, that the bearing carries no load. The consideration of loaded bearings starts in Chapter 3, but the
material covered in the present chapter is vital in laying the foundations for the remainder of the book.

2.1 Types of Ball Bearings

The essential parts of a ball bearing - the inner and outer ring, the balls, and the separator - are shown in Figure 2.3. The inner ring is mounted on a shaft and has a groove in which the balls ride. The outer ring is usually the stationary part of the bearing and also contains a groove to guide and support the balls. The separator prevents contact between the balls and thus reduces friction, wear, and noise from the regions where severe sliding conditions would occur. In a few applications where operating conditions are mild, the rings and separator can be omitted and loose balls interposed between the shaft and housing. This type of bearing is sometimes found in bicycles. The vast majority of ball bearings, however, are preassembled units consisting of four elements. There are many types of bearings because of variations in the design of rings and separators and in the number of balls. They can be divided into classes according to their function: those that support a radial load, those that support a thrust load, or those that support a combination of thrust and radial loads. The last type is termed "angular-contact bearings."
2.1.1 Radial Bearings

Deep Groove

In deep-groove ball bearings, sometimes called Conrad bearings after their German designer, the races are approximately one-fourth as deep as the ball diameter. A cross section of the ball and outer ring of a deep-groove ball bearing is shown in Figure 2.4. Although deep-groove ball bearings are designed to carry a radial load, they perform well under a combined radial and thrust load. For this reason this is the most widely used type of ball bearing.

The assembly of a deep-groove bearing is shown in a four-stage illustration in Figure 2.5. The inner ring is displaced within the outer ring, leaving a crescent-shaped space (Figure 2.5(a)). Balls are inserted into the resultant wide crescent opening (Figure 2.5(b)). By this method only slightly more than half the annular space between the inner and outer pathway can be filled with balls. The inner ring is then moved back to be concentric with the outer ring, and the balls are spaced equally around the bearing (Figure 2.5(c)). A separator is then installed (Figure 2.5(d)) to maintain equal spacing of the balls.
Filled Notch

The most effective way of increasing bearing load capacity is to increase the ball complement. In a filled-notch ball bearing, as shown in Figure 2.6, this is accomplished by cutting a notch in the races of a conventional bearing, thereby permitting the insertion of the balls. The notch need not be as deep as the grooves; or it may be cut only in the outer race, in which case it will be deeper than the outer groove. Once the balls have been inserted into the bearing through the notch, the notch is filled by an insert. The insert should be kept on the unloaded side of the bearing, for it weakens the race. The increased ball complement makes possible a 20 to 40 percent increase in radial-load capacity over that of the normal deep-groove bearing, depending on bearing size. Thrust-load capacity is sacrificed, however, because the filled notch is cut almost to the bottom of the race groove. The bearing can sustain less than one-third of the thrust load of the deep-groove bearing before the ball-race contact ellipse begins to contact the edge of the filled notch.

Other Radial Bearings

There are a number of other forms of radial ball bearings for specific applications, such as instrument bearings, counterbored bearings, or aircraft control bearings. The details of the spe-
Specific design of these special-application bearings can be found in most bearing company catalogs.

Furthermore there are applications where it is advantageous to have a double- rather than a single-row ball bearing. A double-row bearing has greater load capacity than a single-row bearing. However, the applied load that can be carried by a double-row ball bearing is less than twice that of its single-row counterpart because of the practical difficulty of providing for both rows of balls to share the load equally. Proper sharing of the applied load between the rows is a function of the geometrical accuracy of the grooves. Otherwise these bearings behave like a single-row ball bearing.

2.1.2 Angular-Contact Bearings

Unidirectional Thrust

A unidirectional thrust bearing, which is the most popular form of angular-contact bearing, is shown in Figure 2.7. It is made with one shoulder of the outer ring counterbored almost to the bottom of the race groove. The outer ring is usually heated during assembly so that the inner ring and the ball-and-separator assembly can be snapped into place. This type of bearing is designed to support combined radial and unidirectional thrust loads. The amount of thrust load that this bearing can support is a function of the contact angle. A bearing with a large contact angle, like
that shown in Figure 2.7(b), can support heavier thrust loads than the smaller-contact-angle bearing shown in Figure 2.7(a). The contact angle does not usually exceed 40° since high contact angles result in considerable ball spinning and heat generation. A contact angle of zero corresponds to a radial bearing; a contact angle of 90° corresponds to a thrust bearing.

Self-Aligning Bearings

Self-aligning ball bearings normally have two rows of balls that roll in a common spherical race in the outer ring, as shown in Figure 2.8. Because of this design the inner ring, with the ball complement, can align itself freely around the axis of the shaft. The self-aligning feature of this bearing allows it to adjust to small degrees of misalignment between shaft and housing without malfunction. Even when the shaft bends under load, the bearing will follow the deflection of the shaft without resistance. Self-alignment also contributes to smooth running by neutralizing the effect of the balls wobbling in the grooves. This bearing is therefore particularly useful in applications in which it is difficult to obtain exact parallelism between the shaft and housing bores.

Two characteristics limit this bearing's usefulness: It cannot support moments; and because the balls do not conform well to the outer race, which is not grooved, it has reduced load-carrying capacity. This latter limitation is compensated for somewhat by
the use of a very large ball complement, which minimizes the load carried by each ball.

Duplex

Because most applications require thrust-load support in two directions, unidirectional-thrust angular-contact bearings are usually supplied in matched pairs for duplex mounting. Back-to-back and face-to-face mounted bearings make it possible to carry thrust load in either direction. Such bearings are manufactured as matched pairs such that, when they are mated and the races made flush, each bearing is slightly preloaded. The purpose of a pre-load is to keep the bearing loaded at all times, even when the loads are low and the speeds are high. This preloading of the components against each other stiffens the assembly in the axial direction and helps to prevent ball skidding with acceleration at light loads.

Duplex back-to-back mounts (Figure 2.9(a)) and face-to-face mounts (Figure 2.9(b)) are identical in all their characteristics except resistance to moments. For the back-to-back arrangement the lines of contact intersect outside the bearing envelope. This type of mount is therefore more resistant to moments and shaft bending.

Tandem-mounted angular-contact bearings (Figure 2.9(c)) are used in applications requiring very high, unidirectional-thrust-load capacity. With careful manufacture and installation a tandem
bearing pair can have a thrust capacity about 80 percent greater than the capacity of a single bearing. Duplex bearings are supplied in matched pairs and they should never be separated or used with single bearings from other pairs.

Split Inner Ring

Aircraft turbine engines operating in the speed range of greater than 2 million \(d_bN\) (bearing bore in millimeters times shaft speed in revolutions per minute) has brought about the development of the split-inner-ring bearing shown in Figure 2.10. The inner-ring race is normally ground with a shim between the inner-ring halves. The shim is removed before assembly, so that the resulting inner-ring groove is shaped like a gothic arch. This provides a much lower ratio of axial-to-radial play than when both race grooves are circular. This is important where thermal gradients cause an appreciable loss in initial radial play as the bearing comes up to temperature. Sufficient radial play must be built into the bearing without allowing excessive axial play. Excessive axial play results in lower accuracy of shaft location.

