Comparing Surface Failure Modes in Bearings and Gears: Appearances vs. Mechanisms

Charles A. Moyer
The Timken Company, Canton, OH

Table 1 - Damage Classification For Bearings

I. FATIGUE

Contact Fatigue - Subsurface Origin

- Inclusion - with macroshear classical shear stress zone below contact surface
- Inclusion - near surface zone of microshear greatly influenced by surface roughness (asperities)
- Subcase fatigue - origin near case-core interface if yield strength is exceeded by applied stress

Contact Fatigue - Surface Origin

- Origin at the end of contact aggravated by edge geometry
- Multiple origins of micropitting (peeling or frosting)
- Point surface origin - at localized stress risers (dents, grooves, surface inclusions)

II. PLASTIC FLOW

- Brinelling or debris denting
- Load excursions above the material yield point
- Yielding aggravated by high temperature excursions

III. WEAR

A. Adhesive
- Normal - mild or "controlled" - usually identified as "run-in"
- Severe - irreversible scuffing, scoring, smearing, or seizure

B. Abrasive
- Normal - usually 3 body system, medium to fine particles that are also associated with "run-in"
- Severe - grooving, gouging, denting with ridges that cause serious surface stress risers

C. Corrosive
- Water or acidic constituents from lubricant breakdown or temperature chemically aggressive additives in lubricant

D. Fretting
- Microscale adhesive and abrasive wear
- Corrosion involvement depending on environment and contacts

Introduction

In the 1960's and early 1970's, considerable work was done to identify the various modes of damage that ended the lives of rolling element bearings.(1-5) A simple summary of all the damage modes that could lead to failure is given in Table 1. In bearing applications that have insufficient or improper lubricant, or have contaminants (water, solid particles) or poor sealing, failures, such as excessive wear or vibration or corrosion, may occur, rather than contact fatigue. Usually other components in the overall system besides bearings also suffer. Over the years, builders of transmissions, axles, and gear boxes that comprise such systems have understood the need to improve the operating environment within such units, so that some system life improvements have taken place.

Those of us who manufacture bearings realized that identifying the damage modes was not enough, but that an understanding of the causes and underlying mechanisms at play was important to improve the ability to predict bearing life in this variety of systems. However, whenever adverse operating conditions prevailed, and actual life was below expected life, the understanding of modes of failure within bearings could be used to determine what improvements were needed to allow bearings to have extended life even under adverse conditions.

Considerable work has been done to accomplish this, and more remains to be done, but through understanding expected failure modes based on known operating conditions and applying this knowledge to the design cycle, significant improvements in bearing performance have been achieved. For line contact bearings, the prime factors are (1) significantly cleaner steel, (2) much improved surface finishes, which include roughness, waviness, and other factors beyond R_a (the arithmetic center-line average of roughness), and (3) geometry that optimizes internal stress profiles, so that even under high-load or nonaligned conditions much more uniform stresses can be maintained along the contact line.

The same development path seems to have been occurring for gear contacts. Thus, we can review the failure modes that occur in gears and, based on the understanding of what mechanisms underlie the various identified failure modes, see what design changes may be appropriate that would also extend gear performance life beyond present limits.

One way to do this is to compare the modes of damage identified in line contact bearings and gears to determine what is happening within their respective contacts that cause the final failures. Since both bearings and gears function primarily with fluid lubrication, a tribological model of line contact, as shown in Fig. 1, can be the starting point to discuss the mechanisms that contribute to the various failure modes reported.

The Basic Line Contact Model

Whether gears or roller bearings, the contacting surfaces can be represented as two cylinders. It is possible to include transverse profiles or radii on these cylinders and determine the appropriate elliptical or truncated elliptical contact formed; but for this discussion two simple cylinders of the same or different radii are sufficient. Under running conditions, each cylinder has surface velocities either close to the same or considerably different that result in tangential or frictional forces within the contact that can range from less than 1% to well over 20% of the normal force that forms the rectangular contact area between the two cylinders.

The two cylinders under load determine the Hertzian contact area on the X-Y plane shown on Fig. 1. The cylinder length along the Y axis and 2b, the width of contact along the X axis, form the Hertzian contact shown by the straight dashed lines. The Hertzian contact pressure reaches a

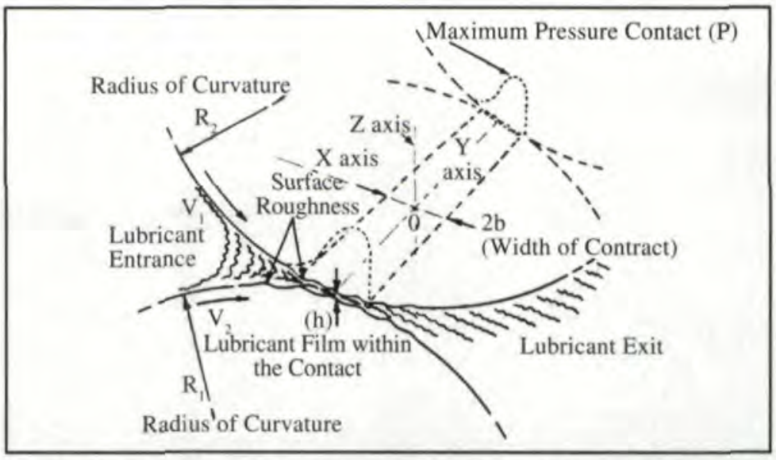

Fig. 1 - The tribology model of line contact. Two cylinders in contact (R_1 and R_2) under load with surfaces moving (V_1 and V_2). X axis (in the direction of motion) and Y axis (at right angles to motion) define the contact plane. The Z axis at right angles to the contact plane refers to the distances into the contacting cylinders. The contact length on the Y axis and the contact width 2b define the Hertzian stress contact area.

