ASM Handbook, Volume 19: Fatigue and Fracture
ASM Handbook Committee, p 331-336



# Contact Fatigue

W.A. Glaeser and S.J. Shaffer, Battelle Laboratories

CONTACT FATIGUE is a surface-pitting-type failure commonly found in ball or roller bearings. This type of failure can also be found in gears, cams, valves, rails, and gear couplings. Contact fatigue has been identified in metal alloys (both ferrous and nonferrous) and in ceramics and cermets.

Contact fatigue differs from classic structural fatigue (bending or torsional) in that it results from a contact or Hertzian stress state. This localized stress state results when curved surfaces are in contact under a normal load. Generally, one surface moves over the other in a rolling motion as in a ball rolling over a race in a ball bearing. The contact geometry and the motion of the rolling elements produces an alternating subsurface shear stress. Subsurface plastic strain builds up with increasing cycles until a crack is generated. The crack then propagates until a pit is formed. Once surface pitting has initiated, the bearing becomes noisy and rough running. If allowed to continue, fracture of the rolling element and catastrophic failure occurs. Fractured races can result from fatigue spalling and high hoop stresses.

Rolling contact components have a fatigue life (number of cycles to develop a noticeable fatigue spall). However, unlike structural fatigue, contact fatigue has no endurance limit. If one compares the fatigue lives of cyclic torsion with rolling contact, the latter are seven orders of magnitude greater (Ref 1). Rolling contact life involves ten to hundreds of millions of cycles.

## Examples of Contact Fatigue

Contact fatigue produces a surface damage that is unique and well recognized. Familiar examples are found in fatigue of ball and roller bearings. A typical failure in a roller bearing is shown in Fig. 1. Although this spall is small, it would grow in size until roller fracture would occur, as bearing operation continues.

One classic shape of a fatigue spall in a ball bearing is a delta shape, as shown in Fig. 2 with a diagram of the pit (Ref 2). The apex of the pit is the initiation point, usually the location of a surface defect like a dent. The pit grows in a fan shape, becoming wider and deeper as it grows in the direction of ball travel. Not all spalls in ball-bearing races are of the shape shown in Fig. 2.

Figure 3 shows a fatigue spall near the race shoulder of a deep-groove ball bearing. The spall appears to have been formed by the joining of several pits. The fact that the spall occurred close to the race shoulder may have distorted the contact state of stress, causing a multiple origin.

Fatigue in roller bearings may differ from ball-bearing contact fatigue. Quite often the pitting occurs in the inner race at the contact zone of the roller ends. In some cases, contact stress peaks at the roller ends and pitting originates in these locations. Roller-end pitting can be a sign of misalignment.

**Cams and Gears.** Valve lifter cams and rollers are subject to contact fatigue. An example is shown in Fig. 4 (Ref 3). The character of the damage is very similar to that found in rolling contact bearings. The example shown in Fig. 4 was found in both cam nose and lifters during automobile engine tests (Ref 3). Lifters were nodular iron, and cams were flake graphite cast iron. Fatigue cracks were associated with cracked carbides, graphite flakes, and hard inclusions.

Contact fatigue occurs in gears along the pitch line. The geometry of tooth mesh is such that rolling occurs at the pitch line while sliding occurs at the addendum as the gears come out of mesh. An example of pitch line contact fatigue is shown in Fig. 5 (Ref 4). The pits seen on the teeth will grow in size and depth, ultimately resulting in tooth fracture.

Another form of contact fatigue, known as micropitting, occurs in bearings. An example is shown in Fig. 6. This feature can show up over the entire raceway surface. It is often the result of too thin a lubricant film or excessive surface roughness and sometimes heavy loading.

In gears, micropitting is termed frosting and in the present ANSI/AGNA standard it is considered a form of contact fatigue. For bearings,