A disadvantage of the split-ring bearing is that it must operate at fairly high ratios of thrust to radial load. Any appreciable radial load causes simultaneous ball contact on both halves of the inner ring, and this results in severe ball skidding and bearing distress.
2.1.3 Thrust Bearings

The simplest form of ball thrust bearing is shown in Figure 2.11. In this type of bearing a single row of balls, set in a separator, runs in two similar grooves formed in the stationary and revolving rings. The revolving ring is fixed to a shaft. These grooves are usually shallower than the groove in a deep-groove radial ball bearing. Thrust ball bearings are not suited to carry any appreciable radial load. However, radial loading of each ball arises from its own centrifugal force when it is revolving in its track. Therefore at high speeds these forces can be considerable relative to the thrust load, and this may restrict the speed of operation. Such bearings are also speed limited because of spinning in the ball-race contact.

Because they cannot support appreciable radial loads, thrust ball bearings must be used in conjunction with radial bearings. Even a slight misalignment can cause drastic load concentrations, so either a high degree of squareness must be maintained between the stationary race seat and the shaft axis or a self-aligning seat must be used. Furthermore it is important in this type of bearing that all balls be very accurately manufactured, both in roundness and diameter, since any ball larger than the others would carry a disproportionate fraction of the total load during the whole revolution.
2.2 Geometry of Ball Bearings

The operating characteristics of a ball bearing depend greatly on the diametral clearance of the bearing. This clearance varies for the different types of bearings discussed in Section 2.1. In the present section the principal geometrical relationships governing the operation of unloaded ball bearings are developed. This information will be of vital interest when such quantities as stresses, deflections, load capacity, and life are considered in subsequent chapters. Bearings rarely operate in the unloaded state, but understanding of this section is vital to the appreciation of the remaining chapters.

2.2.1 Conformal and Nonconformal Surfaces

Before the various geometrical relationships for ball bearings are developed, it is desirable to define conformal and nonconformal surfaces. Hydrodynamic lubrication is generally characterized by surfaces that are conformal. That is, the surfaces fit snugly into each other with a high degree of geometrical conformity, so that the load is carried over a relatively large area. Furthermore the load-carrying surface area remains essentially constant while the load is increased. Fluid-film journal and slider bearings exhibit conformal surfaces. In journal bearings the radial clearance between the shaft and bearing is typically one-thousandth of the
the inclination of the bearing surface to the runner is typically one part in a thousand.

Many machine elements have contacting surfaces that do not conform to each other very well. The full burden of the load must then be carried by a very small contact area. In general the contact areas between nonconformal surfaces enlarge considerably with increasing load but are still small compared with the contact areas between conformal surfaces. Some examples of these nonconformal surfaces are mating gear teeth, cams and followers, and rolling-element bearings.

The load per unit area in conformal bearings is relatively low, typically only 1 MN/m² and seldom over 7 MN/m². By contrast, the load per unit area in nonconformal contacts, such as those that exist in ball bearings, will generally exceed 700 MN/m², even at modest applied loads. These high pressures result in elastic deformation of the bearing materials such that the elliptical contact areas are formed for oil film generation and load support. The significance of the high contact pressures is that they result in a considerable increase in fluid viscosity. Inasmuch as viscosity is a measure of a fluid’s resistance to flow, this increase greatly enhances the lubricant’s ability to support load without being squeezed out of the contact zone.

The nonconformal surfaces of a ball bearing are shown in Figure 2.12. The ball and race conform to some degree in the section shown in Figure 2.12(a), but the sectional view shown in
Figure 2.12(b) clearly exhibits little conformity. This book is concerned only with nonconformal contacts.

2.2.2 Pitch Diameter, Clearance, and Race Conformity

The cross section through a radial, single-row ball bearing shown in Figure 2.13 depicts the radial clearance and various diameters. The pitch diameter $d_e$ is the mean of the inner- and outer-ring race contact diameters and is given by

$$d_e = d_i + \frac{1}{2} (d_o - d_i)$$

or

$$d_e = \frac{1}{2} (d_o + d_i)$$

Also, from Figure 2.13, the diametral clearance denoted by $P_d$ can be written as

$$P_d = d_o - d_i - 2d$$

Diametral clearance may therefore be thought of as the maximum distance that one race can move diametrically with respect to the other when no measurable force is applied and both races lie in the same plane. Although diametral clearance is generally used in connection with single-row radial bearings, equation (2.2) is also applicable to angular-contact bearings.

Race conformity is a measure of the geometrical conformity of the race and the ball in a plane passing through the bearing axis, which is a line passing through the center of the bearing
and perpendicular to its plane, and transverse to the race. Figure 2.14 is a cross section of a ball bearing showing race conformity, expressed as

\[ f = \frac{r}{d} \]  

(2.3)

For perfect conformity, where the radius of the race or groove is equal to the ball radius, \( f \) is equal to 1/2. The closer the race conforms to the ball, the greater the frictional heat within the contact. On the other hand, open race curvature and reduced geometrical conformity, which reduce friction, also increase the maximum contact stresses and consequently reduce the bearing fatigue life. For this reason most ball bearings made today have race conformity ratios in the range 0.51 \( \leq f \leq 0.54 \), with \( f = 0.52 \) being the most common value. The race conformity ratio for the outer race is usually made slightly larger than that for the inner race to compensate for the closer conformity in the plane of the bearing between the outer race and ball than between the inner race and ball. This tends to equalize the contact stresses at the inner and outer race contacts. The difference in race conformity ratio does not normally exceed 0.02.

2.2.3 Contact Angle, Endplay, and Shoulder Height

Contact Angle

Radial bearings have some axial play since they are generally designed to have a diametral clearance, as shown in Figure 2.15(a).
This implies a free-contact angle different from zero. Angular-contact bearings are specifically designed to operate under thrust loads. The clearance built into the unloaded bearing, along with the race conformity ratio, determines the bearing free-contact angle.

Figure 2.15(b) shows a radial ball bearing with contact due to the axial shift of the inner and outer rings when no measurable force is applied.

However, before the free contact is discussed, it is important to define the distance between the centers of curvature of the two races in line with the center of the ball in both Figures 2.15(a) and (b). This distance — denoted by \( x \) in Figure 2.15(a) and by \( D \) in Figure 2.15(b) — depends on race radii and ball diameter. Denoting quantities referred to the inner and outer races by subscripts \( i \) and \( o \), respectively, we see from Figures 2.15(a) and (b) that

\[
\frac{P_d}{4} + d + \frac{P_d}{4} = r_o - x + r_i
\]

or

\[
x = r_o + r_i - d - \frac{P_d}{2}
\]  \(2.4\)

and

\[d = r_o - D + r_i\]

or

\[D = r_o + r_i - d\]  \(2.5\)

From these equations we can write that

\[x = D - \frac{P_d}{2}\]  \(2.6\)

17
This distance, shown in Figure 2.15(b), will be useful in defining the contact angle.

By using equation (2.3), we can write equation (2.5) as

\[ D = B d \] (2.7)

where

\[ B = f_0 + f_1 - 1 \] (2.8)

The quantity \( B \) in equation (2.8) is known as the total conformity ratio and is a measure of the combined conformity of both the outer and inner races to the ball. It will be seen that calculations of bearing deflection in later chapters depend on the quantity \( B \).

