Available online at www.sciencedirect.com
ScienceDirect
Chinese Journal of Aeronautics 23(2010) 123-136
Chinese Journal of Aeronautics
www.elsevier.com/locate/cja

Fundamentals of Design and Technology of Rolling Element Bearings

Franz-Josef Ebert*

FAG Aerospace GmbH & Co. KG, D-97421 Schweinfurt, Germany

Received 26 March 2009; accepted 31 August 2009

Abstract

This article presents an overview of fundamentals of rolling element bearing designs and technologies. The aim is to outline the complex interrelation of all fundamentals with the rolling contact fatigue and the attainable bearing life. Such fundamentals include, amongst others, bearing stressing and life capability. The article shows the different types of rolling bearing stressing and the analysis of the stress distribution (principal stresses and equivalent stresses) in the material under the rolling contact area. It becomes obvious that the contamination of the bearing with foreign particles leads to a drastic reduction in bearing life. Furthermore, it demonstrates the impact of the bearing lubrication and coating as well as the effect of additives on the attainable life and wear. Further fundamentals of rolling element bearing design are the materials, the material cleanliness and the heat treatment. The article reveals the importance of the cleanliness of bearing steels as well as different types of inclusions and their effect on rolling contact fatigue. Additionally the article describes how to optimize the material properties (strength, toughness and residual stress) by the heat treatment processes. The outcome of these investigations is that endurance life of a rolling element bearing can be achieved if specific operating conditions, an adequate lubrication, good system cleanliness and specific bearing stressing are met. The article provides a guideline for bearing engineers on how designs and technologies can be applied to optimizing a bearing for a particular industry or aerospace application.

Keywords: aerospace; rolling contact fatigue; endurance life; materials; heat treatment; lubrication; reliability

---

1. Introduction

It is a continuous goal for bearing engineers to always develop advanced bearings that provide higher efficiency, lower friction, and are less sensitive and more reliable even in applications under adverse operating conditions. The more critical and demanding a bearing application is, the more advanced the applied bearing technologies need to be in order to meet the operating requirements. It is further important for both, the rolling bearing user and the rolling bearing manufacturer, to closely cooperate starting in an early design phase to make sure that the application requirements and the bearing performance capabilities will be brought in line with each other, and to enable the execution of joint value engineering activities in order to achieve the overall best design solution at the lowest system life cycle cost for the user[1] .

Continuous research and development efforts are required by the bearing manufacturer to steadily advance the bearing performance capabilities. This involves the bearing designs and analysis aspects as well as materials, heat treatment and surface engineering technologies, and last but not least the bearing manufacturing technologies, which all together do have a substantial impact on the bearing performance capabilities and operating reliability[1] .

This article presents an overview of some of the fundamental aspects regarding the design and development of rolling element bearings.

2. Bearing Life and Endurance Limit

The FAG Kugelfischer research work in the 1970s and the early 1980s laid the basis for a very important step regarding the knowledge and understanding of bearing life capabilities. It is connected with the discovery of the endurance strength. This discovery

---

*Corresponding author. Tel.: +49-9721-913551.
E-mail address: ebertfra@schaeffler.com

doi: 10.1016/S1000-9361(09)60196-5

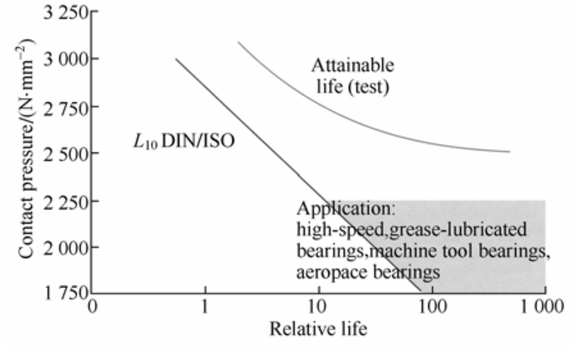
was sheer revolutionary. It immediately had a stimulating effect on the development of bearing technology.

2.1. Rolling bearing endurance strength

Endurance strength has long been known for materials and components. It was not found until some time ago for rolling bearings although researchers had been wondering for a long time why rolling bearings could not achieve endurance life like known for many other components. It was left to FAG to prove that there is an endurance strength for rolling bearings, and to describe the necessary preconditions under which endurance life for bearings can be achieved.

Fig.1 displays the correlations which FAG discovered and introduced in international presentations and publications for the first time in 1981. In comparison to the assumed life behavior at that time, standardized in DIN/ISO, the left straight line, FAG found the actual correlation between the life and the specific bearing stressing, which the right curve in Fig.1 shows[2] . It has been demonstrated that endurance life can be achieved with rolling element bearings. Prerequisites for the validity of the right curve are:1) surface separation by an elastohydrodynamic lubrication film, 2) clean lubricant.

Typical operating stresses of some rolling bearing applications are also entered in Fig.1 in order to demonstrate the importance of that finding for these and many other applications. In these applications the prerequisites for endurance life are either completely or largely met, thereby presenting great opportunities for life improved bearings.



L_{10} DIN/ISO—Rated life with 10% failure probability

Fig.1 Actual attainable life in comparison to DIN/ ISO life calculation (unfactored).

With the high influence of stress, which is clearly visible here, the development aims of minimizing the bearing stressing and maximizing the load carrying capability gain particular importance.

Minimizing the contact stressing in the bearing with a given external load is a particular challenge to the bearing engineer. It is important to optimize the bearing design under consideration of the surrounding parts and construction.

A high degree of cleanliness reduces stress as well. It prevents particle indentations with stress concentration in the rolling contact area during the subsequent cycling.

The operating conditions also influence the bearing stressing greatly. Good surface separation by the lubrication film and additives in the lubricant at mixed friction reduce additional stress arising from friction within the contact area.

Material technology in particular is necessary for maximizing strength. After minimizing stress, the material strength determines the attainable life. Of particularly importance are the material properties and the properties such as residual stresses induced by the machining process as well as by the heat treatment.

The endurance life behavior presented by FAG is meanwhile confirmed and verified in millions of bearing applications that have been designed since FAG introduced the adjusted fatigue life calculation method in 1985.

2. 2. Real fatigue life behavior of rolling bearings

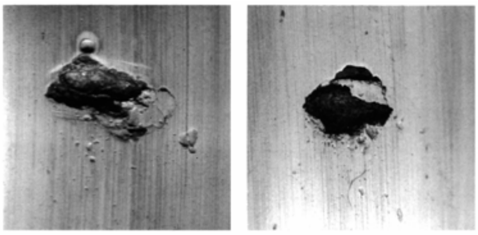
The real fatigue life behavior of the rolling bearing was not discovered until relatively recently.

The causes can be seen in two points principally.

(1) With the high loads which were formerly used to accelerate ageing in life tests, the negative effect of contamination was comparatively small. For this reason it was underestimated.

(2) At the same time the same kind of damage patterns with and without particle cycling disguised the actual damage course. Fatigue cracks develop on damaged surfaces, particularly at raised edges around the particle indentations on the rolling elements in contact. The subsequent cycling causes them to penetrate into the depths to the zone of the highest equivalent stress where they propagate and ultimately cause pitting and flaking of material. The pitting (see Fig.2(a)) has developed from an indentation which was intentionally produced. At its early stage, such flaking cannot be differentiated from classical fatigue pitting where the crack formation starts at a weak spot somewhere below the surface in the most highly stressed subsurface material zone (see Fig.2(b)).