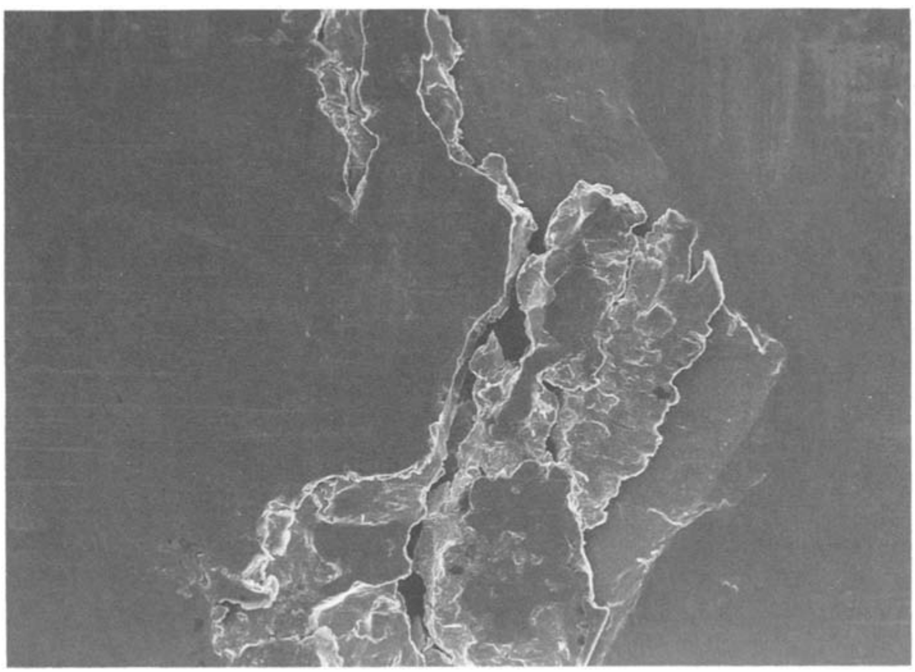


**Fig. 1** Scanning electron micrograph of a fatigue spall on a roller from a roller bearing after 630,000 cycles. Roller is AISI 1060 steel, hardened to 600 HV. Spall is 400 μm wide by 700 μm long.

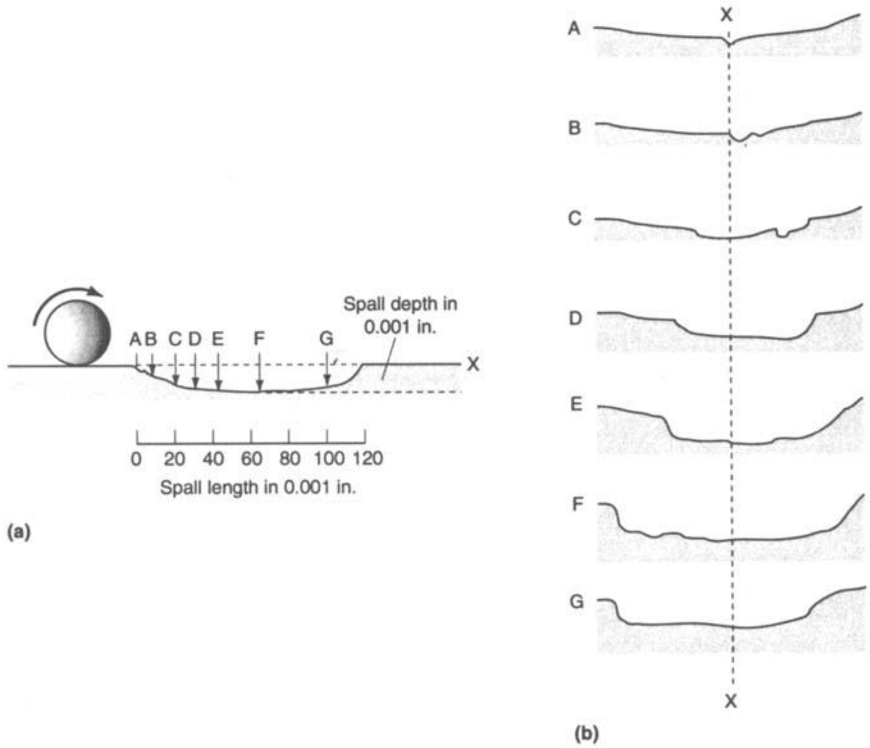



**Fig. 2** Anatomy of a race spall in a ball bearing. (a) Typical delta shape with the apex at the origin. (b) Profiles of the spall. Source: Ref 2

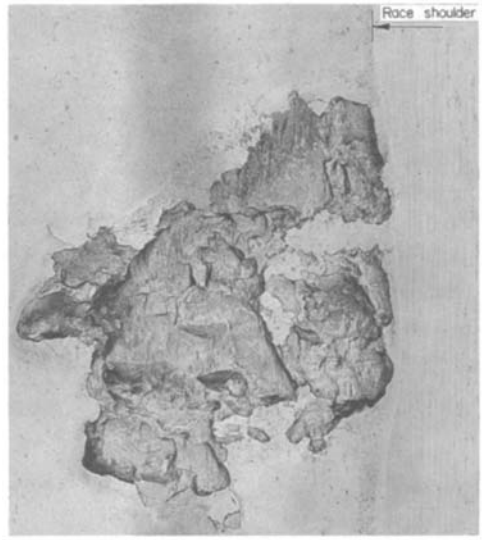



**Fig. 3** Multiple spall near a race shoulder

gears, and any contact, micropitting can be reduced by improved surface finish, reduced temperature or loads, and providing sufficient elastohydrodynamic film.

**Rails.** Spalling and "shelly"-type failures occur on track rails from wheel-track rolling contacts. An example of shelly failure is shown in Fig. 7 from Kilburn (Ref 5). The name comes from the morphology of the fracture surface in the bottom of the spall. Shelly failures are serious because they lead to rail fracture and derailments. Rail spalling has been reduced in recent years by the use of higher carbon steels for rails.