The free-contact angle \( \beta_f \) (Figure 2.15(b)) is defined as the angle made by a line through the points of contact of the ball and both races with a plane perpendicular to the bearing axis of rotation when no measurable force is applied. Note that the centers of curvature of both the outer and inner races lie on the line defining the free-contact angle. From Figure 2.15(b) the expression for the free-contact angle can be written as

\[
\cos \beta_f = \frac{D - P_d/2}{D}
\]

or

\[
\beta_f = \cos^{-1} \left( 1 - \frac{P_d}{2D} \right)
\] (2.9)

By using equations (2.2) and (2.5), we can write equation (2.9) as
Equation (2.10) shows that if the size of the balls is increased and everything else remains constant, the free-contact angle is decreased. Similarly, if the ball size is decreased, the free-contact angle is increased.

From equation (2.9) the diametral clearance $P_d$ can be written as

$$P_d = 2D(1 - \cos \beta_f) \quad (2.11)$$

This is an alternative way of defining the diametral clearance from that given in equation (2.2).

Endplay

Free endplay $P_e$ is the maximum axial movement of the inner race with respect to the outer, when both races are coaxially centered and no measurable force is applied. Free endplay depends on total curvature and contact angle, as shown in Figure 2.15(b), and can be written as

$$P_e = 2D \sin \beta_f \quad (2.12)$$

The variation of free-contact angle and endplay with the ratio $P_d/2d$ is shown in Figure 2.16 for four values of total conformity normally found in single-row ball bearings. Eliminating $\beta_f$ in equations (2.11) and (2.12) enables the following relationships

$$\beta_f = \cos^{-1} \left[ \frac{r_o + r_1 - \frac{1}{2} (d_o - d_i)}{r_o + r_1 - d} \right] \quad (2.10)$$
between free endplay and diametral clearance to be established:

\[ P_d = 2D - \left[ \left(2D\right)^2 - D^2 \right]^{1/2} \]  \hspace{1cm} (2.13)

\[ P_e = \left(4DP_d - D^2\right)^{1/2} \]  \hspace{1cm} (2.14)

Shoulder Height

The shoulder height found in ball bearings is illustrated in Figure 2.17. Shoulder height, or race depth, is the depth of the race groove measured from the shoulder to the bottom of the groove and is denoted by \( s \) in Figure 2.17. From this figure the equation defining the shoulder height can be written as

\[ s = r(1 - \cos \theta) \]  \hspace{1cm} (2.15)

The maximum possible diametral clearance for complete retention of the ball-race contact within the race under zero thrust load is given by

\[ (P_d)_{\text{max}} = \frac{2D}{r} \]  \hspace{1cm} (2.16)

In Chapter 3 the shoulder height requirement under a thrust load is discussed.
2.2.4 Curvature Sum and Difference

The undeformed geometry of contacting solids in a ball bearing can be represented by two ellipsoids. The two solids with different radii of curvature in a pair of principal planes \((x\) and \(y)\) passing through the contact between the solids make contact at a single point under the condition of zero applied load. Such a condition is called point contact and is shown in Figure 2.18, where the radii of curvature are denoted by \(r\)'s. It is assumed throughout the book that convex surfaces, as shown in Figure 2.18, exhibit positive curvature and concave surfaces, negative curvature. Therefore, if the center of curvature lies within the solid, the radius of curvature is positive; if the center of curvature lies outside the solid, the radius of curvature is negative. It is important to note that if coordinates \(x\) and \(y\) are chosen such that

\[
\frac{1}{r_{ax}} + \frac{1}{r_{bx}} > \frac{1}{r_{ay}} + \frac{1}{r_{by}}
\]  

(2.17)

coordinate \(x\) then determines the direction of the semiminor axis of the contact area when a load is applied and \(y\), the direction of the semimajor axis. In this chapter we are concerned only with unloaded contacts, but the notation has been chosen such that it is applicable for all the remaining chapters.

A cross section of a thrust-loaded ball bearing operating at a contact angle \(\beta\) is shown in Figure 2.19. Equivalent radii of
Curvature for both inner- and outer-race contacts in, and normal to, the direction of rolling can be calculated from this figure.

The radii of curvature for the ball - inner-race contact are

\[ r_{ax} = r_{ay} = \frac{d}{2} \quad \text{(2.18)} \]

\[ r_{bx} = \frac{d_e - d \cos \beta}{2 \cos \beta} \quad \text{(2.19)} \]

\[ r_{by} = -f_id \quad \text{(2.20)} \]

The radii of curvature for the ball - outer-race contact are

\[ r_{ax} = r_{ay} = \frac{d}{2} \quad \text{(2.21)} \]

\[ r_{bx} = -\frac{d_e + d \cos \beta}{2 \cos \beta} \quad \text{(2.22)} \]

\[ r_{by} = -f_id \quad \text{(2.23)} \]

Note that the ball - inner-race and ball - outer-race contact inequality (2.17) is satisfied and that the sign convention mentioned earlier has been adopted. In equations (2.19) and (2.22) \( \beta \) is used instead of \( \beta_f \) since these equations are also valid when a load is applied to the contact.

The curvature sum and difference, which are quantities of some importance in the analysis of contact stresses and deformations, are

\[ \frac{1}{R} = \frac{1}{R_x} + \frac{1}{R_y} \quad \text{(2.24)} \]
\[ \Gamma = R \left( \frac{1}{R_x} - \frac{1}{R_y} \right) \quad (2.25) \]

where

\[ \frac{1}{R_x} = \frac{1}{r_{ax}} + \frac{1}{r_{bx}} \quad (2.26) \]

\[ \frac{1}{R_y} = \frac{1}{r_{ay}} + \frac{1}{r_{by}} \quad (2.27) \]

Equations (2.26) and (2.27) effectively redefine the problem of two ellipsoidal solids approaching one another in terms of an equivalent ellipsoidal solid of radii \( R_x \) and \( R_y \) approaching a plane in the manner shown in Figure 2.20. From the radius-of-curvature expressions, the radii \( R_x \) and \( R_y \) for the contact example discussed earlier can be written for the ball-inner-race contact as

\[ R_x = \frac{d(d_e - d \cos \beta)}{2d_e} \quad (2.28) \]

\[ R_y = \frac{f_1 - 1}{2f_1 - 1} \quad (2.29) \]

and for the ball-outer-race contact as

\[ R_x = \frac{d(d_e + d \cos \beta)}{2d_e} \quad (2.30) \]

\[ R_y = \frac{f_0 d}{2f_0 - 1} \quad (2.31) \]
2.2.5 Geometrical Separation of Ellipsoidal Solids

The geometrical separation \((S_{ax} + S_{bx})\) or \((S_{ay} + S_{by})\) between two ellipsoidal solids is thus made equivalent to that between a single ellipsoidal solid and a plane in the manner shown in Figure 2.20. The geometrical requirement is simply that at any value of \(x\) and \(y\) (Figure 2.20(a)) the geometrical separation between the two ellipsoids must be the same as the separation between the equivalent ellipsoid and a plane at the same values of \(x\) and \(y\) (Figure 2.20(b)). From Figure 2.20(a) the following can be written:

\[
\frac{r_{ax}^2}{S_{ax}} = x^2 + (r_{ax} - S_{ax})^2
\]

or

\[
x^2 = S_{ax}(2r_{ax} - S_{ax})
\]

But \(2r_{ax} \gg S_{ax}\), so equation (2.33) becomes

\[
S_{ax} = \frac{x^2}{2r_{ax}}
\]

This is the well-known parabolic approximation to the circular section of the solid and is valid as long as the separation is much smaller than the radius of curvature. Similar expressions for \(S_{ay}, S_{bx}\), and \(S_{by}\) can be written, and the expression for the total separation of an ellipsoidal solid and a plane (Figure 2.20(b)) can thus be written as

\[
S = \frac{x^2}{2R_x} + \frac{y^2}{2R_y}
\]
Dowson and Higginson (1966) have shown in a similar way that the geometrical separation between two cylinders can be adequately described by an equivalent cylinder near a plane with the geometrical separation

\[ s^* = \frac{x^2}{2R_x} \tag{2.35} \]

The difference between equations (2.35) and (2.36) is that the radii of curvature in the \( y \) direction \( r_y \) and \( r_y \) are infinite for the case of two cylinders and thereby imply \( 1/R_y \) is zero.

