

Tribology of Power Train Systems

H. Rahnejat, R. Rahmani, M. Mohammadpour, and P.M. Johns-Rahnejat
Loughborough University, United Kingdom

Introduction

Tribology is the study of friction, lubrication, and wear. It is a multidisciplinary subject covering the mechanics of contacting surfaces, their roughness characteristics, lubrication, and material behavior under normal load as well as in traction. A main parameter of interest in any tribological study is the calculation of lubricant film thickness; an insufficient lubricant film can lead to the direct contact of mating surfaces, which would lead to increased friction, heat generation, and many mechanisms of failure, such as wear, surface scoring, scuffing, and so on. The other parameter of importance is the generated contact pressures, causing elastic deformation of the surfaces.

Issues of component reliability, service life (longevity), and operational efficiency have become progressively more important under pervading global competition. There is a significant volume of literature on a plethora of tribological issues, most of which requires a good knowledge of specialist skills in numerical analysis and/or the use of specialist methods of measurement. The diversity of the subject and the multitude of issues preclude the possibility of including everything in this article.

In dealing with some of the issues in the area of engine and power train tribology, this article is confined to well-established and widely accepted analytical methods and design and analysis charts. This includes coverage on lubricant rheology and prediction of lubricating film thickness. Better lubrication reduces friction, resulting in fewer incidents of wear and thereby improving component reliability and life. Combined with improved contact mechanics, it also reduces stresses and lowers the chance of contact fatigue.

In engine and power trains, as in many other systems, friction as a source of energy loss is of paramount concern for energy efficiency and reduced emissions (see also the preceding article, "Friction, Lubrication, and Wear of Internal Combustion Engine Parts," in this Volume). Frictional losses in an internal combustion engine account for 15 to 20% of all its losses. The main contributor to the frictional power

loss is the piston-cylinder system, accounting for 40 to 50% of all these losses, from piston skirt-cylinder liner and piston ring pack to cylinder liner contacts.

Engine journal bearings, comprising big-end (connecting rod) bearings and crankshaft support bearings, account for 20 to 30% of the frictional power loss, with the valve train contacts, particularly the cam-follower pair, contributing to 5 to 10% of these losses. The remaining frictional losses are from other conjunctions, such as cam gears, camshaft bearings, and small-end (wrist-pin) bearings (sometimes referred to as the gudgeon pin), as well as due to pumping losses.

Contact Configuration

A primary concern is to choose the correct method for a given problem. The most appropriate analytical methods to use for tribological analyses tend to vary according to the type of contact (contact configuration). The classification depends on the degree of conformity of the mating surfaces. When the two surface geometries conform closely to one another (snug closely, as shown in Fig. 1a), the configuration is referred to as conformal or conforming. A relatively large contact area results, and generally pressures of the order of a few to tens of megapascals are generated in gaps of several to tens of micrometers. Good examples are journal bearings supporting the crankshaft and

camshaft, and the connecting rod (big-end) bearing between the connecting rod and the crankshaft. Note that the journal closely conforms to the surface of the bearing bushing or shell. In other words, the convex radius of the journal, R_1 , is close to the concave radius of the bushing, R_2 (of the order 95% or closer). Conversely, the contact of a perfectly spherical ball bearing residing on a perfectly flat surface is just a single point (Fig. 1c). The convex radius of the sphere, R_1 , does not conform at all to the infinite radius of an ideal flat surface. This form of contact is nonconformal, sometimes referred to as counterformal, such as the contact between a pair of touching cylinders having their axes perpendicular to each other. With a normal applied load, the surface of the sphere at the point of contact flattens (its deformation follows Hertzian contact theory, as discussed later) if the flat counterface surface is assumed to be ideally rigid. The footprint shape becomes a small circle of a few tenths of a millimeter radius subjected to high pressures, usually in the range of several hundred megapascals to a few gigapascals. Because the pressures are generated over such a small contact area, this class of contact is often referred to as concentrated counterformal circular point contact.

In an actual ball bearing, the balls are retained in raceway grooves (Fig. 1b), although there is some degree of conformity (the radius of the ball can be as much as 80 to 85% of the raceway groove). The case is still termed

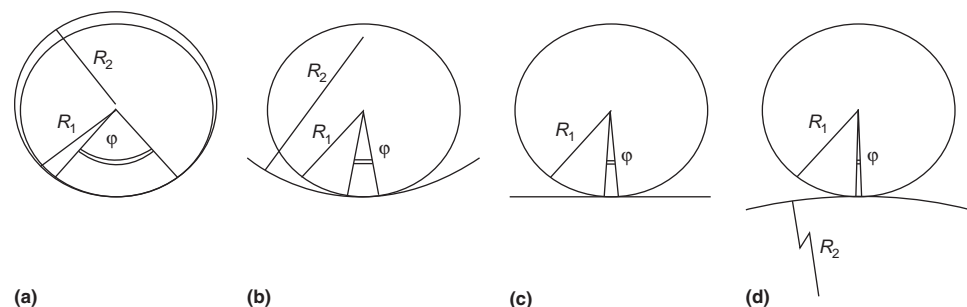


Fig. 1 Various contact configurations

nonconformal, with the resulting contact footprint being elliptical in shape (i.e., concentrated counterformal elliptical point contact). The ball contact with the raceway groove can be viewed in the zy -plane of contact (Fig. 2), where the ball rotational velocity vector is perpendicular to the plane of paper (in the x -direction, shown in Fig. 1d). Therefore, a counterformal contact also occurs in the zx -plane, a two-dimensional footprint of elliptical shape.

There is a larger class of contacts that is neither conformal nor counterformal, for example, a piston compression ring ideally conforming circumferentially to the inner surface of an ideal right circular cylinder but with a parabolic contact face profile not conforming to the cylinder wall surface in the direction of sliding along the cylinder axis. The contact radius of an ideal cylinder along its axis is considered to be infinite, although in practice this is not the case; usually there is a local contact radius of tens of meters. This case and many other similar contacts are referred to as partially conforming (Fig. 1b). For piston compression rings, film thickness of a few tenths to a few micrometers would be expected at contact pressures of a few to several megapascals.

Mechanics of contact depends on contact conformity as well as many other factors. Therefore, care should be taken to use the appropriate analytical method of prediction (as highlighted later) for the cases considered in this article.