## Mechanisms of Contact Fatigue

The state of stress produced by rolling contact is concentrated in a small volume of material and produces intense plastic strain. The strain accumulates as the same volume is stressed with each rolling cycle until a crack is initiated and forms a spall. In the real world of contact fatigue, the mechanisms involved can be quite complex. Most models assume a condition of ideal geometric surfaces and little input by heat generation, environmental conditions, and inhomogeneities of materials. Hertz stress analysis assumes a circular, elliptical, or line contact surface area between curved surfaces (depending on the geometry of the contacts) and a parabolic pressure distribution with the maximum pressure at the center of the contact.

The subsurface state of stress involves a hydrostatic component that inhibits tensile fracture. Analysis of the subsurface stress state indicates that a maximum shear stress exists at a given depth below the surface. The stress distribution is shown in Fig. 8. The maximum shear stress is shown increasing with depth below the surface as discussed by Kloos and Schmidt (Ref 6).

The curves shown are based on two different mathematical approaches to the estimation of contact state of stress. Both approaches produce a shear stress distribution quite close to each other. The $z$ axis scale is $Z/B$, where $B$ is the minor axis length of the contact ellipse and the usual direction of rolling motion. Note in Fig. 8 that the maximum shear stress depth is about the same as the dimension of $B$. However, with increased traction or tangential force, the maximum shear stress moves closer to the surface. In hardened martensitic steels used in ball and roller bearings, the subsurface shear stresses produce plastic deformation in the martensitic structure. The residual strain increases with increase in rolling cycles. This has been shown by x-ray measurements during rolling contact experiments (Ref 7).

Many researchers have studied the microstructural changes that occur as a result of the buildup of subsurface strain (Ref 8-10). In AISI 52100 steel, a common rolling-contact-bearing material, the accumulation of strain initially is associated with the formation of a dark etching zone below the surface. Further strain causes the formation of light etching bands caused by the formation of a new ferrite phase. Then carbides in the high stress region begin to show decay and break up. Other microstructural features include "butterflies" or