FAG had adjusted the life calculation process standardized in DIN/ISO that takes the latest findings well into consideration[3]. For aerospace bearings these new findings and the life calculation method have been adopted accordingly[4-6] .



(a) Induced by one indentation (b) "Classical"(real) fatigue pitting of 1.7 mm width

Fig.2 Pittings with different damage origins.

Fig.3 shows the different life reactions of rolling element contacts, which are well lubricated and clean, to changes in contact pressure, shown in comparison to the unfactored DIN/ISO life. From Fig.3, it is shown that with minor load changes, major changes in the bearing life may result in well-lubricated elements in rolling contact and under a high degree of cleanliness.

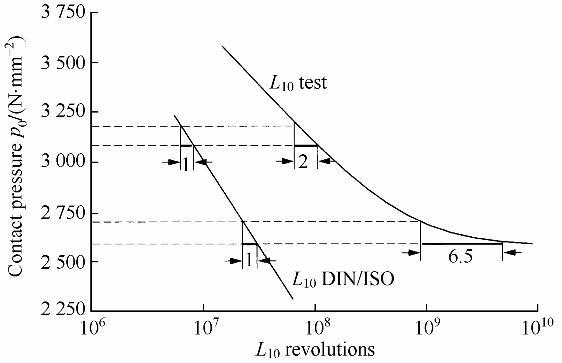


Fig.3 Contact pressure vs L_{10} revolutions.

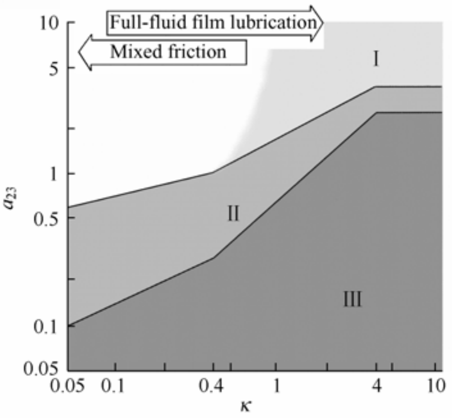
These findings are used to optimize bearing designs and to increase reliability and efficiency of bearings and bearing systems, and have become standardized in today's DIN/ISO codes.

The strong dependence of the bearing life capability on the amount of stress and on the conditions of lubrication and cleanliness is obvious, and the significance in bearing and system designs provides an important guideline to all designers.

3. Lubricants at Mixed Friction Full-fluid Film Lubrication, Mixed Friction

According to the known principles of the formation of a separating lubricating film between surfaces in rolling contact with oil or grease lubrication, rolling bearings reach the longest possible life with full-fluid film lubrication.

As can be seen in the a_{23} diagram (see Fig.4), a multiple of the nominal unfactored L_{10} life according to/DIN/ISO 281 is attainable in the zone of full-fluid film lubrication. The metallic surfaces of the components in rolling contact are separated by a lubricating film which prevents asperity contacts between the roughness peaks to a large extent. Therefore, "physical lubrication" by a sufficiently thick lubricating film should always be aimed for[7] .



a_{23}—Factor for operating conditions; K—Viscosity ratio; v—Operating viscosity; v_1—Rated viscosity

Fig.4 Areas of full-fluid film lubrication and mixed friction in the a_{23} diagram.

Full-fluid film lubrication cannot, however, be reached under certain operating conditions such as low speed, very high load, and at very high temperature a thin, inadequately separating lubricating film is built up. These operating conditions are indicated in the a_{23} diagram by $\kappa<4(\kappa=v/v_1)$ and specially marked by $\kappa<0.4$. In these ranges, metal-to- metal contact and adhesion occur, at least partly, so that high normal and tangential stresses and wear cannot be avoided[8]. The life is substantially reduced by such adverse influence.

Where a "physical lubrication" is not given, separation of the contact surfaces is reached either by boundary layers built by physical adsorption or from the reaction of additives and lubricating oil with the bearing material, or by a dry lubricant layer which is firmly bonded to one or more surfaces; the layer is formed by paste, powder, compound or coating, including bondered and thin chromium coatings.

The formation of boundary layers with oil or grease lubrication is described below.

3.1. Boundary layer

A boundary layer is created by reactions of the additives in the near-surface zone under the conditions existing at the contact area such as pressure, sliding motion and temperature. The repeated reaction yields a constant renewal of the layer in the case of layer loss and thus for a steady separation effect.

The boundary layer thickness varies between a few tenths of a nanometer (nm) and several nanometers. It includes chemical compounds containing, for example, phosphorus, sulphur, oxygen, carbon, zinc, copper, and nitrogen.

Various elements contribute to the layer formation depending on the additive properties. The share of such elements decreases differently with the penetration depth into the surface zone and thus into the boundary layer and can be measured by Auger analysis which works in the nanometer range.

3. 2. *Operating behavior and damages in mixed friction range*

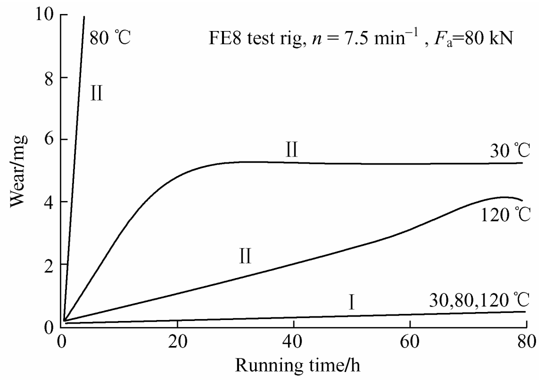
Mixed friction affects friction and wear behavior depending on the lubrication and the extent of sliding friction that occurs within the bearing contact areas, and on other rubbing areas such as the cage piloting surfaces, the rolling element/cage contacts or the roller endface contacts, just to mention the most important ones. Bearing types with higher sliding at mixed friction have coefficients of friction that can be a multiple of the coefficient with full-fluid film lubrication, depending on the type of lubrication. Rolling contact areas with large shares of sliding friction are exposed to severe wear under unfavorable lubrication.

If lubricants with extreme pressure (EP) additives are used in the mixed friction range, the contact surfaces are usually smoothened and the friction reduced to a moderate extent. However, some lubricants lead to wear which changes the geometry, in particular in zones with sliding friction. Depending on additives and temperature, different patterns of wear development can be recognized, as shown by the example of a cylindrical roller thrust bearing (see Fig.5).

1) At temperatures between 30 ℃ and 120 ℃ there is no measurable wear with EP oil I; the additive package contains additives which are effective at different temperatures.

2) With EP oil II at 30 ℃, running-in wear occurs during the first 20 operating hours, then wear stops. The running-in wear is found preferably near the roller ends and the running track edges which are the sliding zones; i.e., protection is insufficient.