Contact Mechanics—Footprint Shape and Elastic Deformation

The first step in any analysis is to determine the contact area, where the applied load is carried. This is the prelude to calculating the generated pressures as well as any localized contact deformation, leading to the determination of the lubricant film thickness. In the field of dynamics/vibration, the general approach of any analytical solution is to reduce the problem to a series of masses, stiffness and damping elements, and any applied forces. Then, for any given mass, Newton's second law of motion is applied to obtain the unknown acceleration:

$$a = \frac{1}{m} \sum F$$

where ΣF is the net force acting on the mass, m , inclusive of the restraining forces of stiffness elements, dissipative damping forces, and any applied forces. This simple approach leads to the calculation of the unknown kinematic quantities: acceleration, a ; velocity, $v = \int a dt$; and displacement, $x = \int v dt$. A similarly simple yet profound model is required for contact mechanics. Hertz provided such a model, where the contact configurations in Fig. 1 are represented by a rigid ellipsoidal solid of revolution (Fig. 2) contacting an elastic half-space of infinite dimensions compared with the dimensions of the rigid ellipsoid, and with equivalent elastic properties to the original contacting pair. The ellipsoidal solid has two principal radii of contact in the planes zx and zy , which represent the curvatures of the original contacting pair at the point or line of their touching contact. The two principal radii of the equivalent ellipsoidal solid, R_{zx} and R_{zy} (Fig. 2), become:

$$\frac{1}{R_{zx}} = \frac{1}{R_{x1}} + \frac{1}{R_{x2}} \text{ and } \frac{1}{R_{zy}} = \frac{1}{R_{y1}} + \frac{1}{R_{y2}} \quad (\text{Eq 1})$$

For the case shown in Fig. 2, because the radii of the original contacting solids differ in the planes of contact, zx and zy , the equivalent solid becomes an ellipsoid, as also shown in the same figure. Note that concave radii, such as that of the raceway groove, R_{y2} , are considered to be negative (a concave curvature). It can be seen that when a concave radius exists, the equivalent radius of the ellipsoidal solid will be larger than that of either of the original bodies. Such a radius is often referred to as an equivalent increased radius (R_{zy} in Fig. 2). Conversely, with a pair of convex radii, the equivalent radius of the ellipsoidal solid will be smaller than both the original contacting solids, referred to as the decreased equivalent radius (R_{zx} in Fig. 2). A special case is a sphere on a plane (Fig. 1c), where for the sphere, $R_{x1} = R_{y1} = R_1$, and for the flat plane, $R_{x2} = R_{y2} = \infty$. Thus, the equivalent radii of the ellipsoidal solid are $R_{zx} = R_{zy} = R_1$, which is the spherical ball itself. Similarly, for a roller on a flat plane, $R_{zx} = R_1$ and $R_{zy} \rightarrow \infty$; therefore, a roller of nominally infinite length results, which, in practice, has finite length (the length of the roller). Referring to Fig. 1(a), assuming that $R_1 =$

$0.95R_2$, and using Eq 1, it can be seen that the equivalent increased radius becomes $R_{zx} = 20R_2$, which yields a large equivalent cylinder with its length equal to the length of the bearing into the plane of paper along the zy -plane for a journal bearing. It can be shown that the resulting gap at the center of contact of this equivalent cylinder with a flat surface is not representative of the original clearance between the journal and the bearing bushing. Therefore, it is important to note that the Hertzian methodology described in this section does not apply to conforming contacts of the type depicted by Fig. 1(a). This point is rarely understood, resulting in a large volume of reported literature applying Hertzian contact mechanics analysis to journal bearings, ball-in-socket joints of a high degree of conformity, and even hip joint prostheses, which is quite erroneous. Furthermore, in wear analysis of counterface materials for such applications, pin-on-disc tribometers with pins of fairly small radii are often used, which is unrepresentative of the contact conditions. The main reason for such misconception is that Hertzian contact mechanics deals with the localized nature of deformation (small elastic strain assumption). Of course, this is not the case for any deformation of larger contact domains, such as journal bearing shells, rings, and any body where the extent of deformation may result in significant changes to its overall geometry. These forms of deformation may be considered as global rather than local. The appropriate theories for these were originally set forth by St. Venant, culminating in specific analytical and numerical methods, including finite-element analysis. As the equivalent rigid solid of any shape (roller/cylindrical, sphere, or ellipsoidal shape) is pressed onto an elastic plane of equivalent elastic properties to those of the original contacting pairs, the resulting contact footprint shape can be visualized. For a sphere this is a circle, for a cylinder it is a narrow rectangular band, and for a generalized ellipsoid it is an elliptical shape. For a ball in a raceway groove, this is an ellipse of semi-half-widths a and b (Fig. 3a); for a sphere on a flat, this is a circle, where $a = b$; and for a cylinder, a narrow rectangular band results (Fig. 3b). The pressure distribution on the elliptical

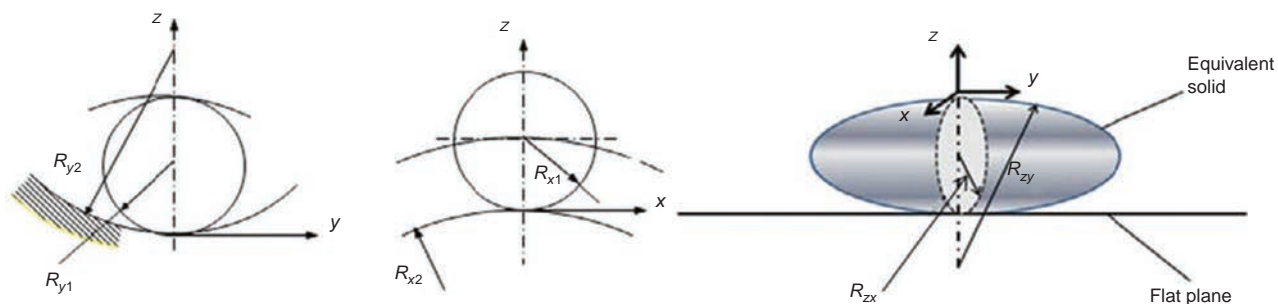


Fig. 2 Principal planes of contact and the equivalent ellipsoidal solid

contact footprint is an ellipsoid, and that over the elastic line contact (rectangular band) is elliptical in cross section of an infinite length according to classical Hertzian theory (Ref 1). However, when a rigid cylinder of finite length is pressed onto an elastic half-space, the contact footprint spreads out at the edges due to the profile discontinuity (sharp edges) of the roller. Therefore, a dog-bone- or dumbbell-shaped footprint of finite length results, as shown by Johns and Gohar (Ref 2) (Fig. 3c). The sharp edges of the roller yield high generated pressures that can lead to wear

and fatigue. This is the reason for blending the edges of roller bearings; they are usually relieved by the introduction of dub-off radii or through crowning. The contact relationships provided in this section disregard these subtleties; otherwise, a numerical analysis would be required. The ratio of the elliptical footprint semi-half-widths a and b is known as the ellipticity ratio (sometimes referred to as the aspect ratio). This is determined as:

$$e_p^* = \frac{a}{b} \sim \left(\frac{R_{zy}}{R_{zx}} \right)^{2/3} \quad (\text{Eq 2})$$

where, for a circular point contact, $a = b$; thus, $e_p^* = 1$. Recall that for the idealized Hertzian infinite line contact, $R_{zy} \rightarrow \infty$; thus, $e_p^* = \infty$. However, in reality, the length of a line contact is finite. It is reasonable to assume a line contact for $e_p^* > 10$ for a narrow rectangular band of width b . The effective or equivalent modulus of the elastic half-space is:

$$\frac{2}{E^*} = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2} \quad (\text{Eq 3})$$

where E_1 and E_2 are the Young's elastic moduli of the contacting bodies, and ν_1 and ν_2 are their respective Poisson's ratios. When the contacting surfaces are made of the same material, then the equivalent elastic modulus becomes:

$$E^* = \frac{E}{1 - \nu^2} \quad (\text{Eq 4})$$

which is usually termed the plane-strain elastic modulus.