maximum P as indicated by the dotted lines above the contact rectangle. P directly relates to the subsurface principal and shear stresses to be discussed later.

If there is sufficient lubricant in the entering meniscus to the contact, an elastohydrodynamic lubricant (EHL) film (h) will be formed. Recognizing that the cylindrical surfaces have some amount of roughness, the lubricant film (h) and roughness (σ) can be compared to give an indication of the quality of the lubricant operating regime within the contact; that is, the EHL film (h) divided by the composite roughness of the two surfaces (σ) provide h/σ or lambda (λ).

Testing on ball and roller contacts, as for example in a tapered roller bearing,[6] demonstrates that the contact itself (2b) actually operates as a functional filter, and only the surface roughness wavelengths that are approximately one-fourth 2b up to twice 2b are active within the contact. Considering gear contacts, Wellauer and Holloway understood this in their paper on the application of EHL lubricant films to gears presented in 1975.[7] A simple approach to address such a concept was presented as a modified lambda ratio in 1989.[8] The concept was included in a paper at the 1989 AGMA Fall Technical meeting[9] and will be in the next revision of the ANSI/AGMA Standard 2001-B88, Appendix A.

The modified lambda can be expressed as:

$$\lambda m = \underline{h}\ \_\ \frac{L}{2b}^{1/2} \qquad (1)$$

where λm = the modified lambda

Charles A. Moyer
was employed at The Timken Company for 38 years prior to his retirement. His last post there was as Senior Scientist. He is the author of over 30 articles on bearing technology.

L = large end wavelength cutoff used to measure surface roughness [usually L = 0.8mm (.030 inch)]

2b = width of contact in the direction of motion

Any further discussion of lambda, especially as used later in Table 3, is primarily based on λm rather than λ not modified.

Under load the two cylinders, besides producing the contact stress profile shown in Fig. 1, also produce stresses below the surfaces. Of most interest are some specific shear stresses formed from the principal stresses σ_x, σ_y, σ_z and the shear stresses T_{xz}, T_{yz}, and T_{xy} that occur within the stressed volume below the Hertzian contact area. The first of these is the maximum shear stress (T_{45}) that is in the center of contact:

$$T_{45} = 0.5\ (\sigma_z - \sigma_x) \qquad (2)$$

where T_{45} = shear stress on 45° plane from the contact surface in the rolling direction and at x (or b) = 0.

σ_x = principal stress in the x direction (Fig. 1).

σ_z= principal stress in z direction (Fig.1).

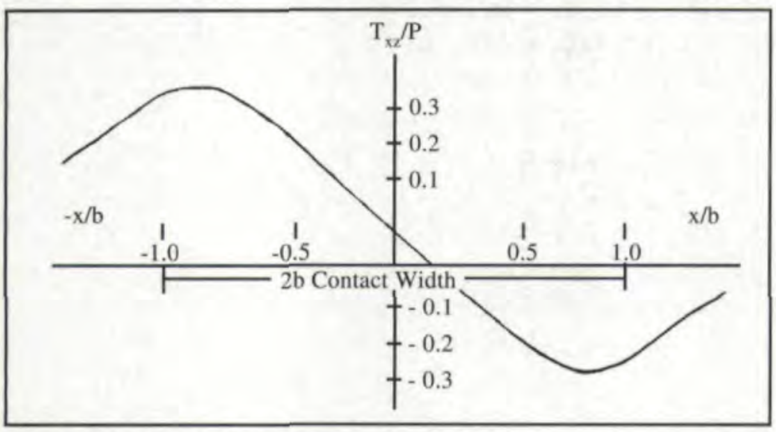

Fig. 2 - The range of orthogonal shear stress T_{XZ} as a fraction of the Hertzian contact pressure P. The values go from a maximum of -0.25P through zero to +0.25P at 0.5b below the contact surface for pure rolling.[10]