3) Wear increases steeply to moderately with EP oil II at 80 ℃ to 120 ℃. The additive package is more effective at 120 ℃ than at 80 ℃.



n—Rotation speed; F_a—Axial load

Fig.5 Wear pattern vs the running time (bearing 81212 under short-term test conditions with EP oils I and II at varying temperatures).

4. Material and Heat Treatment

The demands on the material used for elements in rolling contact may vary just as widely as the types and sizes of rolling bearings[1] . High strength, resistance to wear, high-temperature hardness and dimensional stability, corrosion resistance, certain magnetic properties and low-noise operation are bearing characteristics, some of which tend to have a contradictory effect. Therefore a thorough analysis of the stress condition is required for selecting the appropriate material. With the knowledge gathered as a rolling bearing universalist, FAG is able, in dialogue with the customer, to select the most suitable material for the application in question. Selection criteria are always of the technical-economical type, which means, in addition to quality and reliability, availability and machinability play a decisive part.

4. 1. *Types of rolling bearing stressing*

Stressing of a rolling contact element is always the reaction to an external load acting on the rolling bearing; the load may be merely mechanical, but also thermal or chemical stressing may affect the bearing as well. For the introduction of the external loads into rolling contact elements under the other operating conditions (speed, lubricant type, lubricant amount), all design-specific and production-related possibilities (contact geometry, surface quality) must be considered in order to minimize internal material stress of the elements in rolling contact.

There are four major stressing types of elements in rolling contact (see Fig.6).

(1) Elastohydrodynamic lubrication

The appropriate coordination of bearing load, speed, lubricant viscosity, and surface quality of the contact areas result in a load carrying lubricating

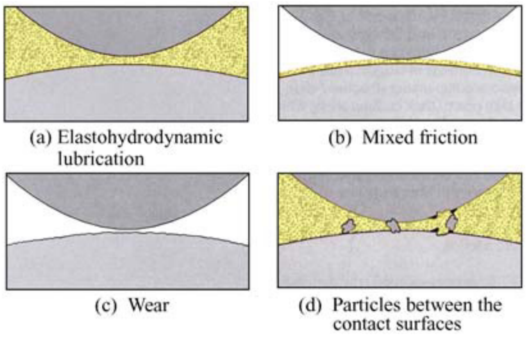

(a) Elastohydrodynamic lubrication
(b) Mixed friction
(c) Wear
(d) Particles between the contact surfaces

Fig.6 Stressing types in a rolling bearing.

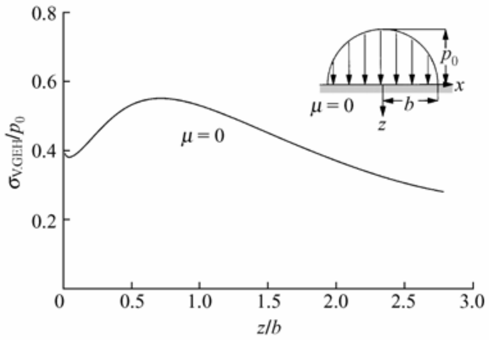
film in between elements in rolling contact which are then fully separated. Under these ideal elastohydrodynamic (EHD) lubrication conditions nearly all loads which are transmitted are normal loads. With the normal loads, the stress distribution in the material can be calculated (see Fig.7).

The depth profile of the equivalent stress which was calculated according to the distortion energy hypothesis, shows the peculiarity of this stressing condition: the maximum stress occurs at a certain depth below the surface. Under such stress conditions, rolling bearings reach the longest lives and even endurance life can be achieved. After very long running time, the life of a rolling bearing is limited by the formation of first subsurface cracks(see Fig.8), which typically progress to macroscopic pitting. Between the initiation and the final phenomena of fatigue, a number of fatigue-accompanying structural changes take place such as dark etching areas, white bands and "butterflies"[9]. Their locations and effects depend more or less strongly on the distribution of internal defects in the material and on the stress distribution within the rolling contact partners which varies in time and location.

(2)Mixed friction

While processes inside the material determine the length of life under EHD conditions, the processes near the surface are more important under the second type of stressing, the mixed friction (see Fig.6). If a load carrying lubricating film is not built up during operation, the roughness peaks of the contact areas will touch one another. This can happen with the stationary bearing, for example, due to an unsuitable lubricant, and also during the starting and stopping phases of a bearing otherwise running under EHD conditions during the steady-state operation.

In addition to the pure normal forces from the Hertzian pressure, additional tangential forces caused by mixed friction are superimposed at the surface. The changes of the surface topography are clearly visible under the SEM (see Fig.9). The calculation of the depth pattern of the equivalent stress shows that, as a consequence of the additional frictional stress, the maximum stressing is moved towards the surface (see Fig.10).



p_0—Contact pressure; μ—Friction coefficient; $\sigma_{V,\ GEH}$—Equivalent stress according to distortion energy hypothesis

Fig.7 Material stressing (according to distortion energy hypothesis) vs depth under contact pressure.

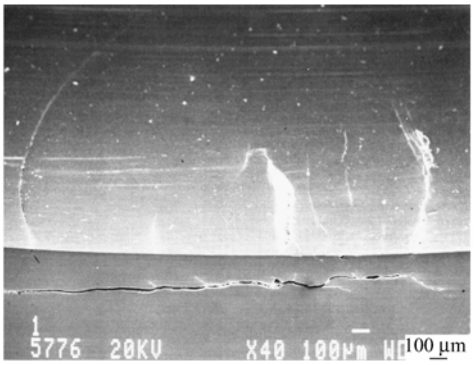


Fig.8 Scanning electron microscope (SEM) micrograph (crack below the surface).

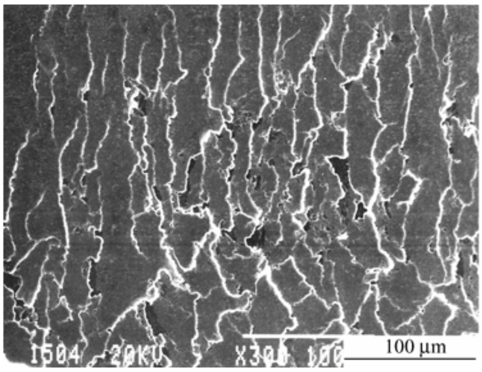


Fig.9 SEM micrograph (surface after mixed friction operation).

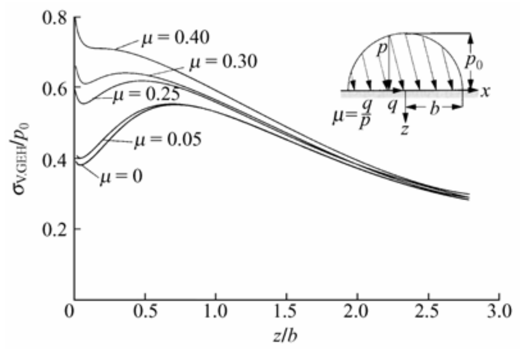


Fig.10 Material stressing(according to distortion energy hypothesis) vs depth under contact pressure for various friction coefficients.

(3) Wear

The deterioration of the lubricating conditions eventually leads to solid body friction in dry condition, namely wear (see Fig.6). Under these conditions, although rare, elements in rolling contact are damaged and put out of function due to wear. Although this condition does not represent a normal operating condition, in cases where short lubrication breakdowns may occur, some beneficial resistance against wear is tried to be provided by protective coatings or by the reaction layers formed by suitable lubricants and their EP additives.