2.3 Kinematics

The motion of the balls, rings, and separator of a ball bearing is important to the friction, temperature rise, life, and lubrication of the bearing. In this section it is assumed that there is no slip between the balls and races during rotation of the ball bearing, and as in the rest of the chapter it is also assumed that there are no distortions. The influence of high speeds, where centrifugal effects cause elastic distortions and a divergence of the inner- and outer-race contact angles from their zero-speed value, is not considered at this stage. If these assumptions were to be considered in the subsequent theory, a complex analysis and computer program would be required to obtain an approximate solution. Useful relations can, however,
be derived within these assumptions, which are relatively easy to apply and which can be used for the bulk of bearing applications, where modest speeds prevail.

Attempts to develop a parameter that would be an accurate gauge of the bearing operating limit have met with limited success. The criterion most frequently used to describe the speed capabilities of a ball bearing is the \( d_bN \) value (bore diameter in millimeters times rotative speed in revolutions per minute). It should only be used as a rough guideline for determining the limiting speed of a bearing because it does not take proper account of the size and centrifugal effects. The variation of limiting speed \( N \) with bore diameter \( d_b \) based on present practice and for limiting values of \( d_bN \) of \( 2 \times 10^6 \) and \( 3 \times 10^6 \) is shown in Figure 2.21 for ball bearings of varying bore diameter, from Bamberger, et al. (1971).

The surface velocities of the solid entering the conjunction between the ball and races are calculated for the general case in which both the inner and outer races rotate, even though in most practical situations only one would be rotating. The simplified solution can therefore be readily deduced from the general results. The various rolling speeds and dimensions to be found in a ball bearing are shown in Figure 2.22. From this figure and the assumptions mentioned earlier, the velocities at the inner- and outer-race contacts, as well as the velocity at the ball center, can be written respectively as.
\[ u_1 = \frac{d_e - d \cos \beta}{2} \omega_1 \quad (2.37) \]

\[ u_o = \frac{d_e + d \cos \beta}{2} \omega_o \quad (2.38) \]

\[ u_c = \frac{d_e}{2} \omega_c \quad (2.39) \]

Knowing these velocities, we can write two equations that define the surface velocities of the ball at the ball-race contact, namely,

Inner:
\[ u_B = \frac{d \cos \beta}{2} \omega_B = \frac{d_e - d \cos \beta}{2} \omega_1 - \frac{d_e}{2} \omega_c \quad (2.40) \]

Outer:
\[ u_B = \frac{d \cos \beta}{2} \omega_B = \frac{d_e}{2} \omega_c - \frac{d_e + d \cos \beta}{2} \omega_0 \quad (2.41) \]

Solving for \( \omega_B \) and \( \omega_c \) in these two equations gives

\[ \omega_B = \frac{1}{d \cos \beta} \left( \frac{d_e - d \cos \beta}{2} \omega_1 - \frac{d_e}{2} \omega_o \right) \quad (2.42) \]

\[ \omega_c = \frac{1}{d_e} \left( \frac{d_e - d \cos \beta}{2} \omega_1 + \frac{d_e + d \cos \beta}{2} \omega_0 \right) \quad (2.43) \]

The velocities of the surfaces entering the ball-race contact can be expressed by letting the coordinate system rotate about the bearing center with velocity \( \omega_c \). This fixes the ball-race geometry relative to the observer. The angular velocities of the
inner and outer races and the ball center after the transformation are

\[ \omega_{ir} = \omega_i - \omega_c \]  \hspace{1cm} (2.44)

\[ \omega_{or} = \omega_o - \omega_c \]  \hspace{1cm} (2.45)

\[ \omega_{Br} = \omega_B + \omega_c \]  \hspace{1cm} (2.46)

Subscript \( r \) refers to the rotating coordinate system. Using equations (2.42) and (2.43) we can write equations (2.44), (2.45), and (2.46) as

\[ \omega_{ir} = \frac{d_e + d \cos \beta}{2d_e} (\omega_i - \omega_o) \]  \hspace{1cm} (2.47)

\[ \omega_{or} = \frac{d_e - d \cos \beta}{2d_e} (\omega_o - \omega_i) \]  \hspace{1cm} (2.48)

\[ \omega_{Br} = \frac{d_e^2 + d^2 \cos^2 \beta}{2d_e d_e \cos \beta} (\omega_i - \omega_o) \]  \hspace{1cm} (2.49)

The surface velocities of the solids entering the ball-inner-race contact (where subscript \( a \) represents the inner race and subscript \( b \) the ball) are then

\[ u_{ai} = \frac{d_e - d \cos \beta}{2} \omega_{ir} = \frac{d_e^2 - d^2 \cos^2 \beta}{4d_e} (\omega_i - \omega_o) \]  \hspace{1cm} (2.50)

\[ u_{bi} = \frac{d \cos \beta}{2} \omega_{Br} = \frac{d_e^2 - d^2 \cos^2 \beta}{4d_e} (\omega_i - \omega_o) \]  \hspace{1cm} (2.51)
Since \( u_{ai} = u_{bi} \), the condition for pure rolling is satisfied.

Likewise, for the ball - outer-race contact

\[
\begin{align*}
    u_{ao} - u_{bo} &= \frac{d_e^2 - d^2 \cos \beta}{4d_e} \left( \omega_o - \omega_i \right) \\
(2.52)
\end{align*}
\]

Note that the condition for pure rolling is again satisfied at the ball - outer-race contact.

2.4 Materials and Manufacturing Processes

The good performance of ball bearings is as dependent on the proper choice and high quality of their materials as it is on their design and precise manufacturing. The operating conditions to which ball bearings are exposed call for the use of materials capable of withstanding high compressive stresses over millions of stress cycles without significant wear. The material must be hard enough to prevent excessive plastic deformation under these high compressive stresses and yet not be so brittle as to cause the bearing to fail when a shock load is applied. In addition to these high compressive stresses, bearings must withstand sub-surface shear stresses frequently in excess of \( 1.4 \text{ GN/m}^2 \) (200,000 lbf/in²). Because the stress at a given point may reach high values and then be reduced to zero at least once during each revolution of the shaft, it is necessary to use metals with a high fatigue strength.
Before this century the development of the ball bearing was greatly retarded by the lack of suitable materials. This situation was somewhat alleviated with the advent of case-hardened steels. However, using such materials makes it difficult to provide a case of great homogeneity. A single soft spot on a ball due to a slight deficiency of absorbed carbon is enough to significantly reduce the life of an otherwise well-made bearing. To combat this shortcoming, the through-hardening process was developed. In this process the hardness is achieved by cold working or by heat treating, rather than by a surface modification such as carburizing or nitriding.