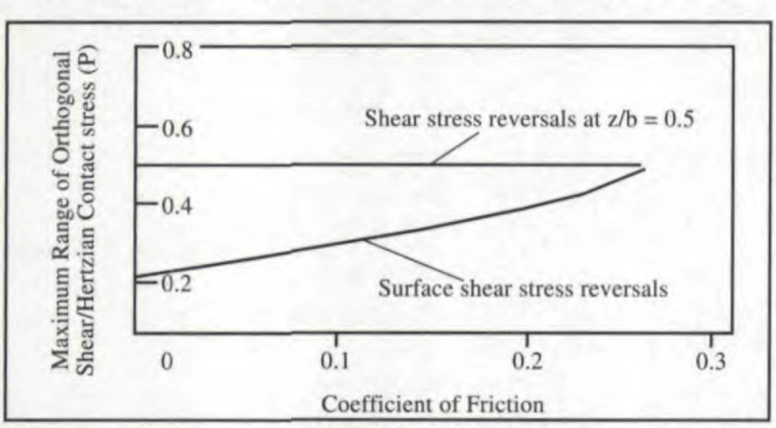

Fig. 3 - Influence of the coefficient of friction on the maximum range of orthogonal shear stress in terms of the ratio $|2T_{XZ}|/P$ where P is the maximum Hertzian contact stress for line contact.[10]

The maximum of T_{45} occurs on the center line of the rectangular contact (b = o) and below the surface at 0.78b, where b is one-half the width of contact. Under rolling conditions, T_{45} is equal to 0.3P (P equals maximum Hertzian pressure in Fig. 1).

While T_{45} seems to relate to the yield limit of stressed material, the orthogonal shear stress, usually T_{xz} (in the direction of motion) seems to be related to rolling or rolling/sliding contact fatigue. Of interest here is the range of orthogonal shear that occurs under a contact area as shown in Fig. 2. This figure is based on the results published by Kannel and Tevaarwerk [10] and shows for pure rolling that the maximum T_{xz} occurs about 0.5b below the contact surface and has opposite values across the contact of ±0.25P; that is, the range of shear stress is 0.5P, and this magnitude along with the depth below surface is strongly related to fatigue. There is not space here to develop the equations used to calculate T_{xz}, but the reader is directed to Refs. 10 and 11 for further details.

Another equivalent stress considered with contact fatigue is the von Mises yield parameter (σ Mises) that is represented by:

$$\sigma\ \text{Mises} = [T_{xy}{}^2 + T_{yz}{}^2 + T_{xz}{}^2 + 1/6\ \{(\sigma_x - \sigma_y)^2 + (\sigma_x - \sigma_z)^2 + (\sigma_y = \sigma_z)^2\}]^{1/2} \qquad (3)$$

where all stresses are the principal or shear stresses on or below the contact surface in either the direction of motion (x-axis) or in the mutually perpendicular directions (y-axis and z-axis).

These three stresses, along with a few others, have been used to evaluate what stress best relates to contact fatigue and the onset of plastic deformation.

When the surface velocity difference between the velocities of the two cylinders is low (low-sliding), and the surfaces are separated by a lubricant film ($\lambda m > 1.0$), the shear stress will be below the surface and contribute only to material related fatigue. As the velocity difference increases (larger slide/roll ratio), and/or with considerable surface contact ($\lambda m < 1.0$), the shear stresses are larger (contribution from higher frictional forces) and, as an example, the orthogonal shear stress can extend up to the surface. This is shown in Fig. 3, taken from Kannel and Tevaarwerk,[10] and indicates how the maximum shear stress reversals in the surface increase with friction force increase. This means that from the increased friction a larger volume of material below the surface sees significant shear stresses.

Surface roughness, besides contributing to the λm calculation, determines the actual contact area that occurs in comparison to the contact calculated from Hertzian theory. Surface roughness is very often represented in two dimensions as a field of asperities, and under load only a finite number are in contact. As the load increases more contacts occur, and some existing contacts grow larger, but the actual contact area never reaches the full theoretical contact. The run-in that occurs for bearings and gears includes the portion of these asperities in contact that are plastically deformed and worn away, depending on the amount of tangential force (sliding) that accompanies the normal load in the contact. It is clear that the size of these asperities, their side slopes, hardness, and the lubricant film that tries to separate the two rough surfaces, all contribute to run-in and to the mode of damage the surfaces may eventually see.

A real surface can be represented by a string of asperities in a line. These multiple contacts form micro-contact stresses and subsurface shear stresses in miniature below each loaded asperity so that, depending on λm, shear stresses may be formed below the surface. Figs. 4 & 5 provide representations of real surfaces that are stressed under two different λm values. Besides the variation of contact stresses that deviate from the Hertzian stress calculated assuming ideal surfaces (zero roughness), these figures show the subsurface shear stresses and the shear stresses in miniature below the surface as modified by the roughness, and illustrate the changes that can occur with a change in λm.