(4) Particles in contact area

The risk of contaminants in the lubricating gap, that is to say, foreign particles at the contact area, is much more frequent than "dry friction". The contaminants can be mineral particles (mould sand, grit)or metallic ones(chips, grinding dust, abraded material). Their hardnesses vary correspondingly from "very soft" to "extremely hard", in relation to the hardnesses of the elements in rolling contact. The particles also vary widely in size, shape and distribution. Foreign particles which enter the bearing together with the lubricant damage the contact surfaces by indentations. The material edged up around the indentation penetrates the lubricant film and causes local mixed friction. The material around the indentation is plastically deformed and work-hardened; favorable residual stresses from heat treatment and/or machining are greatly disturbed.

The consequence of such damage is a bearing life reduction that can be drastic (see Fig.11)[10] . The larger the particles and consequently the indentations caused in the rolling contact area, the greater is the life reduction.

Under certain conditions cycled particles can get embedded in the surface of one of the parts in rolling contact, then act like a cutting tool and wear down the mating part abrasively[11] .

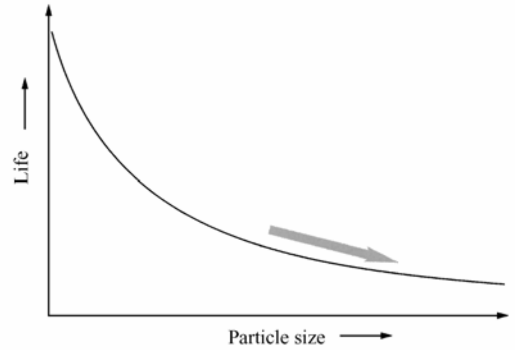


Fig.11 Effect of contamination particle size on life.

FAG has thoroughly investigated the fundamental correlations and uses the results for actual bearing designs, so as to achieve the optimum life capability even under such adverse operating conditions. Apart from the different stress types described, the following factors are taken into account for an optimized bearing design:

1) Thermal stressing in case where the bearing and lubricant are heated by an external heat source.

2) Chemical stressing due to corrosive surrounding media.

3) Additional mechanical stressing from extreme operating conditions (e.g., high hoop stresses from centrifugal forces in the rings caused by very high speeds, stresses from tight fits, etc.) or from integrating the rolling bearing function into a complex statically or dynamically stressed component (e.g., bearing deformations, bending or vibratory stressing, etc).

FAG always makes the greatest efforts to analyze, in close cooperation with the customer, the stresses as precisely as possible and to assess the requirements on material and specific material properties accordingly.

This is done with the help of theoretical means such as FEM calculations as well as with experimental means like material failure analysis.

4.2. Materials for rings and rolling elements

In addition to the technical criteria, also the availability and machinability (deformability, heat treatability) influence the material selection. For every application a solution is prepared which will serve the purpose best from technical and economical viewpoints. This may include even the use of special materials and special production methods.

By the example of the most commonly used rolling bearing steel 100Cr6, the following describes the demands to be made on the material and how the strength and fatigue characteristics can be optimized by the suitable heat treatment.

The specific stresses at the rolling contact area require high static and dynamic strengths which can only be achieved by material conditions at a certain minimum hardness. The fatigue behavior of such high-strength materials is decisively influenced by processes which take place at internal imperfections (nonmetallic inclusions, carbides, pores)[9] .

An important objective of material optimization is to reduce the number and size of these “internal notches” or to produce a positive effect on their distribution or minimize their risk potential. Fig.12 shows the adverse effects of nonmetallic inclusions in steel 100Cr6 on the rolling bearing life. It is illustrated by the number of attainable cycles versus the index of microscopic cleanliness determined for the individual heats. Based on the early discovery of this correlation, rolling bearing steel is the greatest challenge to the steel-making industry. The development of rolling bearings with better life and reliability capabilities required, among other measures, to constantly improve the melting processes.

With the introduction of new technologies, for example the secondary metallurgy, rolling bearing steel has developed in previous years into a material of supreme cleanliness. Under this precondition FAG could, for the first time, prove in tests that endurance strength applies also to rolling bearings, and that endurance life can be achieved (see Figs. 12-13).

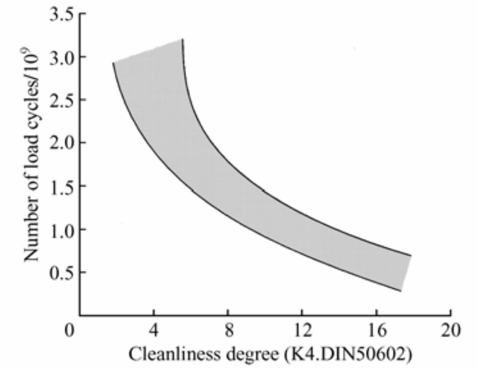


Fig.12 Development of life on steel cleanliness.

The special requirements of the rolling bearing steel have always been a pacemaker in the development of the steel-making industry and are the motor for the development of good quality and reliability in continuous casting, being one of the most modern steel making processes[1]. FAG is in continuous cooperation in order to monitor and further improve the quality of bearing steels. The approval of steel suppliers by FAG is product specific and it is subject to stringent and clearly defined criteria; it is part of a comprehensive quality management concept within FAG.

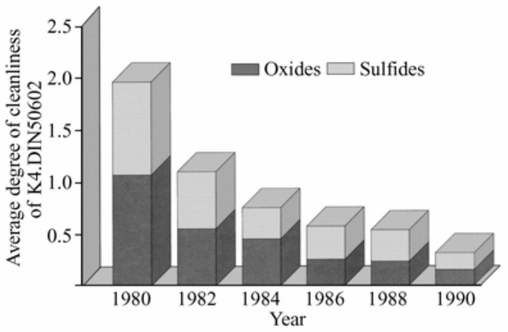


Fig.13 Development of degree of cleanliness of rolling bearing steel.

4.3. *Optimizing material properties by heat treatment*

If quality aspects have been included in the steel making process, the knowhow of FAG takes full effect during further processing, in particular in heat treatment.

Rolling bearing components are heat treated in the most modern, automatic furnaces with full process control. The most common procedure is martensitic hardening; the process parameters are adapted in such a way that the required high hardness is reached at the greatest possible plastic deformability.

(1) Strength and toughness properties

Martensitic heat treatment of through hardening chromium steel allows to achieve a wide range of specifically engineered material properties. Figs. 14-15 show these possibilities by the example of two typical rolling bearing part characteristics, namely the hardness and the retained austenite. By various combinations of the austenitizing and tempering temperatures, the same hardness values can be brought about as shown by the lines of identical hardness in the contour chart (see Fig.14).

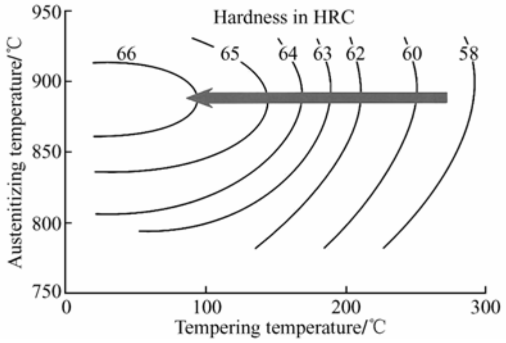


Fig.14 Hardness vs heat treatment (material:100Cr6).