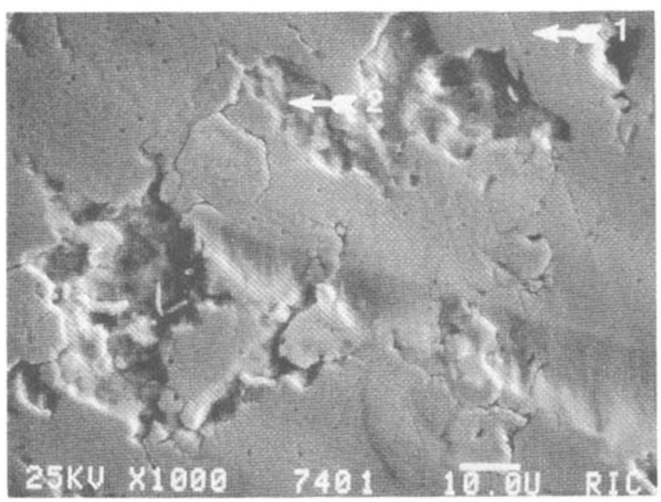



**Fig. 4** Contact fatigue spalling of cam lifter surface. Source: Ref 3

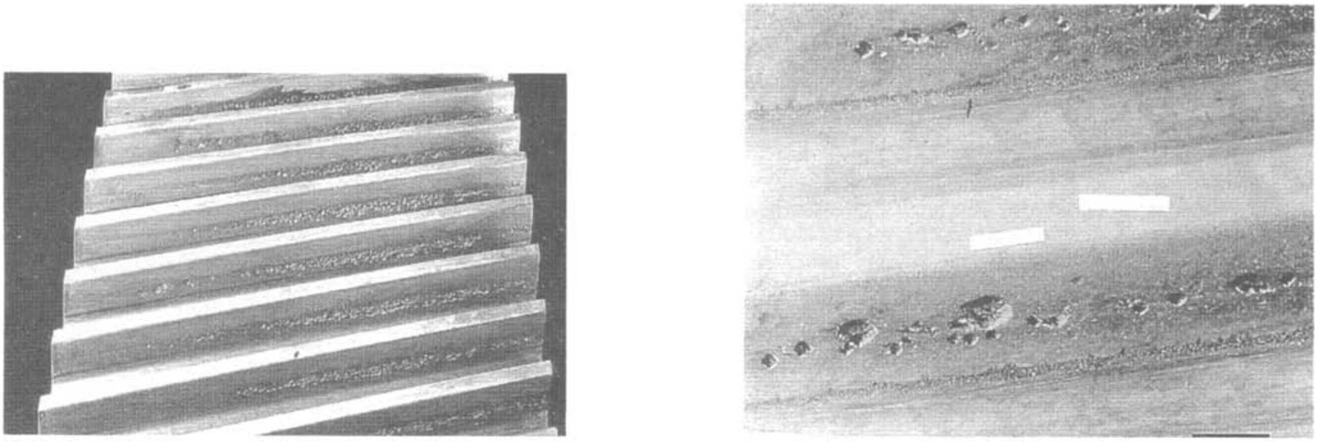
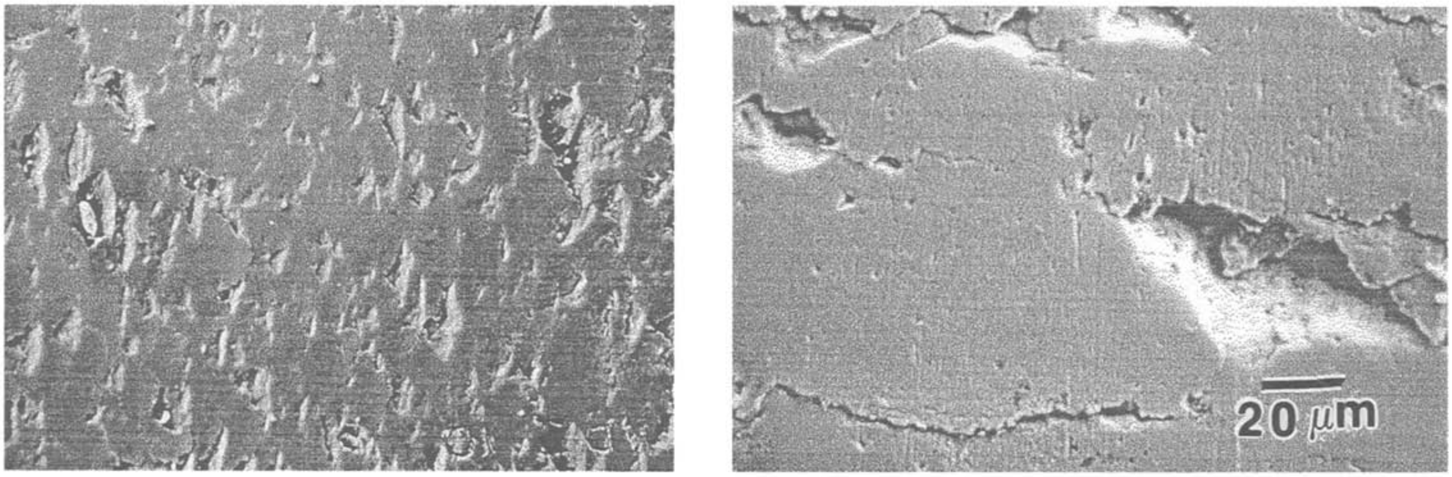


**Fig. 5** Pitch line spalling of medium-hardened gears. Source: Ref 4




(a) (b)

**Fig. 6** Micropitting of roller bearing outer race. Scanning electron micrograph, (a) 57× and (b) at higher magnification