These figures come from the work of Zhou and Cheng[12] based on the actual surface finish on specimens in a two-disk machine described in Ref. 12. They illustrate the subsurface and near surface shear stresses that must be included in any analysis of the basic line contact model. They have used the von Mises stress, and this stress illustrates the influence of the surface asperities as they are able to penetrate the lubricant film. For λm greater than 3.0, the maximum von Mises stress is 620 MPA (90 KSI) and located 338 μm (0.013") below the contact surface. For λm equal to 0.33, the maximum von Mises stress is 1450 MPA (210 KSI) and only 40 μm (0.0016") below

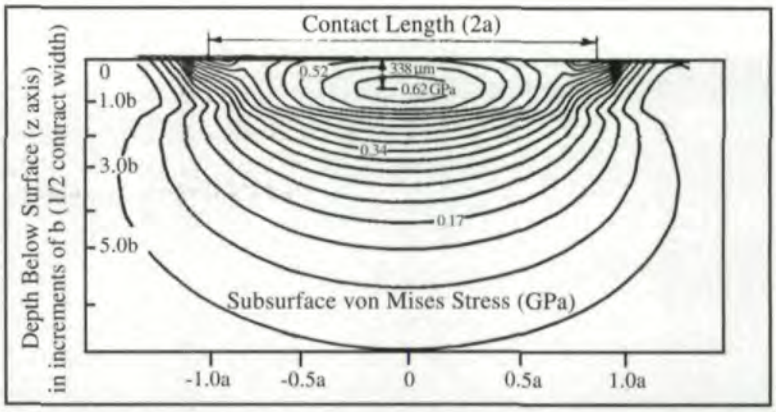

Fig. 4 - Contours of von Mises stress below the contact. Hertzian contact pressure 2.3GPa, lambda ratio above 3.0. Maximum stress is 0.62GPa, occurs 338 μm below the surface or 0.78b (width of contact) below the surface showing no influence from surface asperities.[12]

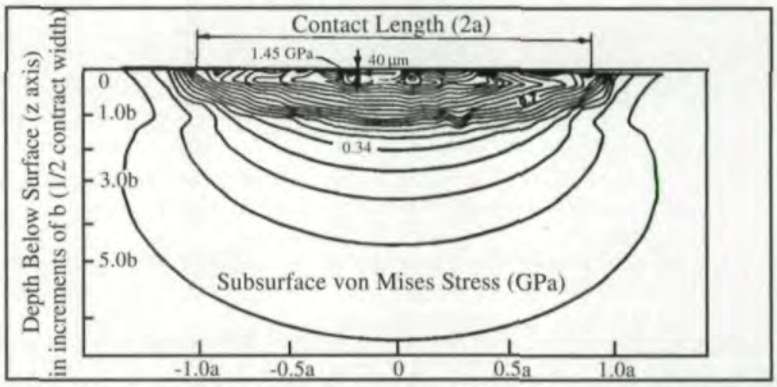

Fig. 5 - Contours of von Mises stress below the contact. Hertzian contact pressure 2.3GPa, lambda ratio 0.33. Maximum stress is 1.45GPa, only 40 μm below the surface or 0.09b below the surface showing considerable asperity interaction.[12]

the surface. All these aspects of the line contact model must be kept in mind as we discuss the fundamentals of wear mechanisms.

The Basic Wear Mechanisms

Within the context of the line contact model, wear can be considered in its broadest terms. As expressed in an excellent review paper on "Wear and Wear Mechanisms" by Vingsbo,[13] wear can be defined as the removal of material from surfaces in a tribology system. And basic to this definition is that the fundamental phenomenon in material removal is fracture. Vingsbo considers four mechanisms.

The first is abrasive wear, termed ploughing or microcutting, that separates a fragment or chip from the main body. Also associated with this mechanism are various amounts of plastic deformation. Abrasive wear also assumes that one surface has harder asperities than the surface it abrades, or that hard particles (three body system) are moving across the surface. This mechanism implies the presence of slip or considerable sliding between the surfaces. In steels, with which we are primarily concerned here, a large portion of energy expended is used in the plastic deformation and primarily forms ridges along the furrow sides.

The second mechanism is adhesive wear or shearing of junctions. This considers asperity-asperity contact or asperity-surface contact, depending on the relative surface finishes. Again considerable plastic strain and frictional work from the localized slip (hot spots) or sliding takes place. Obviously, the energy into the points or areas of contact dissipates as heat that helps remove oxide layers or contaminants on the surfaces. This then allows the asperities to form intimate contact or "weld." If the binding energy of the join (weld) is stronger than material away from the surface contact, fracture will take place within the adjacent material. This causes material transfer, and with continued plastic flow from sliding, the transfer areas may break loose and form third body fragments.

Table 2 - Damage Classification for Gears[18]

I. (4.) SURFACE FATIGUE

(4.1) Pitting - Initial (0.4-0.5mm size pits) to destructive (considerably larger) usually takes place in the dedendum (negative sliding).

(4.2) Spalling - Larger irregular shaped spalls, but quite shallow compared to destructive pitting.

(4.3) Case Crushing - Subsurface fatigue failure just below hardened case with longitudinal cracks ending on the surface.

II. (5.) PLASTIC FLOW

(5.1) Cold Flow - Extreme surface deformation from heavily loaded contacts - appearance of extrusion or peening.

(5.2) Rippling - Periodic, wave-like deformation with fish scale appearance usually seen on hardened gear surfaces.

(5.3) Ridging - Deep ridges from plastic deformation from high stresses.
(-) Indentation - Deep plastic flow from rolling, brusing, peening, or brinelling.[26]

III. (3.) WEAR

(3.1) Polishing - Fairly uniform mild wear - associated with run-in process

(3.2) Moderate Wear - Visible contact pattern - more material loss than in polishing.