The lines of same retained austenite content (see Fig.15) show the obvious trend which prohibits too high austenitizing temperatures for small retained austenite contents (which is a prerequisite for dimensionally stable bearings) and demands high tempering temperatures[12] . These interrelations, some of which contradict each other, are important to understand in order to optimize the bearing material and bearing performance characteristics.

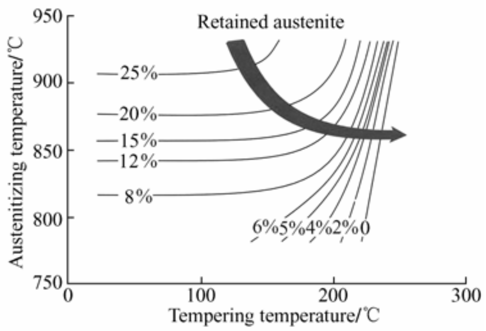


Fig.15 Retained austenite vs heat treatment (material: 100Cr6).

Hardness alone is not a sufficient criterion for maximizing the fatigue capability of a rolling bearing. Bearings which were tested under near-field conditions proved that within the wide range of 58 to 65 HRC the hardness has no significant effect on the life[13] . In the first approximation, material conditions with a very high resistance to plastic deformation offer the highest fatigue strength. The cyclic yield strength is the relevant material parameter, but for simplification reasons it can be replaced by the elasticity or microstrain yield limit for high-strength steels. Fig.16 shows the lines of same elasticity or microstrain yield limit for varying austenitizing and tempering temperatures[12] .

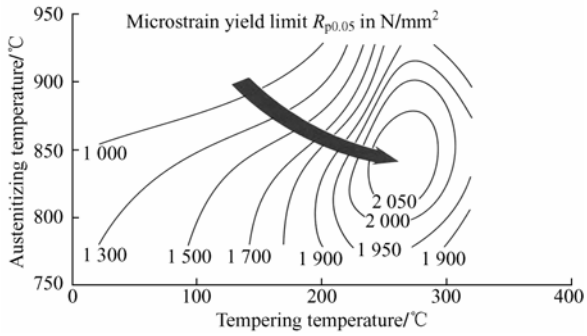


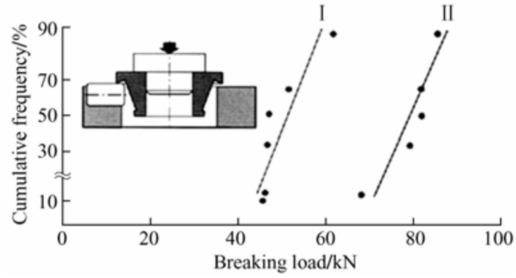
Fig.16 Microstrain yield limit vs various heat treat- ment parameters (material: 100Cr6).

A high cyclic strength delays the occurrence of plastic deformation which is the first stage of the fatigue process. Another material parameter, the plastic deformability, indicates the extent to which a material condition accommodates plastifications un- til first cracks form.

The optimizations of the raw material, the heat treatment parameters and consequently the final properties of bearing material are complex tasks, in particular because the fatigue processes do not take place in a flawless, homogeneous material matrix but are clearly localized in a material which is comprised of different structure types (martensite, retained austenite, carbides) and with nonmetallic inclusions as "internal notches"[9] . The results from FAG research provide the basis for finding the best compromises and a guideline between often contradictory requirements relative to the material condition, so as to achieve the overall best possible solution for a particular bearing application.

(2) Effects of residual stresses on component strength and life capability

The processes employed by FAG for heat treating rolling bearing steels, i.e., martensitic hardening, austempering and induction hardening of surface layers, lead to characteristic residual stress conditions in the components. The residual stress condition has a positive effect on the component strength and life capability (see Figs. 17-18).



I—Martensitic; II—Austempered

Fig.17 Static lip fracture test with a tapered roller bearing inner ring.

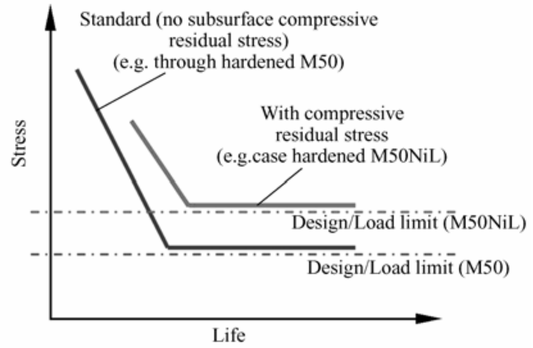


Fig.18 Effects of compressive residual stresses on bearing life.

(3)Dimensional stability of rolling bearings

Another material-dependent property for reliable rolling bearings is their dimensional stability. All efforts to minimize stressing by design and lubrication are useless if the components in rolling contact change their dimensions during operation. Loosening of the press fit on the shaft, clearance alterations or loss of preload may induce premature bearing failure.

Dimensional instabilities are caused by transformation of the retained austenite and martensite in hardened rolling bearing steel. Because of the effects of temperature and time, a reduction of retained austenite leads to an increase in volume, and carbide precipitation in the martensite leads to a decrease in volume. Both processes together result in the total dimensional change of a component (see Fig.19)[14] .

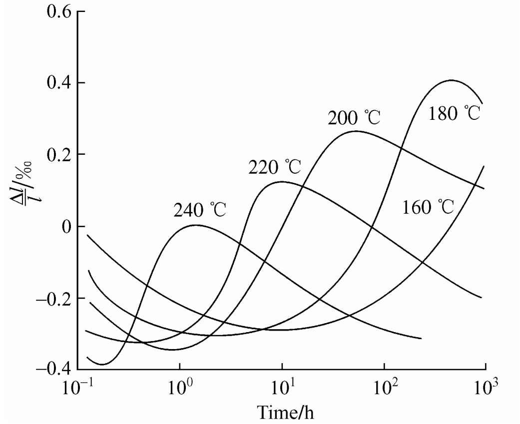


Fig.19 Dimensional changes depending on time and temperature.

Furthermore, the changes in volume described are stress-depending and follow the law of least constraint, that is, tensile stresses support dimensional growth (for example caused by a tight press fit on the shaft), compressive stresses stabilize retained austenite and force the dimensional change into that direction in space where compressive stresses do not act.

5. Rolling Contact Fatigue

Demands on rolling bearings are constantly increasing. The trend towards lower weight and low- or no-maintenance bearing designs, demands more reliability and longer life, which is a great challenge to rolling bearing makers. One very important step forward in meeting these demands is the utilization of all capacity reserves of the bearing material. Material selection and the microstructural state of the material must be optimized; performance and capacity which will be yielded by this optimization must be predetermined as accurately as possible.

All this is only feasible if there is a fundamental knowledge of the material behavior, origin and progress of damage under rolling contact stressing.

The evaluation of weak points in the hardened material, in particular nonmetallic inclusions, is very important for the increase of the endurance strength limit.