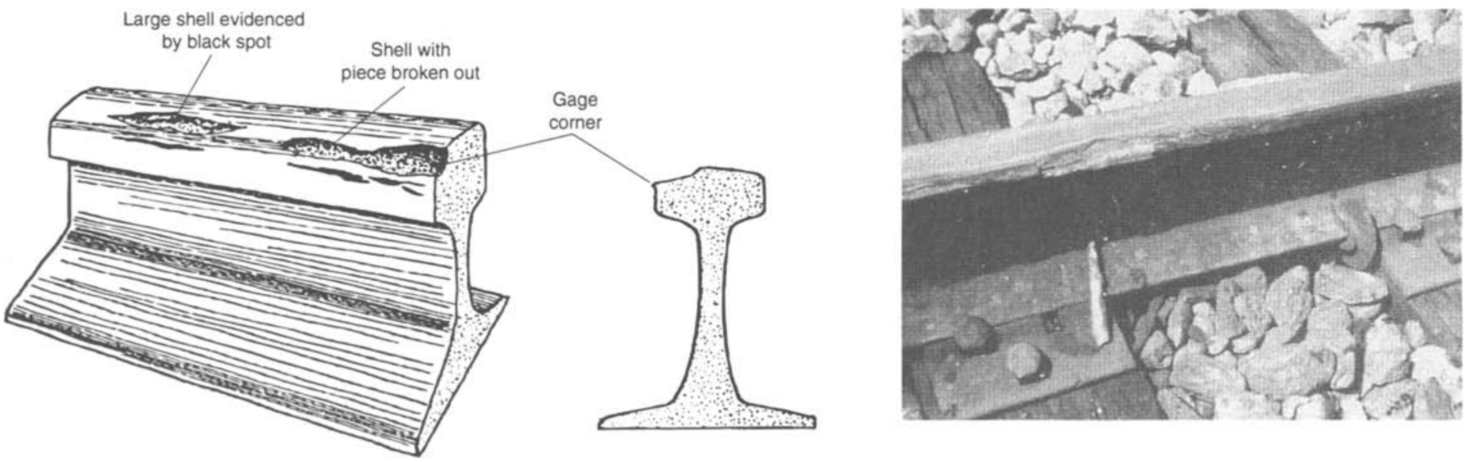



**Fig. 7** Shelly rail spall from wheel-rail contact fatigue. Source: Ref 5

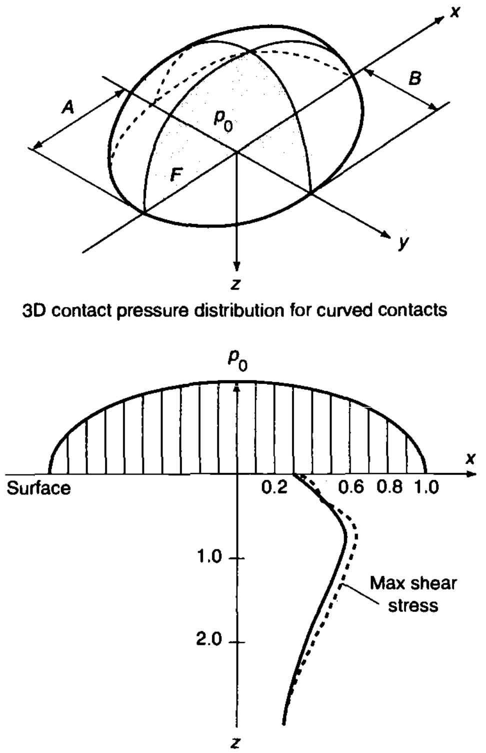



**Fig. 8** Stress distributions at a contact subsurface

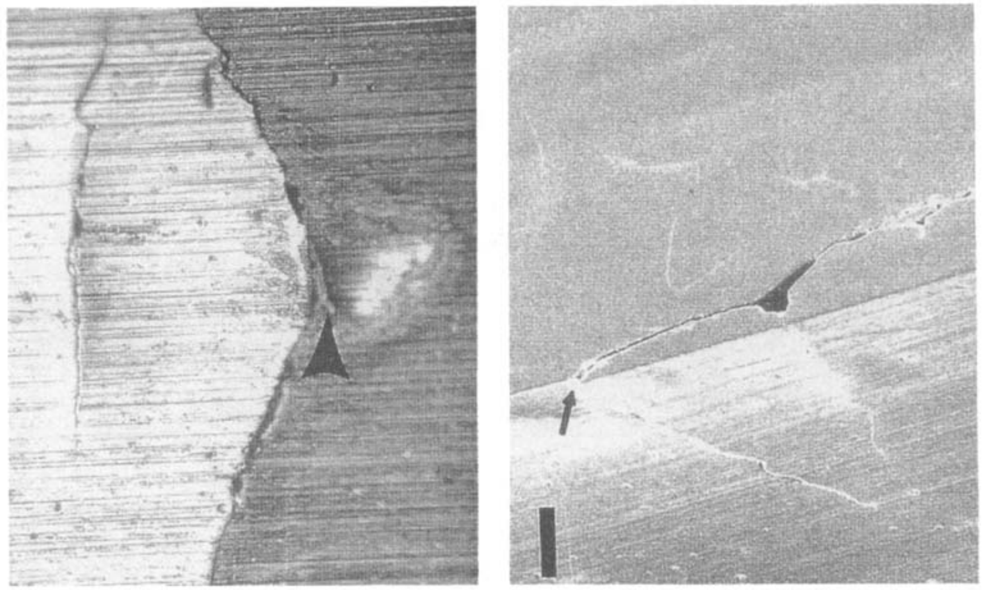


**Fig. 9** Developing spall. (a) Top view of developing spall at race surface dent. (b) Section through developing spall showing subsurface cracking Source: Ref 4

white etching wings radiating out from large hard inclusions. Some features depend on the level of contact stress. For example, butterflies are often associated with high contact stress (2000 to 4000 MPa), as discussed in more detail in the article "Contact Fatigue of Hardened Steel" in this Volume.

In a detailed study of butterfly formations in AISI 52100 steel ball bearings by Becker (Ref 11), both through-hardened AISI 52100 steel and carburized SAE 8620 steel bearing races were used. Contact stress was 3280 MPa (480 ksi). Butterflies were found in sectioned posttest races. They were always oriented at about 40° to the surface and oriented in the rolling direction. The "wings" of the butterflies were found to be composed of a mix of heavily strained, ultrafine-grained ferrite and fine carbide particles. Hardness was measured as close to 1000 HV—harder than the martensitic matrix surrounding the butterflies. Fine cracks were also found on the edges of the butterfly wings. The same structures were found in the carburized case in the 8620 steel. Becker says in Ref 11: "The breakdown of the matrix microstructure to ferrite and carbide is caused by very high stress concentration either at hard inclusions or at pre-existing cracks."