Ignoring the magnitude of the edge pressure peaks (sometimes referred to as pressure spikes or "pips"), whose determination requires numerical analysis, Table 1 provides the contact dimensions, localized elastic deformation, and generated pressures for different concentrated counterformal contact types.

Within the context of this article, the Hertzian relations in Table 1 apply to the contact of gearing systems in vehicular transmissions, differentials, cam-followers, and a host of other power train subsystems, where use is made of ball or rolling-element bearings. In all of these cases, the contact geometry and applied load must be determined before these relations can be used. Other key parameters are not used in classical Hertzian theory, including operational speed and temperature. This is because Hertzian theory is essentially for elastostatic frictionless contacts. Nevertheless, in many cases the theory provides a good estimate of the prevailing conditions for generated pressures and contact deflection that closely estimate the real conditions for lubricated concentrated counterformal contacts. The frictionless assumption of Hertzian theory is often relaxed, and a coefficient of friction is used to estimate friction. This approach is very empirical and should be avoided. Later, appropriate methods are outlined for calculating friction.

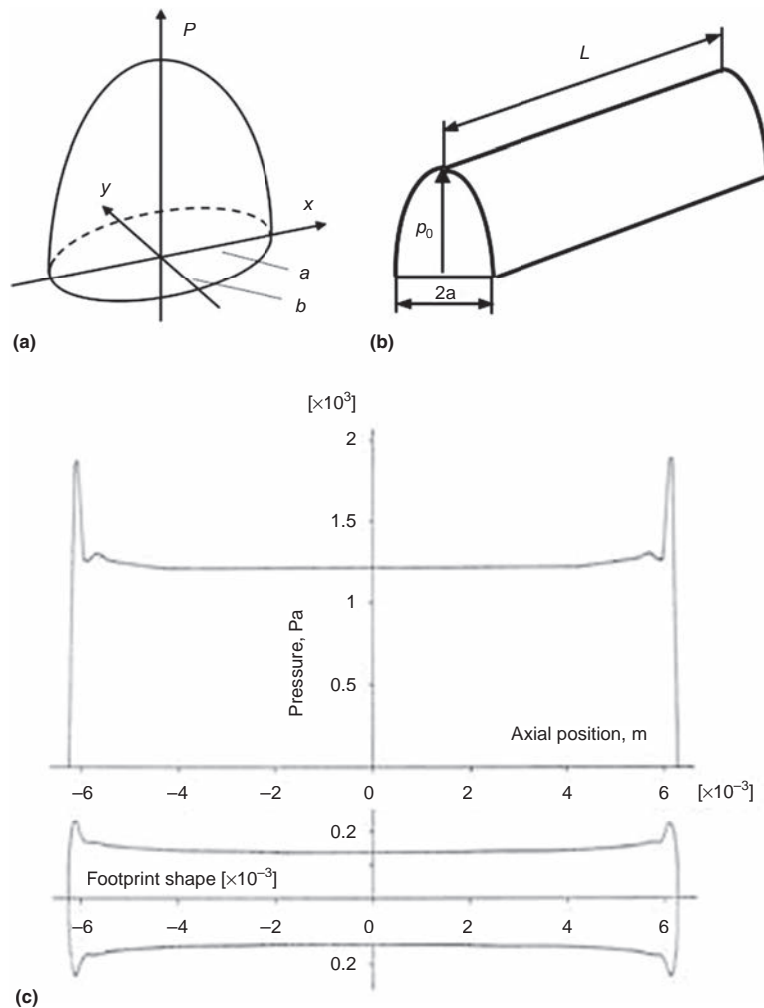


Fig. 3 Contact footprint and pressure distribution. (a) Elliptical point contact. (b) Long line contact. (c) Finite line contact

Table 1 Relationships for Hertzian contacts

Variable	Elastic line contact	Circular contact	Elliptical contact
Contact half-width or radius	$a = \left(\frac{4WR}{\pi LE^*} \right)^{1/2}$	$a = \left(\frac{3WR}{4E^*} \right)^{1/3}$	$\sqrt{ab} = \left(\frac{3W\sqrt{R_{zx}R_{zy}}}{4E^*} \right)^{1/3}$
Maximum and mean contact pressures	$p_0 = \frac{4}{\pi} p_m = \left(\frac{WE^*}{\pi RL} \right)^{1/2}$	$p_0 = \frac{3}{2} p_m = \left(\frac{6WE^*}{\pi^2 R^2} \right)^{1/3}$	$p_0 = \frac{3}{2} p_m = \left(\frac{6WE^*}{\pi^2 R_{zx}R_{zy}} \right)^{1/3}$
Contact load	$W = 2aLp_m$	$W = \pi a^2 p_m$	$W = \pi ab p_m$
Contact center deflection	$\delta = \frac{W}{\pi LE^*} \left[\ln \left(\frac{L^2 \pi E^*}{2RW} \right) + 1 \right]$	$\delta = \frac{\pi p_0 a}{2E^*} = \left(\frac{9W^2}{16E^{*2}R} \right)^{1/3}$	$\delta = \frac{1}{2} \left(\frac{9W^2}{2E^{*2}\sqrt{R_{zx}R_{zy}}} \right)^{1/3}$

Contact Fatigue

With calculated pressures for any of the cases highlighted in Table 1, the integrity of the contact under the prevailing condition can be ascertained. One mode of failure is fatigue of the contacting surfaces. Fatigue occurs as the result of inelastic deformation due to generated stresses beneath the contacting surface reaching a critical value. There are various fatigue failure criteria. One is the maximum subsurface shear stress criterion, known as

Tresca. In a two-dimensional analysis, the maximum shear stress is obtained as:

$$\tau_{\max} = \frac{1}{2} |\sigma_1 - \sigma_2| \quad (\text{Eq 5})$$

where σ_1 and σ_2 are the principal stresses. As the applied load increases, so do the maximum Hertzian pressure, p_0 , and the principal contact stress difference. Therefore, the maximum subsurface shear stress increases. The onset of inelastic deformation occurs when:

$$\tau_{\max} = \frac{\sigma_Y}{2} = 0.3p_0 \quad (\text{Eq 6})$$

where σ_Y is the yield stress of the softer of the two counterface materials. Therefore, for a given applied load, the onset of plastic deformation due to the Tresca yield criterion occurs at the maximum Hertzian pressure of:

$$p_0 = 1.67\sigma_Y = 0.6H \quad (\text{Eq 7})$$

where H is the material indentation hardness (this is often provided in Vickers or Rockwell scale, but the units used in Eq 7 are in pascals). Hence, for a given load, such as the contact load on gear teeth or a flat tappet, and for given contacting materials, if the maximum Hertzian contact pressure exceeds that given in Eq 7, then onset of inelastic deformation occurs at a depth z beneath the contact footprint center in the softer of the two counterface surfaces, after some degree of work hardening, of course. The chance of failure is increased with the existence of any fault such as a crack, pore, void, and so on within the bulk of the material. Commonly, a crack grows to the contact surface, and a small piece of material is removed. The phenomenon is referred to as fatigue spalling (ISO 6336). The depth at which the maximum shear stress occurs is $z = 0.78a$, where a is the half-width of the narrow band in the long line contact (Table 1). Clearly, with the finite line contact of a roller against a raceway or a cam against a flat tappet, the edge pressure spikes (Fig. 2) induce subsurface stress fields of their own. Because these edge stresses are usually higher than the maximum Hertzian pressure, larger maximum subsurface shear stresses occur there nearer to the contact surface, resulting in a greater chance of fatigue failure (Ref 3).