The basis for the investigations is the antagonism between the location of maximum material stressing according to the traditional theory of rolling contact fatigue and the origin of damage in practical cases. According to the accepted theory, the highest material stressing and thus the starting point of classic fatigue damage is at a certain depth below the surface. In practice, however, cracks are generated much closer to the surface or even directly at the surface.

5.1. *Material stressing under contact pressure*

The basic stressing in rolling contact is characterized by the Hertzian contact pressure. At merely elastic deformation, a semi-elliptical load distribution takes place in the contact zone of the two elements in rolling contact. The resulting stress distribution in the material below the contact areas is shown in Fig.20 by the direction and magnitude of the principal stresses. The centre of the contact area is chosen as the origin of coordinates, the y axis represents the direction of rolling, and the z axis, the distance from the material surface into the material. Depending on the y/z position, the principal stresses extend in other directions and have other magnitudes. This is illustrated in Fig.20 with a material section below the contact area. Under cyclic stressing, the contact area moves forward in the direction of the y axis, thus causing a change in magnitude and direction of the stress in every material zone.

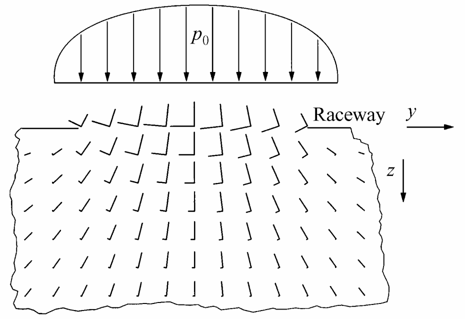


Fig.20 Main stresses under Hertzian contact.

Material stressing is rated best with the equivalent stress $\sigma_{\mathrm{V.Mises}}$ according to the V. Mises hypothesis. The equivalent stress distribution under the contact area belonging to Fig.20 is illustrated in Fig.21 by means of a three-dimensional chart and a contour chart. The axes of the chart $Y=y/b$ (horizontal distance from the contact area centre) and $Z=z/b$ (depth below the surface) are made dimensionless with the help of the small semi-axis b of the contact area. The equivalent stress is shown relative to the contact pressure p_0.

The maximum equivalent stress occurs at a depth of z/b=0.7; it amounts to 0.56p_0. At the surface, in the very centre of the contact area, the equivalent stress is only 0.4 times p_0. The reason can be illustrated with Fig.22. It shows the equivalent and principal stresses down into the material starting from the contact area centre (y=0).

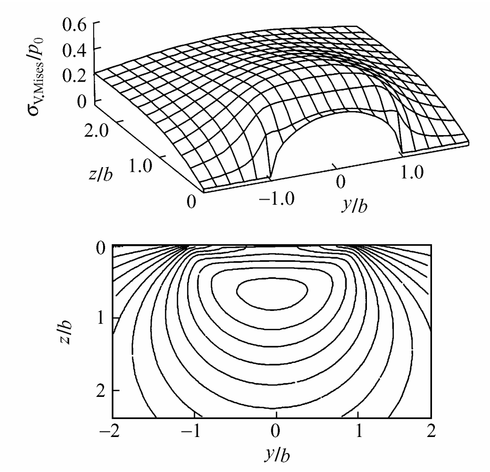


Fig.21 Equivalent stress distribution under contact area.

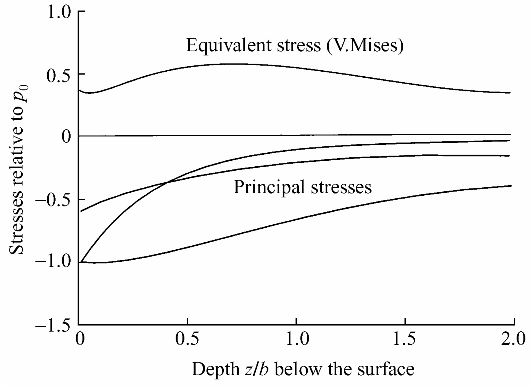


Fig.22 Main and equivalent stresses below contact area centre.

Two principal stresses at the surface reach the same magnitude, which results in the low equivalent stress at that area. Expressed in a simplified way, the equivalent stress reaches its maximum where the differences between the three principal stresses are the greatest. In the case described, this is at the above mentioned depth of z/b=0. 7.

A comparison between this theoretical statement and findings from actual rolling bearing failures shows that in many cases the damage starts near or at the surface even where the bearings are run under ideal EHD conditions and therefore with a very low coefficient of friction within the rolling contact area. An example of the origin of damage near the surface of a rolling bearing ring is shown in Fig.23. It can be clearly seen that the damage started at a weak microstructural point (inclusion) and progressed from there into the depth of the material until it reached the area of greatest equivalent stress(wide crack paths in Fig.23). Later another crack, a thin one, grew to the raceway surface (see Fig.23). Obviously, the material stressing calculable for the homogeneous material does not describe correctly the location and the magnitude of the highest possible stressing. Due to the main importance of these phenomena for judging the capacity, reliability and safety of rolling bearings, greatest efforts have been made by FAG to clarify this matter.

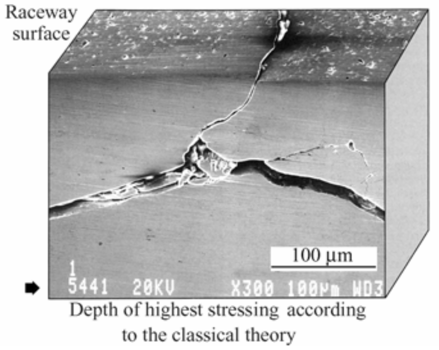


Fig.23 Damage propagating from an inclusion near the surface.

5. 2. *Effects of inclusions on rolling contact fatigue*

Inclusions in the steel cannot be entirely avoided in the steel producing process. It is technically not feasible to remove them completely, the extent of their reduction by specially developed processes has technical and economical limits. It is possible, at great expenditure, to reduce certain inclusion types to very small residues; however, apart from the higher cost, also other drawbacks must be accepted. The share of other inclusions, for example, may rise or a different composition of alloys may be required.

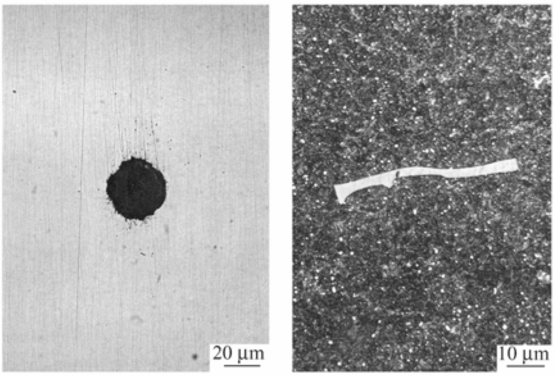
Fig.12 shows the importance of this problem. The number of load cycles depends on the type and number of inclusion types. The various inclusion types have various risk potentials. For a long time FAG has been playing a decisive part in classifying the inclusion types according to their risk potential. Fig.13 illustrates how the constant impulses from the R&D work of the rolling bearing manufacturers on the steel sector have improved the steel quality in recent years. The load rating increase of rolling element bearings in the past was made possible to a large extent by the improved quality of the steel.