Contact fatigue is also surface generated. In fact, surface-originating spalls are more prevalent than subsurface-generated cracks. Proving subsurface fracture origin is difficult because a metallographic section only shows a profile of the crack which, in three dimensions, may have a surface origin. The higher the tangential force or traction, the more likely will be surface-generated contact fatigue. With shear stresses higher and closer to the surface, surface defects (dents, scratches, etc.) all contribute to higher incidence of surface-originating fatigue. Figure 2 shows a race spall that started at a dent in the race. This produced a delta-shaped spall as the cracking progressed from the origin. Sections through the spall show it to be shallow at the origin and deeper at the other end. Photomicrographs of a developing spall (Fig. 9, Ref 4) caused by a dent shows a ridge between the dent and the crack. This is typical and causes disruption in the oil film. The arrow shows the direction of movement of the balls over the race. The section through the developing spall shows the subsurface crack propagating down into the race at an angle to the surface.

As the developing spall matures, a surface layer loosens and eventually breaks out, leaving a pit. While the pit develops, the loose layer batters the fracture surface, obliterating the surface features. Fractographic analysis is not a likely option for investigating contact fatigue.

## Rolling Contact Bearing Life

Ball and roller bearings have been subject to the most extensive life testing of all contact fatigue components. Bearing catalog lives are based on fatigue failure considerations. It is assumed that no ball or roller bearing gives unlimited service. Owing to the special stress state experienced by rolling contact bearings, bending or push-pull tensile fatigue results cannot be applied to their life calculations. There is significant scatter in life tests for rolling contact bearings. The Weibull distribution is used in statistical analysis of bearing-life tests. A typical bearing-life Weibull plot is shown in Fig. 10 (Ref 12).

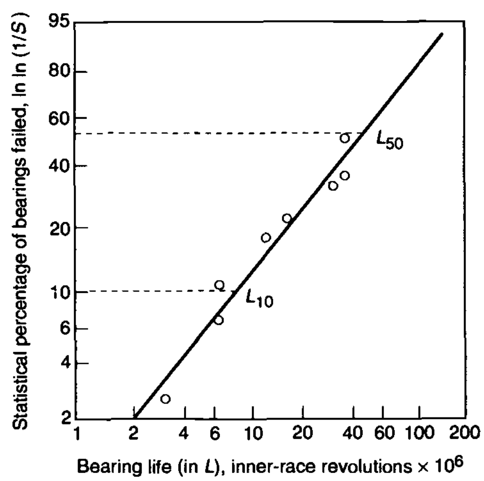



**Fig. 10** Weibull plot of ball-bearing lives distribution. Source: Ref 12

Two life values in the distribution are shown. The $L_{10}$ life, or the life at which 10% of all the bearings have failed is used for bearing selection. Lundberg and Palmgren (Ref 13) developed a relationship that can be used to predict bearing life for any load, using the life for standard load in the relation:

$$L = (C/P)^p$$

where $L$ is the fatigue life in revolutions × $10^6$; $C$ is the standard load ($C$ is defined as the load that gives an $L_{10}$ life of one million revolutions); $P$ is the selected load; and $p$ is 3 for ball bearings and 10/3 for roller bearings.

The predicted life from the above relationship is, of course, based on bearing tests, analyzed statistically. It does not take into account other

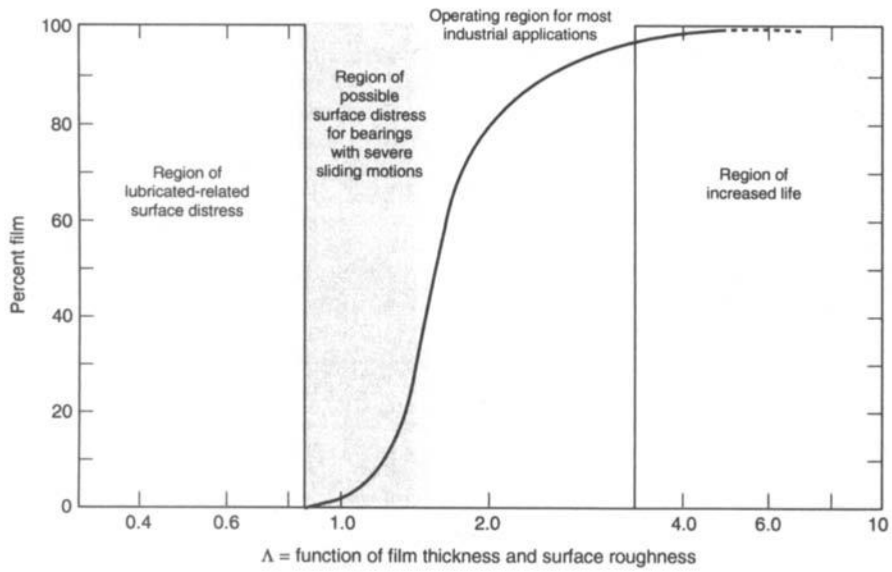



**Fig. 11** Ball-bearing performance map. Source: Ref 2

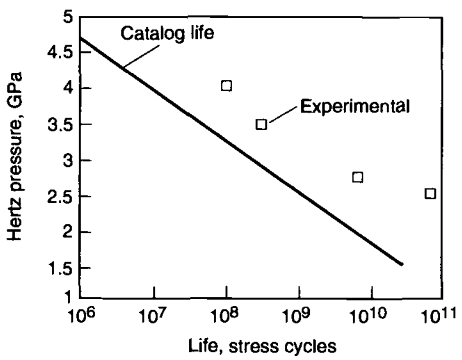



**Fig. 12** Contact stress-life plots for lives based on the inverse power load-life law and bearing tests with ideal operating conditions. Source: Ref 14