The prime ball bearing material since about 1920 has been SAE 52100 (in Europe essentially the same material was called EN 31 and is now designated 534A99 or 535A99), and even today probably 90 percent of all ball bearings are made from this material. The 52100 material is a through-hardened, high-carbon chromium steel that also contains small amounts of manganese, silicon, nickel, copper, and molybdenum.

With the advent of machines like the gas-turbine jet engine, increased thermal demands have been placed on the materials used in ball bearings. This has resulted in the addition of such elements as molybdenum, tungsten, chromium, and vanadium to promote the retention of hardness at high temperatures. Table 2.1 lists the chemical compositions of a number of bearing steels. Of these steels M-50 is widely used by gas-turbine manufacturers in the United States, and 18-4-1 is widely used in Europe. Both
materials (M-50 and 18-4-1) contain elements to promote high hardness and good hardness retention at the high temperatures experienced by bearings in gas-turbine engines. Parker and Zaretsky (1978) found the rolling-element fatigue lives of these two steels to have no significant difference.

Johnson (1964, 1965) developed the AMS 5749 steel listed in Table 2.1. This material combines the tempering, hot-hardness, and hardness-retention characteristics of AISI M-50 steel with the corrosion and oxidation resistance of AISI 440C stainless steel. The impetus for the introduction of this steel, as pointed out by Valori (1978), was that estimates showed that the United States Navy yearly rejected M-50 engine bearings worth over $1 million because of corrosion. AMS 5749 contains a higher percentage of carbon and chromium than AISI M-50 for improved corrosion and wear resistance. The hot hardness and hardness retention of AMS 5749 are better than those of AISI 440C and similar to those of AISI M-50 (Johnson, 1964). Parker and Hodder (1978) experimentally found the rolling-element fatigue life of AMS 5749 to be 6 to 12 times greater than that of AISI M-50.

Bearings made from M-series steels can cost as much as 50 percent more than those made from standard SAE 52100 because of grinding difficulties. At temperatures below about 180º C (350º F) there appears to be no technical or cost advantage of M-series steels over SAE 52100.
2.4.1 Melting Practices

One important cause of rolling-element fatigue is nonmetallic inclusions, as pointed out by Anderson and Zaretsky (1973). These inclusions consist of sulfides, aluminates, silicates, and globular oxides and can act as stress raisers. Incipient cracks can emanate from these inclusions, as shown in Figure 2.23, and can enlarge and propagate under repeated stresses to form a network of cracks that in due course generate a fatigue spall or pit as shown in Figure 2.24.

The vacuum-melting process reduces or eliminates these nonmetallic inclusions, as well as entrapped gases and trace elements present in the metal. Exposing the melt to a vacuum permits deoxidation to be performed effectively by the carbons. The products formed when carbon is used as a deoxidizer are gaseous and thus escape into the vacuum, leaving no harmful residue in the metal. This process therefore results in a substantially cleaner material. Four prime vacuum-melting processes are vacuum induction melting, consumable-electrode vacuum melting, electroslag melting, and vacuum induction melting - vacuum arc remelting.

In vacuum induction melting (VIM), a cold charge is melted in a high-purity refractory crucible in an inductive furnace and subsequently poured while the melt is under a vacuum.

Consumable-electrode vacuum melting (CVM), also known as vacuum arc remelting (VAR), is a technique in which electrodes made by primary heating in air are remelted by an electric arc
process. The product thus remelted solidifies in a water-cooled copper mold under vacuum. The CVM method produces cleaner and more segregation-free material than the VIM process.

Electroslag melting or remelting (ESR), also known as electroflux remelting (EFR), has been used extensively in Europe and the U.S.S.R. for improving bearing steels. The ESR process is a consumable-electrode remelting procedure in which a molten flux blanket is maintained over the molten metal. The flux blanket serves essentially the same purpose as a vacuum atmosphere in protecting the ingot from contamination.

If the primary heat in the VAR process is vacuum induction melting, the process is called double-vacuum melting, or VIM-VAR. The high-quality VIM-VAR AISI M-50 material had a much longer rolling-element fatigue life than VAR material in accelerated fatigue life tests (Schlatter, 1974; Bamberger, 1972) and in ball bearing tests (Bamberger, et al., 1976). The steel produced by the EFR process, with VIM primary heating, also showed longer life than that produced by the VAR process with AISI M-50 in accelerated rolling-element fatigue tests (Schlatter, 1974; Bamberger, 1972), but the improvement was not as great as that with VIM-VAR. Furthermore Parker and Hodder (1978) show that VIM-VAR AMS 5749 steel had a rolling-element fatigue life 14 times greater than that obtained with VIM AMS 5749.

In the last 20 years vacuum-melting techniques have no doubt been most influential in improving the fatigue life of ball bearings. Generally vacuum-melted steels are no more expensive than
previously used air-melted steels. Most ball bearing manufacturers are now using vacuum-melted steels in their bearings. Furthermore, in applications where very high reliability is required, double-vacuum melting is often specified.

2.4.2 Material Hardness

Most ball bearing components are through-hardened within the range 60 to 65 Rockwell C. In general, bearing life can be extended by increasing material hardness in this range, where the material maintains its ductility. Furthermore Zaretsky, et al. (1967) found that bearings assembled with balls and races of various hardnesses had an optimum hardness combination for best fatigue life. Maximum life was obtained when the balls were 1 to 2 points harder than the races. The reason for this is that the residual stresses induced in the races are also a function of the hardness combination. Compressive residual stresses induced in the material reduce the magnitude of the maximum shear stresses.

2.4.3 Fiber Orientation

The races and balls of most ball bearings are formed by forging and thus have a fiber pattern that reflects the flow of metal during the forming process. Fatigue research on balls in-
dictated a strong effect of fiber orientation on fatigue. Materials with the fibers parallel to the stressed surface were strongest; those with fibers normal to the surface were weakest. This effect has been applied to bearings by using forged races. A sketch of the fiber flow in a conventional (top) and a forged race (bottom) is shown in Figure 2.25. Fatigue data reported by Anderson and Zaretsky (1968) show at least a tenfold increase in bearing life with forged races as compared with races cut from tubing. However, forged races are used only in those applications where high reliability is required because of the greater manufacturing cost.

2.4.4 Powder Metals

Another source of rolling-contact fatigue failures is the presence of large carbides in the metal matrix. These act as weak points under the highly cyclic loading experienced in bearing operation and ultimately become the origin of fine cracks, which lead to spalling and bearing failure. Conventional processing techniques have not been able to provide the necessary control of carbon size and distribution.

The use of powder metals is a possible way to produce alloys with highly refined carbides since the carbide size is limited by the size of the powder particles and does not change with subsequent processing. Likewise, a highly uniform distribution of
Carbides should be obtained in the final product because of the random mixing of particles.