A more precise evaluation of the performance and capacity of rolling bearings requires the microscopic analysis of the material stressing under consideration of the presence of nonmetallic inclusions. Their effects on the material strength vary widely. The properties of the inclusions are entirely different from those of the surrounding matrix. The phase boundary is usually incoherent and in positive contact with the matrix. There may be submicroscopically small cracks in the phase boundary area of nonmetallic inclusions. With some inclusion types the different thermal expansions lead to the formation of tensile stresses along the phase boundary. Also the strength and toughness properties of the inclusions have effects on the material strength. Which inclusion properties become particular important in operation depends largely on the operating conditions.

(1) Types of inclusions

Inclusions in steel can have the most widely varying properties. Fig.24(a), for example, shows a calcium aluminate in rolling bearing steel. The globular shape is characteristic of this inclusion type. A titanium carbonitride shown in Fig.24(b), is another important inclusion type. Such inclusions are rectangular to square and have sharp edges.

Table 1 shows the most important inclusion types in steel and their properties. According to their appearances in the micrograph they are classified into:1)streaky sulfides (SS), 2)accumulations of small oxides, produced by breaking large oxides in the rolling process (OA), 3)streaky oxides (OS), 4)globular oxides (OG), 5)sharp-edged rectangular to square titanium carbonitrides with the two basic forms titanium carbide (TiC)and titanium nitride (TiN).



(a) Calcium aluminate (b) Titanium carbonitride

Fig.24 Examples of inclusions in rolling bearing steel.

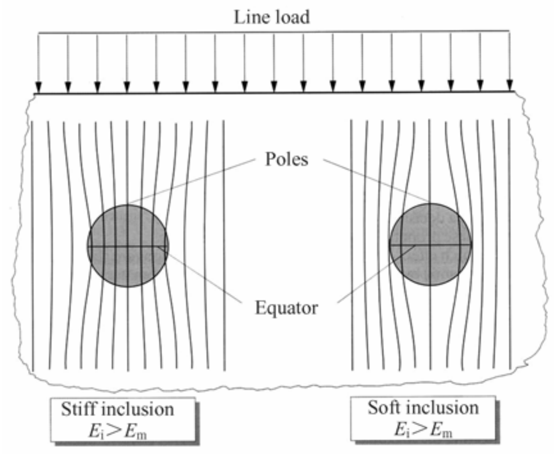
The inclusion types vary in their morphologies and their properties. The condition of the phase boundary, the hardness, brittleness, possible cooling residual stresses, and the internal notch effect are of particular importance for the rolling contact fatigue behavior.

(2) Internal "notch" effects due to disturbed flow of force

The varying elastic properties of the inclusion disturb the flow of force. A stiff inclusion, i. e. , an inclusion whose modulus of elasticity is higher than that of the matrix, supports the matrix because it attracts the flow of force. This does put more stress not only on the inclusion but also on the matrix in the zone where force flows into the inclusion. In the following this area is called pole (see Fig.25). At the pertinent equator the matrix is relieved from load.

Table 1 Important inclusion types in rolling bearing steel and their properties

Inclusion type	Examples	Morphology	Phase boundary	Microhardness	Brittleness	Cooling residual stresses	Notch effect
SS	MnS, CaS	Streaky	Positive contact, ncoherent	150 HV to 170 HV	Very low	None	Small
OA	Al_2O_3	Local accumulation of small foreign particles	Smallest microcracks possible	About 2 200 HV	Very high	Tensile stresses along phase boundary	Marked
OS	Mixed oxide systems	Oblong	Positive contact, incoherent		Marked	Marked tensile stresses	Small to medium
OG	Calcium aluminates	Globular	Frequently microcracks	900 HV to 2500 HV	High	Tensile stresses	Marked
TiC		Oblong to square, sharp-edged	Positive contact, ncoherent sometimes microcracks	About 3000 HV	Marked	Tensile stresses	High
TiN		Oblong to square, sharp-edged	Positive contact, incoherent, sometimes microcracks	About 2500 HV	Marked	Low tensile stresses	Small to medium



E_i—Elastic modulus inclusion;
E_m—Elastic modulus matrix

Fig.25 Flow of force disturbed by inclusions.

With a soft inclusion, with a smaller modulus of elasticity than the matrix, some of the forces by pass the inclusion (see Fig.25, left). The matrix is relieved at the pole area and more loaded at the equator than would be the case without an inclusion.

Even at an external, uniaxial load, very inhomogeneous stress conditions prevail around inclusions. Fig.26 illustrates the complicated stress conditions near inclusions. It shows the photoelastic examination of an inclusion model. The model was loaded in the vertical direction with a constant line load. The complex and inhomogeneous stress conditions in the matrix near the inclusion can be clearly seen without evaluating the picture in detail. The inclusion model itself has by far not such a great inhomogeneity in its stress state.

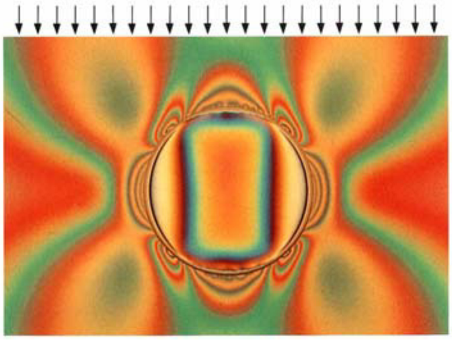


Fig.26 Photoelastic presentation of stress condition around an inclusion model.

Apart from the photoelastic analysis, the effects of inclusions can also be assessed with finite element method(FEM) calculations and in special cases with analytical calculations. FAG makes use of these possibilities to assess the effects of inclusions in areas under contact pressure. Fig.27, for example, shows the pattern of the two principal stresses in the y/z plane and the pattern of the equivalent stress (V. Mises)for an inclusion which is located at a depth of z/b=0.24 below the centre of the contact area. The principal stress acting in the x direction was not entered in the diagram for the sake of clearness; however, it was taken into account for determining the equivalent stress. The missing principal stress can be easily estimated from the condition of plane strain with the help of the two plotted curves.

The stress patterns in the undisturbed matrix, plotted in full in Fig.22, are shown as dashed curves for better comparison. It can be seen how the inclusion changes the principal stress pattern. The equivalent stress is increased in the inclusion as well as in the matrix near the inclusion. It is remarkable that the increase of the equivalent stress is higher in the matrix than in the inclusion itself.

Fig.27 suggests why most cracks are generated in the matrix near the surface. In the undisturbed matrix the equivalent stress at the surface is lower than at a certain depth due to the nearly identical two principal stresses in this area. If there are inclusions in this area this even pattern is disturbed in such a way that the major differences between the principal stresses become greater. In the matrix near the inclusion, the lower stress decreases somewhat while the larger stress increases clearly.