The Tresca criterion is one of a number of proposed yield criteria and is particularly suitable for relatively hard-wearing but brittle contacting surfaces, such as many coatings of high hardness. Teodorescu et al. (Ref 4) provide

numerically obtained subsurface stress fields for different coated surfaces, which show that the coincidence of maximum shear stress with the interfacial layer between a coated layer and the original substrate material can lead to their exfoliation at given maximum pressures or tangential traction.

Using the Tresca yield criterion for spherical surfaces, the maximum Hertzian pressure (for elliptical and circular point contacts) at the onset of yield becomes:

$$\tau_{\max} = 0.31p_0 = \frac{\sigma_Y}{2} \quad (\text{Eq 8})$$

Thus, the maximum Hertzian pressure at the onset of yield for an elliptical point contact becomes:

$$p_0 = 1.60\sigma_Y = 0.6H \quad (\text{Eq 9})$$

The depth at which the maximum shear stress occurs is $z = 0.47a$.

There are other failure criteria, including the distortion energy hypothesis, also known as the von Mises criterion. This criterion is more suited to ductile substrates, hypothesizing the onset of yielding to be when the defined equivalent stress, σ_e , reaches a certain limit (Ref 5, 6). The equivalent stress is defined in terms of the subsurface stress field, $\sigma_e = f\{\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \tau_{xy}, \tau_{xz}, \tau_{zy}\}$ (Ref 3). Various yield hypotheses are used with the equivalent stress. One that is progressively favored for bearings, gears, and cam-follower pairs is the alternating shear stress hypothesis (Ref 3, 7, 8). This is based on the shear stresses τ_{xz} , τ_{zy} , which occur on orthogonal planes beneath the contact footprint in a cyclic manner, alternately stretching (tensioning) and compressing the bulk material (Fig. 4). Although these shear stresses have lower magnitudes than the maximum subsurface shear stress, their repetitive cyclic reversals are often responsible for limiting the useful life of the bearing (Ref 7, 8). Using the alternating shear stress hypothesis with the distortion energy criterion, the equivalent stress at the onset of yield becomes:

$$\sigma_e = 2|\tau_{z\max}/p_0| \quad (\text{Eq 10})$$

where the alternating subsurface shear stress distribution is given by Johnson (Ref 9), requiring numerical evaluation, and its maximum double amplitude can be found as shown in Eq 10.

The contact mechanics approach highlighted here is applicable to many counterformal

concentrated contacts, including cam-followers, transmission gears, rolling-element, and ball bearing supports in some engines, transmissions, driveline, and axle components. In other words, the theory is generic and widely applicable for estimating applied pressures, contact stresses, and localized surface deformation and for predicting the onset of fatigue. However, surface topographical information, lubricant properties, and an estimate of lubricant film thickness are still required to predict friction, power loss, and generated contact temperature, all of which are critical in the assessment of wear and system efficiency. Therefore, it is necessary to determine the prevailing regime of lubrication.

Regimes of Lubrication

Lubrication is an important component of tribology. Its main purpose is twofold. First, it must provide a thin low-shear-strength layer of lubricant (termed the lubricant film thickness) to ideally completely separate the mating contacting surfaces and carry the applied contact load. Secondly, the flow of lubricant should cool the contacting surfaces by carrying a proportion of the generated heat away through convection cooling.

Real engineering surfaces are rough (not smooth and frictionless, as assumed in classical Hertzian theory). Boundary friction is caused by the interaction of roughness peaks on opposing contacting surfaces. (There is also viscous friction arising from shear of a thin lubricant film formed in the contact.) This means that in boundary friction the relative motion of contacting surfaces is resisted by their roughness features, which must be sheared for continuance of motion. A much greater effort is usually needed for this than for shearing of a fluid (a lubricant film, i.e., viscous friction). Viscous friction can become significant at high loads when non-Newtonian lubricant behavior occurs, and lubricant shear stress tends to be its limiting value. The mechanisms underlying friction are discussed later in the section "Friction and Power Loss" in this article. It suffices to state that, depending on the thickness of the lubricant film, various degrees of direct interaction of rough surface topography occur. The regime of lubrication depends on the extent of these interactions. Stribeck (Ref 10) devised a convenient way of relating roughness and film thickness by a parameter, now called the oil film ratio or Stribeck's oil film parameter:

$$\lambda = \frac{h}{\sigma_s} \quad (\text{Eq 11})$$

where h is the lubricant film thickness, and $\sigma_s = \sqrt{\sigma_{s1}^2 + \sigma_{s2}^2}$ is the root-mean-square roughness of the two surfaces (the composite surface roughness); σ_{s1} and σ_{s2} are the individual surface roughnesses (usually the mean roughness heights of each contacting surface,

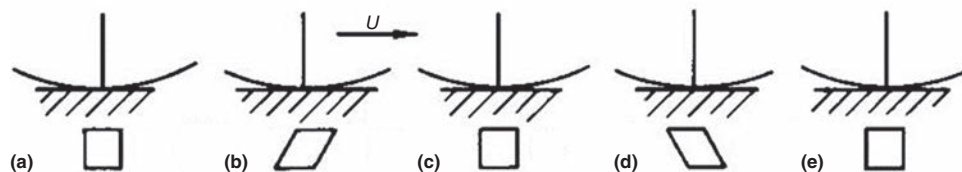


Fig. 4 Material subjected to reversing orthogonal shear stress. U , direction of rolling

although in many cases a better statistical parameter would be more appropriate, such as that discussed in the section “Surface Topography” in this article).

Stribeck proposed a curve relating the coefficient of friction to his oil film parameters (Fig. 5). The coefficient of friction is $\mu = \frac{f}{W}$, where f is friction, and W is the applied contact load. The Stribeck curve is only instructive, not quantitative. For a pair of surfaces and a particular lubricant type, such curves are obtained using tribometers (devices that measure the resistance to motion under various operating conditions: applied load, sliding speed, and bulk surface temperature). There are many types of tribometers, including pin-on-disc, roller-on-disc, and sliding strip-on-plate, each representing a different contact configuration. Depending on the operating conditions, it is now more usual to construct the Stribeck curve with its abscissa being the Hersey parameter, $\frac{\eta U}{p}$, where for a line contact, $p = \frac{W}{L}$ (load per unit length), making the Hersey number nondimensional. Changing the contact speed, U , for a given lubricant dynamic viscosity, η , and a given load and bulk temperature, the lubrication condition alters as shown by the Stribeck curve. For all other operating conditions remaining the same, counterface surfaces of different composite root-mean-square roughness yield a Stribeck curve as well (Fig. 5).