The starting materials for powder-metal processing are conventionally processed ingots prepared by either vacuum induction melting or vacuum induction melting - vacuum arc remelting. Powder-metal processing has been described by Brown and Potts (1977):

Conversion of the ingots into powder was accomplished with equipment which is schematically represented in the sketch shown in Figure 7. (See Figure 2.26.) The atomizing system is enclosed and is evacuated to a pressure level which removes all traces of atmospheric contamination prior to remelting of the ingot. Melting is accomplished by induction heating techniques in a crucible in the upper section of the unit. After the metal becomes molten it is poured through an atomizing nozzle into another vacuum chamber located below the remelt chamber. The molten metal is discharged as refined droplets from the nozzle. High velocity inert gas bled into the chamber cools the droplets as they emerge from the nozzle. Vacuum pumping acts continuously to maintain the chamber pressure well below atmospheric. The cooled droplets solidify and drop to the bottom of the lower chamber. This powder now contains carbides that are many times smaller than those that were present.
in the original large ingot. The powders are to be re-consolidated into new ingots at temperatures well below alloy melt temperatures in order to ensure that no coalescence or growth of the alloy carbides will occur. By means of this process it is assured that the final product will exhibit refined carbides.

Notwithstanding the limited success (Brown and Potts, 1977) in using this process to date, it seems to offer promise for applications where rolling-element fatigue is a critical problem and extreme reliability is desired.

2.4.5 Ausforming

A thermomechanical process termed "ausforming" has been developed in the hope of extending bearing fatigue life. Ausforming consists of re-forming a bearing steel by using a hot-rolling procedure while the material is in a metastable austenitic condition. Bamberger (1967) was the first to apply the process to rolling-element bearings. To apply the ausforming process to bearing steels, the material must have a sluggish martensitic transformation behavior. The M-type steels exhibit such behavior and can be ausformed, whereas the SAE 52100 material cannot be treated in this way because of its austenite-to-martensite transformation characteristics. Bamberger's (1967) tests indicated that bearings made from ausformed AISI M-50 were superior to those
of normally processed M-50. Additionally a relationship was shown to exist between the amount of deformation suffered during ausforming and the fatigue life of the material. A 75 to 80 percent work, or reduction of area, was required to realize the maximum benefit.

2.5 Separators

Ball bearing separators, sometimes called cages or retainers, are bearing components that, although never carrying load, are nevertheless capable of exerting a vital influence on the efficiency of the bearing. In a bearing without a separator the balls contact each other during operation and in so doing experience severe sliding and friction. The separator therefore maintains the proper distance between the balls to ensure proper load distribution and to prevent sliding contact between the balls and thus permits safe operation of the bearing at all higher speed ranges. Furthermore a separator is necessary for several types of bearings to prevent the balls from falling out of the bearing during handling.

The materials used for separators vary according to the type of bearing and the application. The most common type of separator used in ball bearings is made from two strips of carbon steel that are pressed and riveted together. Called ribbon separators, they are the least expensive to manufacture and are entirely suitable for many applications.
The design and construction of the angular-contact bearing allow the use of a one-piece separator. The simplicity and inherent strength of one-piece separators permit their fabrication from many desirable materials. Reinforced phenolic and bronze are the two most commonly used materials. Bronze separators offer strength and low-friction characteristics and can be operated at temperatures to 230° C (450° F). Reinforced phenolic separators combine the advantages of light weight, strength, and nongalling properties; they are used for such high-speed applications as gyro bearings. Phenolic separators are, however, limited to a maximum operating temperature of about 135° C (275° F).

The temperature limits of the separator materials discussed in this section are shown in Figure 2.27. The bars represent the temperature ranges throughout which these materials can be expected to operate satisfactorily for extended periods of time. Where high speeds and high temperatures are encountered, separators of phosphor-bronze or iron-silicon-bronze are sometimes used. These materials have excellent antifriction properties and favorable strength-to-weight ratios.

At temperatures above 320° C (600° F) some success has been obtained with nickel-based alloys. Among the nickel-based alloys the most widely used is Monel. S-Monel has been used successfully for retainers in bearings at temperatures of 540° C (1000° F), as shown in Figure 2.27. This figure also shows that alloy and tool steels, suitably plated for low friction, perform well at temperatures to 540° C (1000° F).
2.6 Closure

In this chapter the most popular and basic types of ball bearings have been described. The geometric and kinematic relationships of the ball bearing under a condition of no load have also been discussed. These topics are extremely important in laying the foundation for the remaining chapters. The final part of the chapter dealt with bearing materials and manufacturing processes, as well as the separators used in ball bearings. With this background we now proceed to discuss loaded contacts.
SYMBOLS

A
 constant used in equation (3.113)

\{A^*, B^*, C^*\}
 relaxation coefficients

D^*, L^*, M^*

A_v
 drag area of ball, \(m^2\)

a
 semimajor axis of contact ellipse, m

\(\bar{a}\)
 \(a/2m\)

B
 total conformity of bearing

b
 semiminor axis of contact ellipse, m

\(\bar{b}\)
 \(b/2m\)

C
 dynamic load capacity, N

C_v
 drag coefficient

C_1, \ldots, C_8
 constants

c
 \(19,609 \text{ N/cm}^2 (28,440 \text{ lbf/in}^2)\)

\(\bar{c}\)
 number of equal divisions of semimajor axis

D
 distance between race curvature centers, m

\(\bar{D}\)
 material factor

\(\bar{D}\)
 defined by equation (5.63)

De
 Deborah number

d
 ball diameter, m

\(\bar{d}\)
 number of divisions in semiminor axis

\(d_a\)
 overall diameter of bearing (Figure 2.13), m

\(d_b\)
 bore diameter, m

\(d_e\)
 pitch diameter, m

\(d'_e\)
 pitch diameter after dynamic effects have acted on ball, m

\(d_i\)
 inner-race diameter, m

\(d_o\)
 outer-race diameter, m
modulus of elasticity, N/m²

effective elastic modulus, \( \frac{1 - \nu_a^2}{E_a} + \frac{1 - \nu_b^2}{E_b} \), N/m²

internal energy, m²/s²

processing factor

\( \frac{[(H_{\text{min}} - H_{\text{min}})/H_{\text{min}}] \times 100} \)