(3.3) Excessive Wear - Pronounced wear with sufficient material loss to degrade gear design life.

(3.4) Abrasive Wear - Lapped, radial scratch or grooving on tooth contact surface from hard or metallic debris particles.

(3.5) Corrosive Wear - Surface deterioration from chemical action from liquid contaminants, lubricant breakdown, acid products, or moisture.

(3.6) Frosting - Etched appearance, micropitting with field of micropits 2.5m deep (0.0001 in.).

(3.7) Scoring - From moderate to destructive or localized, advanced occurrence of adhesive wear, alternate welding and tearing, ad vanced frosting, or significant material loss.

The third mechanism is fatigue. It is normally found in high-cycle stress applications, usually with stresses below the yield point of the material. Thus, fatigue seems to lack the extended plastic strains that occur during the operation of the first two mechanisms. However, there is ample evidence that within bearings, especially those run at higher loading, there is microstructural alteration within the stressed volume due to cyclic stressing.[14] At the stress levels that cause fatigue, evidence of plastic deformation occurs in the highest stressed zone and around localized stress risers, such as nonmetallic inclusions. Fatigue can occur without sliding (without significant frictional force) and cracks can nucleate below the contact surface at either the subsurface shear stress regions or below the asperities at miniature shear stress areas as noted in Figs. 4 and 5, and in contribution with subsurface stress risers for either shear stress locations. These stress risers and those on the surface all contribute to the statistical nature of contact fatigue; it tends to follow a different life distribution than wear.

The fourth mechanism Vingsbo considers in his paper is impact, which may have some relevance here; however, most impact fractures tend to occur in brittle material, so it will not be considered further.

The underlying phenomenon that ties these wear mechanisms together is material fracture. The factors that influence the wear rate and wear severity in different degrees for each of these mechanisms are the forces involved, quality of lubricant, ranging from full separating film to boundary lubrication, surface roughness, surface velocity in terms of rolling speed and sliding, and the surrounding environment that inhibits or promotes surface absorption, oxide formation, and other chemical reactions.

Besides material loss (wear) and debris or particle generation, these factors also influence the friction levels, heat, and temperature, which can alter the surface and subsurface areas. The additional changes that take place can directly influence the failure modes, especially for steels.

Microstructural changes of the material can occur just below the contacting surface, and even compositional changes can occur as heat input enhances diffusion into the near surface material. Probably the greatest changes occur from the plastic deformation that accompanies almost all of these mechanisms and also modifies the near surface material structure leading to the microcracks that can be the starting point of surface/subsurface degradation.

Table 3 - Relating Contact Fatigue Damage Mode to Modified Lambda

Modified Lambda	Contact Fatigue Mode (Initiation)	Material Influence	Surface Roughness Influence	Geometry Influence
Ratio > 3.0	Subsurface Fatigue Inclusion(1)	Important	Minor	Important
Ratio 3-1.0	Subsurface/Near Surface Mostly Inclusion	Important	"Sharp"/High Asperities Important	Important
Ratio 1-0.3	Some Inclusion, Some Surface Related	Somewhat Important	Important for Surface & Near-Surface Fatigue	Somewhat Important
Ratio < 0.3	Surface Related(2)	Minor	Important	Less Important
Any Ratio	Localized Stress Risers	Mixed	Mixed	Mixed
	(PSO(3) - dents, grooves, and surface inclusions)	Minor	Somewhat Important	Minor
	(GSC(4) - edge fatigue misalignment)	Somewhat Important	Minor	Important

(1) Fatigue originates at non-metallic inclusion in the maximum shear zone below the surface for both bearings and gears.

(2) Called peeling or micorpitting for bearings[4]; spalling for gears.[18]

(3) PSO - point surface origin - fan-shaped spall propagation starting on the surface.

(4) GSC - geometric stress concentration starting at end of line contact.

Observations of Structural/ Mechanisms Interaction

In recent years, considerable work has been done to understand the changes occurring within the material from repeated stressing from surface contacts. Some of the results indicated that with significant tangential forces present, the mechanism leading to wear particle formation was less adhesion-related, but primarily due to plastic deformation and fracture under a delamination mechanism.[14-15]

This concept considers that sliding surfaces go from asperity-asperity contact with adhesion, ploughing, and deformation in the contacts to form larger actual contact areas that continue the subsurface deformation until cracks nucleate somewhat parallel to the surface. As the cracks propagate, some migrate to the surface (or possibly start from the surface), then flat, thin, plate-like wear particles are formed. Although there is not universal agreement to the delamination theory,[16] there is certainly evidence that very close to the surface a deformed, "fragmental" layer forms, and below a layer of deformation microstructure forms that is "cellular," smaller grained, and relatively free of dislocations.[16] Some excellent work has also been done to relate the significance of plastic deformation of the surface layers to actual gear pitting.[17]

Cummins and Doyle[17] ran annealed and case-hardened gears under conditions to produce surface distress and concluded that some of the debris formed during running-in consisted of submicron wear particles that by themselves would produce only mild wear and contribute to the mutual polishing of the gear surfaces. In addition, they found larger platelets and wedge-shaped particles that were formed after the run-in when larger areas of gear surfaces were in contact. In careful examination of the wear surfaces and the debris particles, as observed in both scanning electron and transmission electron micrographs, it was clear that the fine debris came from surface layers that had been significantly deformed and consisted of very fine grain material. The large particles were primarily thin platelets or larger flat pieces with wedge-shaped undersides. Based on the pits on the gear surface, the thin platelets come from the shallow pits lacking fracture-type bottoms, and the wedge pieces seemed to come from the deeper, more severe pits with some fractographic topography on the pit bottom.