The coefficient of friction of unity on the vertical axis is quite arbitrary but represents very clean, smooth counterface surfaces, where adhesion would take place. When the abscissa is the Hersey number, the intercept with the vertical axis represents the static coefficient of friction, where the contact is dry because no lubricant film is present when there is no relative motion of the surfaces. In fact, it is often quite difficult to measure friction at very low speeds, because stick-slip motion occurs under these conditions.

The region between A and B is termed the boundary regime of lubrication and has relatively high friction, because the lubricant film is insufficient, being less than the average roughness of the counterface surfaces. For ground steel counterfaces, the measured coefficient of friction in this region is often reported to be 0.2 to 0.3. Because an oxide film is formed on the surfaces in normal atmosphere, the coefficient of friction is reduced (surface oxides have lower shear strength than the parent surface material). With an increasing speed, a progressively thicker lubricant film is formed, and the value of λ increases; thus, a lower number of contact roughness heights (asperities) interact. Therefore, the generated friction in the region B to C is due to the viscous shear of a lubricant film and a decreasing level of direct solid boundary friction. The lubrication in this region is termed a mixed regime of lubrication (a mix of fluid film and boundary lubrication, formed oxides, or, in some cases, lubricity of surface coatings).

The region C to D and beyond corresponds to $\lambda \geq 3$, which for a Gaussian distribution of surface asperity heights is shown to be subject to a fluid film regime of lubrication. Generated friction in this region is entirely due to the shear of a lubricant film, which would completely separate the rough counterface surfaces. There are various forms (modes) of fluid film regimes of lubrication. The two main forms are hydrodynamics and elastohydrodynamics. The former refers to the formation of a coherent lubricant film between the contacting surfaces without any localized contact deformation. This occurs at low-to-medium applied loads, with pressures not exceeding several to tens of megapascals and with surfaces of high elastic modulus. The film thickness increases with contact velocity, thus increasing the λ value (for a given surface roughness). The coefficient of friction is much lower than that with mixed and boundary regimes of lubrication (a representative value is shown in Fig. 5). If the coefficient of friction is conceived to be the fractional energy loss in a contact, then for a boundary regime of lubrication a value of $\mu = 0.25$ represents 25% energy loss or, in other words, a contact of 75% efficiency, while under a fluid film regime of lubrication with $\mu = 0.005$ the energy loss would be merely 0.5%. This simple example shows the importance of the principle of lubrication. The dip in region C to D occurs in cases where localized deformation of surfaces occurs (elastohydrodynamic lubrication, or EHL; surface deformation and piezo-viscous action of the lubricant, described in the section “Lubricant Rheology” in this article), increasing the gap filled with a film of lubricant and reducing the chance of direct contact of surfaces (greater values of λ). Hence, the coefficient of friction is reduced.

As the contact velocity increases beyond point D in Fig. 5, the lubricant film is enhanced, but the generated heat due to increased shear lowers the lubricant viscosity, η , and thus the generated pressures, thereby reducing the contact deformation. Hydrodynamic conditions with increased friction and drag would ensue.

Not all types of contact, including those in power train systems, undergo all the aforementioned regimes of lubrication under normally

specified operating conditions. For instance, crankshaft journal bearing supports often operate under a hydrodynamic regime of lubrication, apart from under start-up conditions or under intermittent stop-start in congested traffic, where entrainment of the lubricant into the contact is interrupted. During such periods there may be insufficient film thickness, and some degree of direct surface interactions occurs. To mitigate the resulting excess friction and wear under these conditions, the surface of the bearing bushing is often coated with a thin wear-resistant layer, for example, with bismuth or indium (Ref 11). Sometimes the bushing surface has several soft and hard layers; the soft layers allow some deformation (overlay bearings) and enhance the gap and improve film retention, while hard top layers resist wear (Ref 8, 11). With pressures of the order of tens of megapascals, the soft layers such as aluminum, copper, or tin deform, but the lubricant acts under iso-viscous conditions. This regime of lubrication is known as iso-viscous-elastic (soft EHL) as opposed to elastohydrodynamics (or hard EHL), where high pressures increase the viscosity of the lubricant as well (piezo-viscous effect). Therefore, the fluid film lubrication after point D in Fig. 5 can be classified as (Fig. 6):

- *Iso-viscous rigid*: Under this condition, the generated pressures are insufficient to bring about viscous action of the lubricant (see the section “Lubricant Rheology” in this article) or cause appreciable deformation of contacting surfaces. This regime of lubrication is termed hydrodynamics in the traditional sense. Applications include thick-shell journal bearings, often used in large turbomachinery, as well as nominally lightly loaded but interacting gears, such as the rattling unselected gears of vehicular transmissions (the gear rattle phenomenon) (Ref 12). Note that the first word, *iso-viscous*, refers to the state of the lubricant, while the second, *rigid*, refers to the state of deformation of the contacting surfaces. The same adopted definition is true for the other following categories as well.

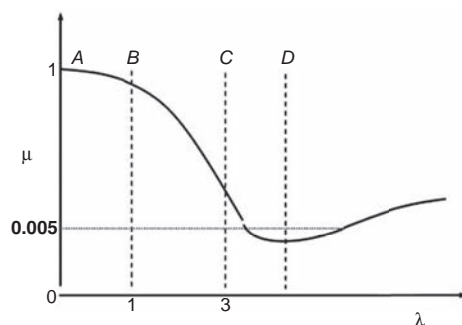


Fig. 5 Stribeck curve (not to scale)

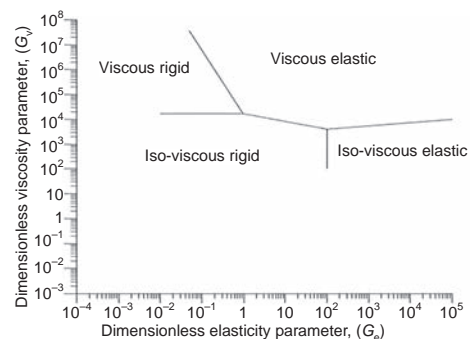


Fig. 6 Fluid film regimes of lubrication

- *Viscous rigid*: The generated pressures alter the lubricant viscosity, but the contiguous surfaces remain undeformed. These conditions are also referred to as hydrodynamic. They dominate in applications with hard surface materials, such as in some journal bearings, for example, the piston pin-bore bearing, often hardened or coated with materials of high elastic modulus.
- *Iso-viscous elastic*: The lubricant viscosity remains unaltered, but the contiguous surfaces undergo elastic deformation. This condition is sometimes referred to as soft EHL, such as in the case of thin-shell or overlay journal bearings.
- *Viscous elastic*: In addition to the lubricant viscosity changing, the contiguous solid surfaces deform. This condition is often referred to as hard EHL. This is the main regime of lubrication for loaded gears, rolling-element bearings, and cam-follower pairs.