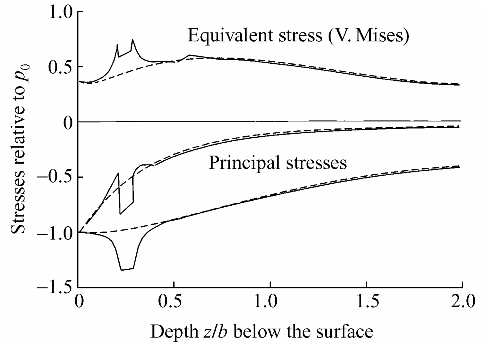


Fig.27 Main and equivalent stresses at an inclusion in a depth of z/b=0.24.

It is now logical to ask for the depth below the surface at which the inclusions have the most dangerous effect.

The answer is complex. All important types of inclusions have to be determined, taking into account their properties and the maximum equivalent stress in the matrix depending on their positions relative to the centre of the contact area. FAG was greatly interested in answering this question due to its importance for understanding the rolling contact fatigue, for a further optimization of the rolling bearing materials, and for the increase of bearing performance and capacity.

FAG worked out, in extensive investigations, that inclusions are particularly dangerous near the surface, and largely independent of their properties. The critical inclusions are usually not directly at the surface, but they are clearly closer to it than assumed by the classical "macroscopic" theory. It is insignificant whether the flow of force is disturbed by stiff or soft inclusions. In any case the equality of the two highest principal stresses near the surface without inclusion is disturbed by an inclusion.

The notch effects of different inclusions may vary widely, of course. Here, the elastic properties of the inclusions and in particular their hardness and brittleness play a decisive role. It depends on these properties whether the internal notch effect caused by an inclusion is softened by plastic deformation or whether a crack is generated inside or near the inclusion.

Fig.28 illustrates the risk potential of a globular Al_2O_3 depending on its depth z/b below the contact area centre (z=0). The equivalent stress curves in the matrix are shown for three identical inclusions located in different depths.

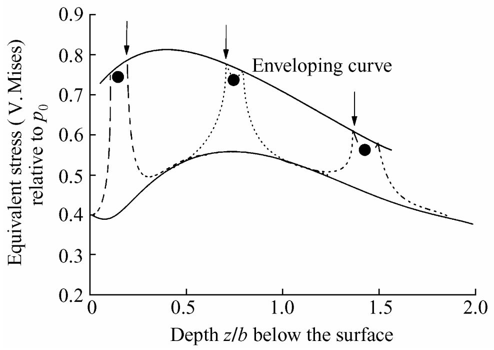


Fig.28 Risk potentials of inclusions depending on their location.

By connecting the maxima of equivalent stresses of all inclusion locations, which are marked with arrows with an enveloping curve, the risk potential of an inclusion depending on its depth can be clearly seen. The maximum equivalent stress for the inclusion type considered occurs at a depth z/b of about 0.4; the maximum value without inclusions is at z/b=0.7 (see Fig.22). A comparison of the enveloping curve and the equivalent stress curve in the inclusion-free material shows that the distance between the two curves is at its greatest just below the surface. This shows clearly at which depth inclusions have the most negative impact, hence the worst effect.

Since inclusions are statistically distributed in the material, some of them will be always found in the critical zone near the surface. First cracks will, of course, occur in this zone[15] .

4. Summary

Rolling element bearings can achieve endurance life if bearing stressing, lubrication and system cleanliness are adequately taken into consideration. The FAG research work of the early 1980's and the advanced life calculation method meanwhile adopted according to these findings present valuable knowledge and a practical methodology to the bearing engineers for designing a system that provides a long-time and reliable performance in the application.

Operation of rolling bearings in die mixed friction range cannot be prevented in practice. In order to eliminate life reducing effects, the lubricant plays a major part in this field.

Investigations have shown that with an appropriate adjustment of the tribological system consisting of bearing, lubricant, operating and environmental conditions including the bearing temperature, a long bearing life can be achieved reliably.

Rolling bearings prove their reliability under most diverse operating conditions in many applications. With the range of materials and special treatment processes, operating temperatures between -200℃ to 600 ºC can be accommodated.

The importance of nonmetallic inclusions in rolling bearing steel has been demonstrated. The effects of inclusions on the endurance strength limit under cyclic stressing are much greater than their effect on the life for operation in the fatigue strength region, i.e., for rolling bearings designed for a limited time of operation. The stress increasing and strength reducing effects of nonmetallic inclusions must be taken into account for assessing rolling bearing materials capabilities. The local strength, i.e., the weakest point in the material, is decisive for service life and endurance strength. Widely varying inclusion types displace the origin of damage under rolling contact fatigue towards the surface, which is different from the classical "macroscopic" theory of the contact stressing.

FAG Aerospace has developed bearing technologies, design solutions and a broad fundamental bearing know-how in order to be able to support the customers and to meet the very demanding application requirements in aircraft engines, helicopters, turbopumps and many other aerospace applications.

References

[1] Ebert F J. An overview of performance characteristics, experiences and trends of aerospace engine bearings technologies. Chinese Journal of Aeronautics 2007; 20(4):378-384.

[2] Loroesch H K. The life of the rolling bearing under

varying loads and environmental conditions. Schweinfurt, Germany:FAG Kugelfischer, FAG Publication No. EA1981. 1.

[3] Loroesch H K. Einfluss von festen Verunreinigungen auf die Lebensdauer von Waelzlagern. Antriebstechnik 1984; 23(10):63-69. [in Germany]

[4] FAG. A practical method of calculating the attainable life in aerospace bearing applications. Schweinfurt, Germany:FAG Kugelfischer, FAG Publication No. FL 40 134 EA.

[5] Ebert F J, Poulin P. The effect of cleanliness on the attainable bearing life in aerospace applications. STLE Tribology Transactions 1995; 38(4):851-856.

[6] Chen G. The analysis of aero-engine mechnical design. Beijing:Beijing University of Aeronautics and Astronautics Press, 2006. [in Chinese]

[7] von Walzlagern S. FAG publication no. WL 81 115/3DA.

[8] Broszeit K. Verschleiβschäden durch Oberflächenermüdung. VDI-Berichte 1975;243:189-204.[in Germany]

[9] Schlicht H, Schreiber E, Zwirlein O. Ermüdung bei Walzlagern und ihre Beeinflussung durch Werkstoffeigenschaften. Antriebstechnik 1987; 26:49-54. [in Germany]

[10] Lorosch H K. Einfluβ von festen Verunreinigun-gen auf die Lebensdauer von Wälzlagern. Antriebstechnik 1984; 23:63-69. [in Germany]

[11] Lorosch H K. Wärmebehandlung minimiert Lagerverschleiβ trotz Verunreinigungen. Der Konstrukteur S. 5/88, 110-116. [in Germany]

[12] Stangner H. Werkstoff für Wälzlager-Festigkeitsverhalten, Auswahlkriterien. VDI-Berichte 1986; 600(1):421-438.

[13] Zwirlein O, Schlicht H. Rolling contact fatigue mechanism:accerlerated testing versus field performance. In:Hoo J ed. Rolling Contact Fatigue Testing of Bearing Steels. ASTM STP 771, 1982; 358-379.

[14] Schlicht H. Material properites adapted to the actual stressing in a rolling bearing. Ball and Roller Bearing Engineering 1981(1):24-29.

[15] Boehmer H J. Rolling contact fatigue. Schweinfurt, Germany:FAG Kugelfischer, FAG Publication No. WL 40 205 EA.