A cross section of the gear surface showed developing cracks at about the depth of the large pit bottoms and crack branching that was parallel to the surface. This would indicate some repeat load occurrences were required to propagate the cracks and separate the particles from their original surface. Severe deformation, along with significant localized tangential forces, preceded the near-surface cracking, and that multiple contacts were required before the particle fractured enough to break loose from the surface. Considering individual asperity, random, repeat contacts, the developing of the surface micropits, and the larger pitting were probably related to a fatigue mechanism.

This review of the basic wear mechanisms that can occur within the tribo line contact model underlines the similarity of the operating conditions that promote damaging wear, plastic flow, or fatigue. The lack of a good lubricant film, rougher contacting surfaces, higher operating temperature, the severity of the loading, and operating speeds either too low or too high promote gear or bearing degradation. Table 2 is taken from the ANSI/AGMA 110.04-1980 National Standard, "Nomenclature of Gear Tooth Failure Modes".[18] Included are the three failure modes, Surface Fatigue, Plastic Flow, and Wear, placed in the same order as in Table 1. As to be expected, there is considerable similarity.

The greatest loading difference is the amount of tangential force that accompanies the normal force for gears in comparison to bearings. The difference this makes is that the near-surface zone with asperity interaction is probably more critical for gears, and this confounds any clear differences between polishing and moderate wear, micropitting, and initial pitting and spalling and possibly the propagation of these into more destructive failure modes. However, with the possible overlap of some of these modes in mind, next consider Table 3, which was based on our recent bearing experience and suggests what consequences this approximate guide might have for gears.

Relating Contact Fatigue Damage Model To Modified Lambda

The concept of h/σ has been popular for over 20 years as a means to assess the contact lubricant condition within either bearings and gears.[8] The λm values are most helpful within the lubricant conditions that are primarily described as mixed EHL. This occurs when asperity contact and some lubricant separation of the surfaces exist together, but in varying degrees. Appropriate lambda ratio values for this cover the range of λm from less than 0.3 to greater than 1.0. A simplified summary of this is shown in Table 3. The lubricant regime range represented by λm is a continuum, and the specific λm values picked to represent portions of the continuum are somewhat arbitrary. However, Table 3 indicates a λm range that goes from essentially complete surface separation to boundary lubrication. Also included in the table are the three prime factors for altering bearing performance, material (primarily cleanness), surface roughness, and for line contact, internal geometry.

When λm is greater than 3.0, almost complete separation of the contact surfaces occur. Under this lubricant condition fatigue damage initiates subsurface to the contact, primarily at inclusions within the material. Thus, under this condition material cleanness is important, and fatigue life can vary greatly, especially in the direction of increased life. Geometry is also important, since proper surface profiles can minimize contact stress and permit the full potential of today's clean steel to be achieved. The influence of surface roughness is minor for this condition. For line contact bearings, the fatigue mode is observed as an elliptical spall with fairly slow propagation. In gears, this fatigue mode is called surface fatigue - destructive pitting in Table 2. This pitting originates subsurface at discontinuities (i.e. non-metallic inclusion or other stress-points) near the maximum subsurface shear stress region.

Next, considering λm in the range from 3.0 to 1.0, contact fatigue mode is still predominantly material related, but depending on the traction occurring on the surface, initiation can move closer to the surface. Surface asperities and asperity slope start to play a role when λm approaches 1.0

When λm is in the middle range 1.0-0.3, fatigue damage may be mixed, with surface roughness being important. Depending on the sharpness (asperity slopes) and relative heights of the surface asperities, fatigue can be surface-related or near-surface at the higher microstress regions just below the higher asperities. Fatigue life within this near surface regime is also influenced by material cleanness and very careful damage analysis is required to assess the actual crack initiation sites in this mixed EHL regime. Thus, the recent understanding of this near surface stress region helps explain the role of clean steel in improving fatigue performance under these essentially poorer operating conditions as illustrated by Fig. 5.

For gears, because the increased sliding in the surface contact area permits higher coefficients of friction, the shear stress region occurs over a wider range and closer to the surface (as in Fig. 3) than in bearing contacts that have nearly pure rolling.[10-12] Because of this, the fatigue mode may change and "spalling" occurs as given in Table 2. According to Drago[19] and Ref. 18, spalling is a different fatigue mode for gears. "In its early stages...cracks occur and spread from the origin, in a fan-shaped manner, in the direction of sliding until a piece of material is removed from the surface."[19]

It would seem that, depending on the λm value, the fatigue modes for gears in the light of the discussion under basic wear mechanisms, destructive pitting or spalling could compete, so that damage analysis takes careful evaluation. Such care in doing this is reported by Clark et al.[20] on geared roller test rollers. They detected subsurface pitting and surface spalling for tests with λm of about 0.1 to 0.4 and with 35% sliding between the rollers.