The abscissa and the ordinate for Fig. 6 are defined as:

$$\text{Elastic parameters: } G_e = \frac{W^{8/3}}{U^{*2}} \quad (\text{Eq 12})$$

$$\text{Viscous parameters: } G_v = \frac{G^* W^{*3}}{U^{*2}} \quad (\text{Eq 13})$$

Both parameters are nondimensional and depend on the dimensionless groups G^* , W^* , and U^* , which depend on the operating contact conditions. These parameters depend on the contact configuration (see the sections “Contact Configuration” and “Contact Mechanics—Footprint Shape and Elastic Deformation” in this article) such as the materials parameter:

$$G^* = E^* \alpha \quad (\text{Eq 14})$$

where E^* is given by Eq 3, and α is the pressure-viscosity coefficient of the lubricant (see the section “Lubricant Rheology” in this article).

The load parameter is:

$$W^* = \frac{W}{E^* R_{zx} R_{zy}} \quad (\text{Eq 15})$$

where Eq 15 is for the general case of an elliptical contact footprint. For a circular point contact, $R_{zx} = R_{zy} = R$, and for a finite line contact geometry (see the section “Contact Fatigue” in this article), $R_{zy} = L$.

The speed (or rolling viscosity) parameter is:

$$U^* = \frac{U \eta}{E^* R_{zx}} \quad (\text{Eq 16})$$

where U is the speed of entraining motion of the lubricant into the contact; $U = \frac{u_1 + u_2}{2}$, which is the average of the surface speeds of the contacting bodies; η is the dynamic viscosity of the lubricant (see the section “Lubricant Rheology” in this article); and R_{zx} is the equivalent principal radius of the ellipsoidal solid (see the section “Contact Fatigue” in this article) in the

direction of rolling/sliding. This depends on the contact geometry. In some cases, the lubricant entrainment is at an angle to the principal radii of contact, such as in hypoid gears. The dimensionless parameters U^* , G^* , and W^* are often used in film thickness equations found through an extensive series of numerical predictions and formed through regression analyses for nonconforming contacts. As can be seen, the dimensionless parameters depend on some lubricant properties (see the section “Lubricant Rheology” in this article). All the expressions stated with regard to Fig. 6 are valid for nonconforming or partially conforming contacts only. For conformal contacts, such as journal bearings, a different approach is used (see the section “Piston-Cylinder Conjunctions” in this article).

Lubricant Rheology

Lubricants used in internal combustion engines comprise a base oil and a multitude of additives to improve its performance. These include viscosity modifiers to improve load-carrying capacity, antiwear agents to reduce wear of surfaces, antioxidants to inhibit lubricant thermal degradation, friction modifiers to reduce the effect of boundary friction, and others. The additives are organic or inorganic molecules and compounds at certain small percentages (or parts per million) added to the base oil to enhance certain desired measures of performance.

The main properties of the lubricant in terms of load-carrying capacity correspond to the formation of a thicker microscale film, due to its effective viscosity as well as its density at low-to-medium pressures (below its solidification pressure, usually approximately 200 to 400 MPa, or 29 to 58 ksi). Lubricant density plays a role of lesser importance than its viscosity. Above solidification pressure, the lubricant becomes incompressible (under hard EHL), resembling an amorphous solid. Bulk rheological properties of lubricants are discussed here, not the actions of the various additives.

Shear Characteristics of Lubricants. The most important properties of a lubricant are its capability to carry the applied contact load and its viscous shear characteristics, affecting contact friction. In these regards, viscosity is the most prominent of lubricant properties. Viscosity is a measure of the resistance to flow of a fluid. Newton described it as the internal friction of layers of fluid in flow. In 1673, Newton stated that “The shear stress between adjacent fluid layers is proportional to the negative value of the velocity gradient between the two layers.” Thus, viscosity was defined as:

$$\eta = \frac{\tau}{\dot{\gamma}} \quad (\text{Eq 17})$$

where τ is the shear stress, and $\dot{\gamma} = \frac{U}{h}$ is the shear strain rate, where U is the contact velocity, and h is the film thickness. Therefore, for a

Newtonian fluid (according to the previous definition), η is invariable with respect to shear rate and is the slope of the $(\tau - U)$ characteristics, as shown in Fig. 7. However, in practice, the viscosity of most fluids alters with shear rate; some exhibit shear thinning (pseudoplastics), while others become more viscous (dilatants), as shown in Fig. 8. Both of these characteristics deviate from the Newtonian definition of viscosity. Therefore, they are often referred to as non-Newtonian fluids. However, note that non-Newtonian characteristics occur after a certain shear rate, which varies according to the fluid.

Most lubricants used in engine and power train systems act with non-Newtonian shear thinning behavior. In some applications, such as viscolock systems and limited slip differentials (Ref 13, 14), silicone-based fluids shear thin with increased slip of contacting surfaces to reduce friction induced by sliding. Another application in power train systems is the dispersed silica particles in a polyethylene glycol solution, which flocculate (particles come out of solution with increasing shear and raise its viscosity, thus friction). They are used in friction materials, for example, in brake pads.

Most lubricants used in other power train contacts act in a non-Newtonian shear thinning manner. This occurs with increased shear rate and particularly with thin films in highly loaded, concentrated counterformal contacts, as in gears and cam-follower pairs. When the applied shear stress is removed, lubricant

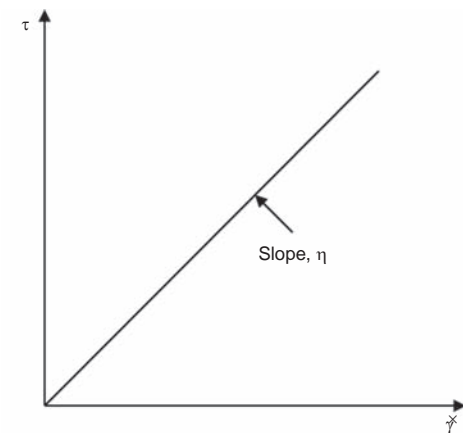


Fig. 7 Shear characteristics of a Newtonian fluid

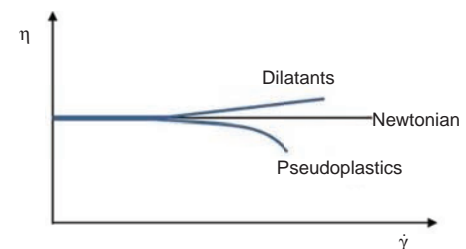


Fig. 8 Newtonian and non-Newtonian shear

viscosity returns almost instantaneously to its original state, such as when it exits the contact. Sometimes a finite period of time (relaxation time) is required for this process to occur. In such cases, the lubricant behavior is referred to as thixotropic. For various lubricants, shear thinning is a function of their composition and occurs when the linearity in $(\tau - \dot{\gamma})$ characteristics is lost (rather similar to the stress-strain relationship after the Hooke's elastic limit for solids at σ_Y). In the case of lubricants, this point is defined as $\tau = \tau_0$, termed as the Eyring shear stress (after Eyring, Ref 15). In terms of shear strain rate, a good guide for the onset of non-Newtonian behavior is $\dot{\gamma} \approx 10^6 \text{ s}^{-1}$ for most engine lubricants and transmission fluids, and $\tau_0 = 2$ to 10 MPa (0.3 to 1.5 ksi). Of course, these values vary with contact pressure and temperature.