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/m²

hardness factor

gravitational constant, m/s²
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_E$</td>
<td>dimensionless elasticity parameter, $W^{8/3}/U^2$</td>
</tr>
<tr>
<td>$g_V$</td>
<td>dimensionless viscosity parameter, $GW/3/U^2$</td>
</tr>
<tr>
<td>$H$</td>
<td>dimensionless film thickness, $h/R_X$</td>
</tr>
<tr>
<td>$\hat{H}$</td>
<td>dimensionless film thickness, $H(W/U)^2 = F^2h/u^2n_0^2R_X$</td>
</tr>
<tr>
<td>$H_C$</td>
<td>dimensionless central film thickness, $h_C/R_X$</td>
</tr>
<tr>
<td>$H_{C,s}$</td>
<td>dimensionless central film thickness for starved lubrication condition</td>
</tr>
<tr>
<td>$H_f$</td>
<td>frictional heat, $N \cdot m/s$</td>
</tr>
<tr>
<td>$H_{\text{min}}$</td>
<td>dimensionless minimum film thickness obtained from EHL elliptical-contact theory</td>
</tr>
<tr>
<td>$H_{\text{min},r}$</td>
<td>dimensionless minimum film thickness for a rectangular contact</td>
</tr>
<tr>
<td>$H_{\text{min},s}$</td>
<td>dimensionless minimum film thickness for starved lubrication condition</td>
</tr>
<tr>
<td>$\tilde{H}_C$</td>
<td>dimensionless central film thickness obtained from least-squares fit of data</td>
</tr>
<tr>
<td>$\tilde{H}_{\text{min}}$</td>
<td>dimensionless minimum film thickness obtained from least-squares fit of data</td>
</tr>
<tr>
<td>$\bar{H}_C$</td>
<td>dimensionless central-film-thickness-speed parameter, $H_C^{U^{-0.5}}$</td>
</tr>
<tr>
<td>$\bar{H}_{\text{min}}$</td>
<td>dimensionless minimum-film-thickness-speed parameter, $H_{\text{min}}^{U^{-0.5}}$</td>
</tr>
<tr>
<td>$\tilde{H}_0$</td>
<td>new estimate of constant in film thickness equation</td>
</tr>
<tr>
<td>$h$</td>
<td>film thickness, m</td>
</tr>
<tr>
<td>$h_C$</td>
<td>central film thickness, m</td>
</tr>
<tr>
<td>$h_i$</td>
<td>inlet film thickness, m</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>$h_m$</td>
<td>film thickness at point of maximum pressure, where $dp/dx = 0$, m</td>
</tr>
<tr>
<td>$h_{min}$</td>
<td>minimum film thickness, m</td>
</tr>
<tr>
<td>$h_0$</td>
<td>constant, m</td>
</tr>
<tr>
<td>$I_d$</td>
<td>diametral interference, m</td>
</tr>
<tr>
<td>$I_p$</td>
<td>ball mass moment of inertia, m N s$^2$</td>
</tr>
<tr>
<td>$I_r$</td>
<td>integral defined by equation (3.76)</td>
</tr>
<tr>
<td>$I_t$</td>
<td>integral defined by equation (3.75)</td>
</tr>
<tr>
<td>$J$</td>
<td>function of $k$ defined by equation (3.8)</td>
</tr>
<tr>
<td>$J^*$</td>
<td>mechanical equivalent of heat</td>
</tr>
<tr>
<td>$\bar{J}$</td>
<td>polar moment of inertia, m N s$^2$</td>
</tr>
<tr>
<td>$k$</td>
<td>load-deflection constant</td>
</tr>
<tr>
<td>$k$</td>
<td>ellipticity parameter, $a/b$</td>
</tr>
<tr>
<td>$\tilde{k}$</td>
<td>approximate ellipticity parameter</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>thermal conductivity, N/s °C</td>
</tr>
<tr>
<td>$k_f$</td>
<td>lubricant thermal conductivity, N/s °C</td>
</tr>
<tr>
<td>$L$</td>
<td>fatigue life</td>
</tr>
<tr>
<td>$L_a$</td>
<td>adjusted fatigue life</td>
</tr>
<tr>
<td>$L_t$</td>
<td>reduced hydrodynamic lift, from equation (6.21)</td>
</tr>
<tr>
<td>$L_{1,\ldots,L_4}$</td>
<td>lengths defined in Figure 3.11, m</td>
</tr>
<tr>
<td>$L_{10}$</td>
<td>fatigue life where 90 percent of bearing population will endure</td>
</tr>
<tr>
<td>$L_{50}$</td>
<td>fatigue life where 50 percent of bearing population will endure</td>
</tr>
<tr>
<td>$z$</td>
<td>bearing length, m</td>
</tr>
<tr>
<td>$\bar{z}$</td>
<td>constant used to determine width of side-leakage region</td>
</tr>
<tr>
<td>$M$</td>
<td>moment, Nm</td>
</tr>
</tbody>
</table>
\( M_g \) | gyroscopic moment, Nm
--- | ---
\( M_p \) | dimensionless load-speed parameter, \( W/U - 0.75 \)
\( M_s \) | 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
\( \bar{m} \) | 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, \( qn_0/\rho_0E'\beta R^2 \)
\( q_f \) | reduced pressure parameter
\( q_x \) | volume flow rate per unit width in x direction, m^2/s
volume flow rate per unit width in y direction, m^2/s

curvature sum, m

arithmetic mean deviation defined in equation (4.1), m

operational hardness of bearing material

effective radius in x direction, m

effective radius in y direction, m

race curvature radius, m

radii of curvature, m

cylindrical polar coordinates

spherical polar coordinates

defined in Figure 5.4

geometric separation, m

defined in Figure 5.4

defined in Figure 5.4

defined in Figure 5.4

empirical constant

shoulder height, m

tangential (traction) force, N

temperature, °C

ball surface temperature, °C

average lubricant temperature, °C

ball surface temperature rise, °C

viscous drag force, N

time, s

auxiliary parameter

velocity of ball-race contact, m/s

46
$u_c$ velocity of ball center, m/s
$U$ dimensionless speed parameter, $n_0u/E'R_x$
$u$ surface velocity in direction of motion, $(u_a + u_b)/2$, m/s
$ar{u}$ number of stress cycles per revolution
$\Delta u$ sliding velocity, $u_a - u_b$, m/s
$v$ surface velocity in transverse direction, m/s
$W$ dimensionless load parameter, $F/E'R^2$
$w$ surface velocity in direction of film, m/s
$X$ dimensionless coordinate, $x/R_x$
$Y$ dimensionless coordinate, $y/R_x$
$X_t, Y_t$ dimensionless grouping from equation (6.14)
$X_a, Y_a, Z_a$ external forces, N
$Z$ constant defined by equation (3.48)
$Z_1$ viscosity pressure index, a dimensionless constant
$x, \bar{x}, \bar{X}, \bar{X}_1$ coordinate system
$y, \bar{y}, \bar{Y}, \bar{Y}_1$
$z, \bar{z}, \bar{Z}, \bar{Z}_1$
$a$ pressure-viscosity coefficient of lubrication, $m^2/N$
$a_a$ radius ratio, $R_y/R_x$
$\beta$ contact angle, rad
$\beta_f$ free or initial contact angle, rad
$\beta'$ iterated value of contact angle, rad
$\Gamma$ curvature difference
$\gamma$ viscous dissipation, $N/m^2$ s
$\dot{\gamma}$ total strain rate, $s^{-1}$
$\dot{\gamma}_e$ elastic strain rate, $s^{-1}$
$\dot{\gamma}_v$ viscous strain rate, $s^{-1}$
\( \gamma_a \)  
flow angle, deg

\( \delta \)  
total elastic deformation, m

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

\( \delta_0 \)  
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

\( \delta \)  
approximate elastic deformation, m

\( \tilde{\delta} \)  
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, \( N \text{s/m}^2 \)

\( \bar{n} \)  
dimensionless viscosity, \( n/\eta_0 \)

\( \eta_0 \)  
viscosity at atmospheric pressure, \( N \text{s/m}^2 \)