Thus, the fatigue mode and material/surface influences given in Table 3 for the λm range of 1.0-0.3 are fairly judgmental, but have validity based on Ref. 20 damage analyses.

Finally, when λm drops below approximately 0.3, surface related micropitting tends to predominate so that surface roughness is the prime influence on fatigue. Under these low modified lambda conditions, increased surface traction and localized frictional heating can take place with increased lubricant and additive interaction with the surfaces. For gears with increased sliding the traction forces, combined with the normally rougher surfaces, seem to require EP additives in the lubricant as protection for the contacting surfaces.

There is some evidence in Ref. 21 that EP additives improve fatigue life under λ values less than 1.0 and reduce life for λ values above 1.0, in disk-on-disk tests run with 10% sliding. The tests were with and without 5 wt. percent sulphur-phosphorus EP added to the same base oil. However, in the presence of debris, increased wear has been shown with S-P/EP in lubricants both for gears[22-23] and tapered roller bearings.[24]

For line contact bearings with much lower slip between the contacting surfaces, the role of additives is unclear. However, in bearings operating with low lambda values that promote boundary lubrication, chemical reactions and higher tractions can lower fatigue life to less than is predicted by using λm value assessments alone.

In bearings, the terms peeling and micropitting are used to describe the surface-related fatigue as Table 3 indicates. For gears, the term spalling is used for surface fatigue as given above. In Ref. 18 λ seems primarily applied to scoring and wear. From Table 2, frosting is considered surface distress that resembles bearing micropitting. There are similarities between the approaches to lessening peeling-micropitting in bearings and those to alleviating the occurrence of micropitting in gears. Both forms of surface distress are reduced by a reduction in surface roughness or increasing the oil film. That micropitting could be related to fatigue was demonstrated by Macpherson and Cameron[25] in some specific sliding disk tests they conducted in the early 1970's. They called the "new form" of gear failure "fatigue scoring." The work of Cummins and Doyle[17] reinforces this conclusion that micropitting is fatigue.

There are also localized stress risers that can occur either from handling damage before running, large debris particles going through the contact, or serious misalignment that causes end of contact edge damage. Within the bearing contact, failures from these stress risers are termed point surface origin (PSO); i.e., spalling from a localized point that propagates much as the "spalling" Drago describes for gears. At the end of contact, the failure is called geometric stress concentration (GSC). These stress risers are usually generated well beyond the elastic limit for the contact material so, based on their size relative to the lubricant film, they may influence fatigue life almost independently of λm. In bearings, the defects from debris are observed as isolated, pronounced dents, while in gears the debris develop long scratches or a series of grooves and ridges from the considerable sliding present in the contacts.

A fair volume of material around these "localized" stress risers may also be highly stressed and may be influenced by material cleanness. Of course, GSC and edge stress problems can be overcome with adequate design and proper profile geometry for the application. For line contact, as in gears or tapered roller bearings, from a fatigue standpoint only, the largest debris particles are capable of life-limiting dents in bearings or deep scratches in gears.

Summary

It is possible with the tribology line contact model to consider all the factors that influence and interact within the loaded concentrated contact. Two of the important factors are the film thickness (h) and combined surface roughness (σ). Together in the modified lambda, they are an indicator of the lubricant regime operating in a gear contact or roller bearings. A corollary to this model is the stress field developed below the concentrated contact that consists of subsurface shear stress and near surface shear stresses related to the surface asperities. These shear stresses are important in the understanding of near-surface plastic deformation accumulation and near-surface and subsurface fatigue.

The three basic mechanisms that lead to metal loss in steel gears or bearings are abrasive wear, adhesive wear, and fatigue in the presence of various levels of plastic deformation. Very often in real operating contacts, all three of these mechanisms exist together, and all contribute to the nucleation and propagation of the fractures that lead to the loss of material. An aid to separating out the primary mechanism is a careful examination of the contact surface and the debris formed. Optical microscopes, scanning electron microscopes, particle size analysis, and ferrography may be required to clearly identify the actual mechanisms present.

Table 3 presents the approximate relationship of contact fatigue mode for bearings and the modified lambda ratio. Although it is based on fatigue, an understanding of the other mechanisms at work that degrade gear and bearing performance and that as relate to λm may allow Table 3 to be used in a broader sense. Our review of the various surface and near surface damage modes in gears and the development of cracks and fractures that precede the formation of debris, when compared to the development of fatigue cracks for bearings, makes the relationships in Table 3 a reasonable starting point for defining the lubricant regime in gears. ■

Acknowledgement:
Presented at the 1991 AGMA Fall Technical Meeting. Reprinted with permission. The opinions, statements, and conclusions presented are those of the Author and in no way represent the position or opinion of AGMA.