Cross (Ref 16) provided a flow curve for dynamic viscosity of pseudoplastics as a function of shear rate:

$$\eta = \eta_\infty + \frac{\eta - \eta_\infty}{1 + \left(\frac{\dot{\gamma}}{\dot{\gamma}_c}\right)^n} \quad (\text{Eq 18})$$

where η is the low-shear viscosity of the lubricant at a particular temperature and pressure, and η_∞ is its limiting high-shear viscosity. For most mineral-based and synthetic engine lubricants, the value of $\eta_\infty = 6.31 \times 10^{-5} \text{ N} \cdot \text{s/m}^2$. $\dot{\gamma}_c$ is the shear rate at which the viscosity is midway between the values η_0 and η_∞ . Also:

$$\dot{\gamma}_c = \frac{1}{\chi} \quad (\text{Eq 19})$$

where χ is the relaxation time of the fluid. In the case of mineral oils, this is approximately 1 ms. The exponent n in Eq 18 can be set to unity.

Variation of Lubricant Viscosity with Pressure and Temperature. Lubricant viscosity alters with pressure. The first expression for this variation was provided by Barus (Ref 17), which is widely used in literature:

$$\eta = \eta_0 e^{\alpha p} \quad (\text{Eq 20})$$

where α is the pressure-viscosity coefficient in m^2/N (sometimes called piezo-viscosity). This is a measure of lubricant response to applied pressure. As Eq 20 shows, it affects the lubricant viscosity in an exponential manner; thus, it is critically responsible for the piezo-viscous behavior (pressurization) of the lubricant. The pressure-viscosity coefficient is itself a function of temperature and the lubricant molecular structure. Its value for most lubricants is in the range of 1.5 to $3.5 \times 10^{-8} \text{ Pa}^{-1}$ (for most engine oils and transmission fluids) at 30°C (85°F) and 1.2 to $2 \times 10^{-8} \text{ Pa}^{-1}$ at 100°C (210°F). The Barus law (Eq 20) yields viscosity values that are fairly accurate for pressures up to the lower levels of solidification pressure of fluids (which is generally in the range of 200 to 400 MPa, or 29 to 58 ksi). It can be seen that the exponent $\alpha p \approx 1$ for contact pressures of the order of 100 MPa (14.5 ksi), where viscosity

nearly triples from its atmospheric value (using Eq 20), $\eta \approx 2.718\eta_0$ (Note: $e \approx 2.718$). At 400 MPa (58 ksi), the viscosity is increased by a factor of $\frac{\eta}{\eta_0} \sim e^4 \approx 2981$, which is clearly impractical. Therefore, for pressures of 200 MPa (29 ksi) and above, alternative relationships should be used. In practice, viscosity of the lubricant increases with pressure by 2 orders of magnitude at the utmost (~ 200 to 300 times), but at the same time it reduces because of the accompanying generated contact heat due to shear. The Barus law is suitable for conforming and most partially conforming contacts, such as journal bearing supports, big-end bearings, and piston rings, where pressures are in lower megapascals, indicating that usually $\alpha p \ll 1$, where the conditions may be considered as iso-viscous (meaning that lubricant piezo-viscous action is not induced, which is really an unfavorable outcome). On the other hand, in counterformal contacts, generated pressures are of the order of gigapascals, and it can be seen that $\alpha p \gg 1$, which means that the lubricant would usually be solidified to an amorphous, incompressible solid (elastohydrodynamic conditions). Any further rise in pressure simply deforms the solid surfaces and extends the footprint dimensions. Typical ball-to-race pressures in main shaft bearings of aero-engines can reach pressures of the order of 2.5 to 3 GPa (0.36 to 0.44×10^6 psi) in take-off conditions, similar to hypoid gears of vehicular differentials of trucks at high loads of the order of 6 to 10 kN (corresponding to contact pressures of the order of 1.5 to 2 GPa, or 0.22 to 0.29×10^6 psi). Higher pressures can also be experienced under impacting conditions. Therefore, piezo-viscous action of the lubricant should be enhanced for some applications, with pressure-viscosity coefficients of the order of 10^{-9} Pa^{-1} . These are high-pressure oils.

Roelands (Ref 18) provided an alternative expression for lubricant-pressure dependence, suitable for medium-to-high pressures experienced in elastohydrodynamic contacts:

$$\eta = \eta_0 e^{A\{[1+B(p-p_0)]^Z - 1\}} \quad (\text{Eq 21})$$

where $A = \ln \eta_0 + 9.67$, $B = 5.1 \times 10^{-9}$, and Z is a lubricant-specific constant, sometimes referred to as a viscosity-pressure index or exponent:

$$Z = \frac{\alpha C_p}{(\ln \eta_0 - \ln \eta_\infty)} \quad (\text{Eq 22})$$

where C_p is a pressure-viscosity coefficient, usually with a value of $\sim 1.96 \times 10^8 \text{ N/m}^2$.

Houpert (Ref 19) extended the expression provided by Roelands to include the effect of temperature:

$$\eta = \eta_0 e^{\left(A \left[\left(\frac{\theta - \theta_0}{\theta_0 - \theta_c} \right)^{-S_0} (1+Bp)^Z - 1 \right] \right)} \quad (\text{Eq 23})$$

where A , B , and Z were provided previously, $C = 138 \text{ K}$, θ is the contact temperature, and θ_0 is the reference or ambient temperature, both in degrees Kelvin. The exponent S_0 is given as:

$$S_0 = \frac{\beta(\theta_0 - C)}{A} \quad (\text{Eq 24})$$

where β is the temperature-viscosity coefficient, usually in the range of 0.005 to 0.05 per degree Kelvin.

For viscosity variation with temperature alone, the originally proposed relationship was provided by Reynolds:

$$\eta = \eta_0 e^{-\beta \Delta \theta} \quad (\text{Eq 25})$$

where the exponent β is the same as that defined previously, $\Delta \theta = \theta - \theta_0$, and η_0 is the dynamic viscosity at the reference temperature, θ_0 .

As for Eq 20, the result of Eq 25 is quite limiting. The most commonly used equation for viscosity variation with temperature is due to Vogel (Ref 20):

$$\eta = a e^{\left(\frac{b}{\theta - c} \right)} \quad (\text{Eq 26})$$

where a , b , and c must be obtained for each lubricant from three sets of supplied data by the lubricant manufacturer. The temperature, θ , is usually in degrees Kelvin. However, if this is used in degrees Celsius, then b as the inherent viscosity-temperature dependence of oil is usually in the range of 500 to ~ 2000 , $c = 50$ to ~ 150 , and a is approximately the density of the lubricant. Alternative expressions for viscosity variation with temperature are also available, such as that by Walther with two parameters. The viscosity of lubricants changes with blending and with the addition of thickening additives, such as polymers and viscosity modifiers. A discussion of these issues is provided in Ref 8.