factors that impact on bearing life. Lubrication is a powerful factor in bearing life. Since the discovery of thin-film lubrication, elastohydrodynamic (EHD) lubrication of rolling contact bearings, the effect of film thickness on bearing life, has received considerable attention. Tests have shown that the lubricant film thickness is influenced by bearing speed and lubricant viscosity and less by load. A bearing performance map was developed by Harris (Ref 2). The performance map is shown in Fig. 11. It has been in general use for a number of years. The lubricant film coefficient, Λ, determined by dividing EHD film thickness by a surface roughness factor, relates to the present film or percentage time the surfaces are totally separated by a lubricant film. If Λ is less than 1, the bearing is likely to not attain the $L_{10}$ life predicted by the Weibull distribution. If Λ exceeds 4, then one might expect longer life than predicted.

Steel microstructure also has a significant effect on bearing life. Of greatest importance is the cleanliness of the steel. Because hard inclusions have been found to enhance the fatigue crack process, steelmaking methods have been modified to eliminate or substantially reduce the production of hard inclusions. Consumable electrode vacuum melting has produced bearing-grade steels that have dramatically improved bearing reliability. In many cases, bearing failure is now related to wear rather than to contact fatigue. Good surface finish is necessary for long bearing life. As was noted, contact fatigue is initiated by surface defects like dents and deep scratches. Surface defects not only cause asperity contact in thin-film lubrication, but dents have been shown to disturb the EHD film and cause local film breakdown.

The possibility of a fatigue limit for rolling contact (deviation from the inverse load power law) has been investigated. Tallian (Ref 14) has analyzed test data from bearing tests run at high Λ values under conditions free of contaminants and debris and found deviations from the theoretical life suggesting a fatigue limit. This is shown in the plot in Fig. 12. Further information on bearing life is described in the article "Fatigue and Life Prediction of Bearings" in this Volume.

## Minimizing Contact Fatigue

The study of rolling contact behavior has indicated new approaches that might further improve the contact fatigue resistance of these systems. One important problem in the application of rolling contact systems is the wide scatter in failure lives. Zaretsky (Ref 15) indicates that in a group of 30 ball bearings the ratio of the longest to the shortest life may be as much as 20 times. Bearing-catalog ratings are based on $L_{10}$ lives or the time in which 10% of the bearings have failed. The same scatter can be expected in other rolling contact components such as gears and cams. Bearing-fatigue life is sensitive to bearing load. Generally, it is assumed that life is inversely proportional to maximum Hertz stress to the 9th power. The exponent can be as high as 12 for roller bearings. Lubricant composition, microstructure, and geometry of rolling contacts can influence these exponents of life. Zaretsky, Poplawski, and Peters (Ref 15) give a good summary of life exponents from various tests for a number of bearing configurations. This also includes case-carburized consumable electrode vacuum arc remelted AISI 9310 spur gears. The $L_{10}$ lives varied inversely with the stress to the 8.4 power for the spur gears.

Reducing the scatter in bearing lives so that the distribution would be compressed toward the longer lives would intrinsically improve bearing lives. Improving rolling-element precision, surface finish, and homogeneity of microstructure should reduce scatter somewhat. Lubrication also is effective. Ensuring that a rolling contact component is operating within satisfactory EHD conditions relative to surface finish is essential. Cleanliness of the operating environment and reasonable protection from corrosion are also important.

Because of the sensitivity of contact fatigue life to contact stress, reduction of contact stress can significantly improve bearing life. Of course, accurate estimation of the actual operating contact stress is important. Contact stress can be reduced by spreading out the area contact with a soft solid thin film applied to the surfaces (bearing races, for instance). Conversely, hard coatings have been used to improve fatigue life of bearing steels (Ref 16).

High-speed ball bearings have an increased ball contact stress owing to centrifugal forces. Such increased stress levels are sufficient to cause significant reduction in fatigue life even in very clean precision ball bearings. Reduction in the ball mass can reduce this effect and increase life to reasonable levels. Significant advances have been made by the use of silicon-nitride balls for high-speed bearings. Because of the lower density of silicon nitride, centrifugal forces in the bearing are reduced. Hybrid ball bearings with silicon-nitride balls have surpassed bearing grade steel in rolling contact performance (Ref 17-19). These bearings are finding use in gas turbines and high-speed machine tools.