References

1. Littmann, W. E. "The Mechanism of Contact Fatigue," NASA Symp. Interdisciplinary Approach to the Lubrication of Concentrated Contacts, New York, July 1969.
2. Tallian, T. E., Baile, G. H., Dalal, H. and Gustafsson, O. G. *Rolling Bearing Damage, A Morphological Atlas*, SKF Industries, Revere Press, Phil, 1974.
3. Wren, F. J. and Moyer, C. A. "Modes of Fatigue Failures in Rolling Element Bearings," *Proceed. I. Mech. E.* Vol. 179, pt 3D, pp 236-247, 1964-65.
4. Widner, R. L. and Littmann, W. E. "Bearing Damage Analysis," Symp. Mechanical Failures by Mech. Failures Prevention Group, NBS, May, 1974.
5. Fitzsimmons, B. and Clevenger, H. D. "Contaminated Lubricants and Tapered Roller Bearing Wear," *ASLE Trans.* Vol. 20, 2, pp 97-107, 1977.
6. Leaver, A. H., Sayles, R. S., and Thomas T. R. "Mixed Lubrication and Surface Topography of Rolling Contacts," *I. Mech. E. Proc.* 188, pp 461-469, 1974.
7. Wellauer, E. J. and Holloway, G. A. "Application of EHD Oil Film Theory to Industrial Gear Drives," *Trans. ASME JEI*, pp 626-634, May, 1976.
8. Moyer, C. A. and Bahney, L. L. "Modifying the Lambda Ratio to Functional Line Contacts," *STLE Trib. Trans.*, Vol. 33,4, pp 535-542, 1990.
9. Moyer, C. A. "Power Density Development: The Role of Improved Line Contact Performance," AGMA Tech. Paper 89 FTM3, Nov., 1989.
10. Kannel, J. W. and Tevaarwerk, J. L. "Subsurface Stress Evaluations Under Rolling/Sliding Conditions," *ASME Trans. JOT*, Vol. 106, pp 96-103, Jan., 1984.
11. Fessler, H. and Ollerton, E. "Contact Stresses in Toroids Under Radial Loads," *British Journal of Applied Physics*, Vol. 8, pp 387-393, Oct., 1957.
12. Zhou, R. S., Cheng, H. S. and Mura, T. "Micropitting in Rolling and Sliding Contact Under Mixed Lubrication," *ASME Trans. JOT*, Vol. 111, pp 605-613, Oct., 1989.
13. Vingsbo, O. "Wear and Wear Mechanisms," *ASME Proceedings of Wear of Materials Conference*, Ed. K. Ludema et al., pp 620-635, 1979.
14. Jahanmir, S. "The Relationship of Tangential Stress to Wear Particle Formation Mechanisms," *ASME Proceedings of Wear of Materials Conference*, Ed. K. Ludema., pp 238-247, 1985.
15. Suh, N. P. "An Overview of the Delamination Theory of Wear," *Wear*, 44, pp 1-16, 1977.
16. Rigney, D. A. and Glaeser, W. A. "The Significance of Near Surface Microstructure in the Wear Process," *ASME Proceedings of Wear of Materials Conference*, Ed. Ludema, pp 41-46, 1977.
17. Cummins, R. A. and Doyle, E.D. "Interpretation of Gear Wear," *ASLE Trans.*, Vol. 25, 4, pp 502-510, Oct. 1982.
18.____. "National Standard - Nomenclature of Gear Tooth Failure Modes," AGMA Standard, ANSI/AGMA 110.04-1980.
19. Drago, R. J. "Failure Modes," Chapter 6, *Fundamentals of Gear Design*, Butterworths Publishers, Stoneham, MA, pp 215-263, 1988.
20. Clark, T. M., Miller, G. R., Keer, L. M. and Cheng, H.S. "The Role of Near-Surface Inclusions in the Pitting of Gears," *ASLE Trans.*, Vol. 28, 1, pp 111-116, Jan., 1985.
21. Phillips, M. R. and Quinn, T. F. J. "The Effect of Surface Roughness and Lubricant Film Thickness on the Contact Fatigue Life of Steel Surfaces Lubricated With a Sulphur-Phosphorus Type of Extreme Pressure Additive," *Wear*, 51, pp 11-24, 1978.
22. Errichello, R., "Lubrication of Gears, Part 2," *Lubrication Engineering*, pp 117-121, Feb. 1990.
23. Milburn, A., Errichello, R., and Godfrey, D. "Polishing Wear," AGMA Tech Paper 90 FTM5, 1990.
24. Fitzsimmons, B. and Clevenger, H. D. "Contaminated Lubricants and Tapered Roller Bearing Wear," *ASLE Trans.*, Vol. 20, 2, pp 97-107, 1977.
25. Macpherson, P. B., and Cameron, A. "Fatigue Scoring a New Form of Lubricant Failure," *ASLE Trans.* Vol. 16, 1, pp 68-72, Jan., 1973.
26. Errichello, R. "Lubrication of Gears, Part 1," *Lubrication Engineering*, pp 11-13, Jan., 1990.