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

\( \theta \)  
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

48
\( \mu \) coefficient of sliding friction
\( \mu^* \) Poisson's ratio
\( \nu \) divergence of velocity vector, \((\text{au}/\text{ax}) + (\text{av}/\text{ay}) + (\text{aw}/\text{az}), \text{s}^{-1} \)
\( \rho \) lubricant density, \( \text{N s}^{2}/\text{m}^{4} \)
\( \bar{\rho} \) dimensionless density, \( \rho/\rho_0 \)
\( \rho_0 \) density at atmospheric pressure, \( \text{N s}^{2}/\text{m}^{4} \)
\( \sigma \) normal stress, \( \text{N/m}^{2} \)
\( \sigma_1 \) stress in axial direction, \( \text{N/m}^{2} \)
\( \tau \) shear stress, \( \text{N/m}^{2} \)
\( \tau_0 \) maximum subsurface shear stress, \( \text{N/m}^{2} \)
\( \tau_e \) shear stress, \( \text{N/m}^{2} \)
\( \tau_l \) equivalent stress, \( \text{N/m}^{2} \)
\( \phi \) limiting shear stress, \( \text{N/m}^{2} \)
\( \phi* \) ratio of depth of maximum shear stress to semiminor axis of contact ellipse
\( \phi_1 \) auxiliary angle
\( \phi \) primary angle
\( \theta_I \) thermal reduction factor
\( \psi \) angular location
\( \psi_l \) limiting value of \( \psi \)
\( \Omega_i \) absolute angular velocity of inner race, \( \text{rad/s} \)
\( \Omega_O \) absolute angular velocity of outer race, \( \text{rad/s} \)
\( \omega \) angular velocity, \( \text{rad/s} \)
\( \omega_B \) angular velocity of ball-race contact, \( \text{rad/s} \)
\( \omega_d \) angular velocity of ball about its own center, \( \text{rad/s} \)
\[ \omega_c \] 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:
- (\( \sim \)): 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.


58

ESDU (1965) "General Guide to the Choice of Journal Bearing Type,"
Engineering Sciences Data Unit, Item 65007, Institute of Mechanical Engineers, London.

ESDU (1967) "General Guide to the Choice of Thrust Bearing Type,"
Engineering Sciences Data Unit, Item 67033, Institution of Mechanical Engineers, London.


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, Windlasses of Ships, etc., Together with Metal Proper for the Machine, the Full Directions for Doing It."


Table 2.1 Chemical Compositions of Various Bearing Steels

<table>
<thead>
<tr>
<th>Bearing steel</th>
<th>C</th>
<th>Mn</th>
<th>Si</th>
<th>Mo</th>
<th>W</th>
<th>Cr</th>
<th>V</th>
<th>Co</th>
</tr>
</thead>
<tbody>
<tr>
<td>AISI 440C</td>
<td>1.00</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>----</td>
<td>17.0</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>SAE 52100</td>
<td>1.00</td>
<td>0.35</td>
<td>0.30</td>
<td>----</td>
<td>----</td>
<td>1.45</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>AMS 5749</td>
<td>1.15</td>
<td>0.50</td>
<td>0.30</td>
<td>4.00</td>
<td>----</td>
<td>14.5</td>
<td>1.20</td>
<td>----</td>
</tr>
<tr>
<td>AISI M-50</td>
<td>0.85</td>
<td>0.35</td>
<td>0.25</td>
<td>4.25</td>
<td>----</td>
<td>4.00</td>
<td>1.00</td>
<td>----</td>
</tr>
<tr>
<td>18-4-1</td>
<td>0.70</td>
<td>0.20</td>
<td>0.20</td>
<td>0.28</td>
<td>18.2</td>
<td>4.33</td>
<td>1.08</td>
<td>----</td>
</tr>
<tr>
<td>AISI M-10</td>
<td>0.85</td>
<td>0.25</td>
<td>0.30</td>
<td>8.00</td>
<td>----</td>
<td>4.00</td>
<td>2.10</td>
<td>----</td>
</tr>
<tr>
<td>AISI M-2</td>
<td>0.83</td>
<td>0.30</td>
<td>0.30</td>
<td>5.00</td>
<td>6.15</td>
<td>3.85</td>
<td>1.90</td>
<td>----</td>
</tr>
<tr>
<td>AISI M-1</td>
<td>0.80</td>
<td>0.30</td>
<td>0.30</td>
<td>8.00</td>
<td>1.50</td>
<td>4.00</td>
<td>1.00</td>
<td>----</td>
</tr>
<tr>
<td>WB-49</td>
<td>1.07</td>
<td>0.30</td>
<td>0.02</td>
<td>3.90</td>
<td>6.80</td>
<td>4.40</td>
<td>2.00</td>
<td>5.2</td>
</tr>
</tbody>
</table>

Rubbing brings oil-impregnated porous metal bearings
Rolling brings hydrodynamic oil-film bearings

Figure 2.1. General guide to journal bearing type. (Except for roller bearings, curves are drawn for bearings with width equal to diameter. A medium-viscosity mineral oil lubricant is assumed for hydrodynamic bearings.) From Engineering Sciences Data Unit (1965).
Figure 2.2. General guide to bearing type. (Except for roller bearings, curves are drawn for typical ratios of inside to outside diameter. A medium-viscosity mineral oil lubricant is assumed for hydrodynamic bearings.)
From Engineering Sciences Data Unit (1967).
Figure 2.3. - Elements of typical ball bearing.

Figure 2.4. - Cross section of ball and outer race.

Figure 2.5. - Assembly of Conrad bearing.
Figure 2.6. - Filled-notch ball bearing.

(a) Small angle.
(b) Large angle.

Figure 2.7. - Angular-contact ball bearing.

Figure 2.8. - Self-aligning ball bearing.
Figure 2.9. - Duplex pairs of angular-contact ball bearings.

Figure 2.10. - Split-inner-ring ball bearing.

Figure 2.11. - Thrust ball bearing.
Figure 2.12. - Ball bearing components. Example of nonconformal surfaces.

Figure 2.13. - Cross section through radial, single-row ball bearings.

Figure 2.14. - Cross section of ball and outer race, showing race conformity.

(a) Initial position.

(b) Shifted position.

Figure 2.15. - Cross section of radial ball bearing, showing ball-race contact due to axial shift of inner and outer rings.
Figure 2.16. - Free-contact angle and endplay as function of $P_d/2d_c$ for four values of total conformity.

Figure 2.17. - Shoulder height in ball bearing.

Figure 2.18. - Geometry of contacting elastic solids.
Figure 2.19. - Cross section of ball bearing.

Figure 2.20. - Equivalent ellipsoidal solids.
Figure 2.21 - Limiting $d_p N$ values for present practice and $3 \times 10^6 d_p N$ for ball bearings of various bore diameters. Based on information from Bamberger, et al. (1971).

Figure 2.22 - Rolling speeds in ball bearing.
Figure 2.23. - Fatigue crack emanating from inclusion. From Anderson and Zaretsky (1973).

Figure 2.24. - Typical fatigue spall in bearing race. From Anderson and Zaretsky (1973).
Figure 2.25. - Fiber flow in bearing races. From Anderson and Zaretsky (1968).

Figure 2.26. - Process for making pre-alloyed powder. From Brown and Potts (1977).

<table>
<thead>
<tr>
<th>Separator materials</th>
<th>Temperature, °C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Phenolic</td>
<td></td>
</tr>
<tr>
<td>Bronze</td>
<td></td>
</tr>
<tr>
<td>Iron silicon bronze</td>
<td></td>
</tr>
<tr>
<td>Phosphor bronze</td>
<td></td>
</tr>
<tr>
<td>S-Monel</td>
<td></td>
</tr>
<tr>
<td>Special tool steels</td>
<td></td>
</tr>
<tr>
<td>Alloy steel</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.27. - Temperature limits of separator materials.