Race fracture in high-speed ball bearings can be avoided by using a carburizing grade steel with increased fracture toughness (M50 NiL) (Ref 20) instead of through-hardening steels like AISI 52100. Carburizing to a depth below the estimated maximum shear depth will provide the required resistance to contact fatigue. Cleanliness of the steel will still be an important factor in bearing life.

The residual stress state in the near surface of rolling contact elements resulting from heat treatment and machining have an influence on contact fatigue life. By imposing compressive residual stresses, gear life can be improved. This can be accomplished by shot peening and burnishing. Nitriding gear steel will produce the desired compressive residual stresses to inhibit crack propagation.

As power systems become lighter and more compact, bearings, gears, and other rolling elements will have to operate at higher speeds. Although even at this time not all is understood about the mechanisms of contact fatigue, advances in improved reliability and component life are being made. Research and testing continue to try to narrow the life scatter and increase the predicted life of rolling contact parts.

## REFERENCES


1. V. Bhargava, G.T. Hahn, and C.A. Rubin, Rolling Contact Deformation and Microstructural Changes in High Strength Bearing Steel, *Wear*, Vol 133, 1989, p 69
2. T. Harris, "The Endurance of Modern Rolling Bearings," AGMA paper 269.01, Oct 1964, *Rolling Bearing Analysis*, John Wiley, 1966
3. S.H. Roby, Investigation of Sequence IIIE Valve Train Wear Mechanisms, *Lubr. Eng.*, Vol 47 (No. 5), p 413-430
4. T.E. Tallian, *Failure Atlas for Hertz Contacts*, ASME, 1992
5. K.R. Kilburn, An Introduction to Rail Wear and Rail Lubrication Problems, *Wear*, Vol 7, 1964, p 255-269
6. K.H. Kloos and F. Schmidt, Surface Fatigue and Wear, *Metallurgical Aspects of Wear*, K.H. Zum Gahr, Ed., Deutsche Gesellschaft fur Metallkunde, 1981, p 163-182
7. K.H. Kloos and F. Schmidt, Surface Fatigue and Wear, *Metallurgical Aspects of Wear*, K.H. Zum Gahr, Ed., Deutsche Gesellschaft fur Metallkunde, 1981, p 163-182
8. J.A. Martin, S.F. Borgese, and A.D. Eberhardt, *Trans. ASME*, Vol 59, 1966, p 555
9. H. Swahn, P.C. Becker, and O. Vingsbo, *Met. Sci.*, Jan 1976, p 35
10. W.D. Syniuta and C.J. Corrow, *Wear*, Vol 15, 1970, p 187
11. P.C. Becker, Microstructural Changes Around Non-metallic Inclusions Caused by Rolling-Contact Fatigue of Ball-Bearing Steels, *Met. Technol.*, June 1981, p 234-243
12. R.J. Boness, W.R. Crecelius, W.A. Ironside, C.A. Moyer, E.E. Pfaffenberger, and J.V. Poplawski, Current Practice, *Life Factors for Rolling Bearings*, E.V. Zaretsky, Ed., Society of Tribologists and Lubrication Engineers, 1992, p 5-7
13. G. Lundberg and A. Palmgren, Dynamic Capacity of Rolling Bearings, *Acta Polytechnia*, Mechanical Engineering Series 1, R.S.A.E.E., No. 3, 7, 1947
14. T.E. Tallian, Unified Rolling Contact Life Model with Fatigue Limit, *Wear*, Vol 107, 1986, p 13-36
15. E.V. Zaretsky, J.V. Popiawski, and S.M. Peters, "Comparison of Life Theories for Rolling-Element Bearings," Preprint 95-AM-3F-3, Society of Tribologists and Lubrication Engineers, May 1995
16. A. Erdemir, Rolling Contact Fatigue and Wear Resistance of Hard Coatings on Bearing Steel Substrates, *Surf. Coat. Technol.*, Vol 54-55 (No. 1-3), 1992, p 482-489
17. R.J. Parker and E.V. Zaretsky, Fatigue Life of High-Speed Ball Bearings with Silicon Nitride Balls, *J. Lubr. Technol. (Trans. ASME)*, 1975, p 350-357
18. F.J. Ebert, Performance of Silicon Nitride Components in Aerospace Bearing Applications, *Proc. Gas Turbine and Aeroengine Congress*, 11-14 June 1990, American Society of Mechanical Engineers
19. M. Hadfield, S. Tobe, and T.A. Stolarski, Subsurface Crack Investigation of Delaminated Ceramic Elements, *Tribol. Int.*, Vol 27 (No. 4), 1994, p 359-367
20. C.A. Moyer and E.V. Zaretsky, Failure Modes Related to Bearing Life, *Life Factors for Rolling Bearings*, E.V. Zaretsky, Ed., Society of Tribologists and Lubrication Engineers, 1992, p 67