Variation of Lubricant Density with Pressure and Temperature. Dowson and Higginson (Ref 21) provided an expression for the variation of lubricant density with pressure. This is the most widely used expression for density-pressure dependence:

$$\rho = \rho_0 \left[1 + \frac{0.6 \times 10^{-9}(p - p_0)}{1 + 1.7 \times 10^{-9}(p - p_0)} \right] \quad (\text{Eq 27})$$

This expression was subsequently expanded to take into account the effect of contact temperature (Ref 22):

$$\rho = \rho_0 \left[1 + \frac{0.6 \times 10^{-9}(p - p_0)}{1 + 1.7 \times 10^{-9}(p - p_0)} - c(\theta - \theta_0) \right] \quad (\text{Eq 28})$$

where c is the thermal expansivity of the lubricant, usually in the range of 6.4 to $7.3 \times 10^{-4} \text{ K}^{-1}$.

Variation of Lubricant Thermal Conductivity with Pressure and Temperature. Another important lubricant rheological property is its thermal conductivity, because this affects the ability of the lubricant to convect some of the generated heat away from the contact. Larsson et al. (Ref 23) provided a relationship for the variation of thermal conductivity with pressure:

$$K = a \left(1 + \frac{bp}{1+cp} \right) \quad (\text{Eq 29})$$

where the constants are $a = 0.137$, $b = 1.72$, and $c = 0.54$.

The variation of lubricant thermal conductivity with temperature can be stated as:

$$K = \frac{a}{s} (1 - b\theta) \quad (\text{Eq 30})$$

where, in this case, $a = 0.12$, $b = 1.667 \times 10^{-4}$, and s is the specific gravity at 15.6°C (60.1°F) for temperature θ , in degrees Celsius.

Variation of Lubricant Specific Heat Capacity and Thermal Diffusivity with Temperature. The specific heat capacity at any pressure, p , also varies with temperature. In its simplest form, it is roughly estimated as:

$$c_p = \frac{a + b\theta}{s^n} \quad (\text{Eq 31})$$

where, in this case, $a = 1630$, $b = 3.4$, $n = 0.5$, and s is the specific gravity at 15.6°C (60.1°F) for temperature θ , in degrees Celsius.

For thermal diffusivity:

$$\alpha = \frac{K}{\rho c_p} \quad (\text{Eq 32})$$

Predicting Lubricant Film Thickness

The first step in determining the prevailing conditions is to predict the lubricant film thickness. The usual way is to simultaneously solve the Reynolds equation, the gap shape, and the lubricant rheological state equations for density and viscosity. The gap shape includes any nominal clearance, the undeformed profile of contacting surfaces, and any deformation of contacting surfaces using the elasticity potential equation (Ref 8). However, the method of solution, particularly for counterformal contacts, involves detailed numerical analysis and does not lend itself to a readily accessible outcome to everyone. Fortunately, the results of multiple solutions for various nonconforming contact configurations (an ellipsoidal rigid solid loaded onto a semi-infinite elastic half-space) have been used in regression analyses to provide film thickness formulas for a broad range of operating conditions, based on the dimensionless parameters G^* , U^* , and W^* (see the section "Regimes of Lubrication" in this article).

Figure 9 shows a micrograph of a lubricated line contact of a roller against a flat plane (Ref 24). This image shows half of the footprint shape from the contact center of the roller, with the central lubricant film thickness along the centerline of the contact in the y -direction, h_0 ; x is the direction of lubricant entrainment into the contact from the inlet to the outlet, beyond which the exit flow breaks into cavitation. Figure 10 shows a typical film thickness variation in the x -direction. The minimum film

thickness actually occurs in the vicinity of the outlet.

For counterformal long line contact, the central lubricant film thickness can be predicted as (for rollers, spur gear teeth pairs) (Ref 21):

$$h_0^* = 1.93U^{*0.69}G^{*0.56}W^{*-0.1} \quad (\text{Eq 33})$$

where:

$$h_0^* = \frac{h_0}{R_{zx}}$$

A few important points should be noted. First, the lubricant film thickness is almost independent of load, w (low exponent of W^*). This is a feature of the hard EHL condition (recall the incompressible nature of the lubricant as an amorphous solid). On the contrary, in the case of hydrodynamics, $h \propto w^{-1}$. Secondly, the power indexes for U^* and G^* show that the film thickness is crucially dependent on contact speed and the material combination. Thirdly, all of these equations are obtained with the assumption of a fully flooded (or drowned) inlet condition. Finally, the regime of lubrication is hard elastohydrodynamics. Thus, these equations do not apply to the other modes of fluid film lubrication in Fig. 6.

For a finite line contact, the absolute minimum film thickness at the edges of the contact, h_{sc} , referred to as the side constriction film thickness, should be obtained through numerical analysis. However, the central lubricant film thickness is given as, not considering any lubricant film squeeze (Ref 8, 25):

$$h_0^* = 1.67G^{*0.421}U^{*0.541}W^{*0.059} \quad (\text{Eq 34})$$

Similar observations can be made, as in the case of Eq 33.

For the general case of elliptical point contact, as experienced by ball-to-raceway contact in ball bearings, hypoid and bevel gears, and cam-roller follower contact (Ref 26):

$$h_0^* = 4.31U^{*0.68}G^{*0.49}W^{*-0.073} \left\{ 1 - e^{-1.23 \left(\frac{R_{zy}}{R_{zx}} \right)^{2/3}} \right\} \quad (\text{Eq 35})$$

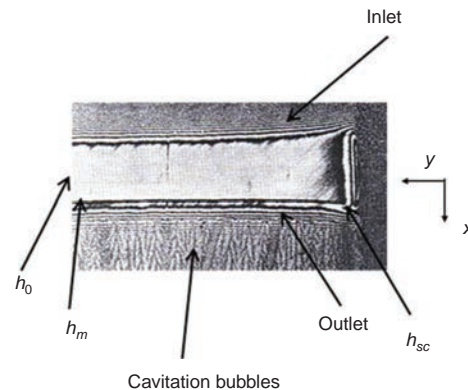


Fig. 9 Micrograph of lubricant film contour for a roller against a flat. Source: Ref 24

where, in this case, the dimensionless parameters are:

$$W^* = \frac{\pi W}{2E^*R_{zx}^2}, U^* = \frac{\pi\eta_0 U}{4E^*R_{zx}}, G_e = \frac{2}{\pi}(E^*\alpha) \quad (\text{Eq 36})$$

Note that for the case of a circular point contact, $R_{zy} = R_{zx} = R$.

Therefore, the lubricant film thickness for various counterformal contacts under EHL can be estimated using Eq 33 to 35. Once the lubricant film thickness is obtained, the Stribeck oil film parameter, λ , can be calculated (Eq 11), if the root-mean-square composite surface roughness is known. For conforming contact of journal bearings, the appropriate lubricant film thickness equation is provided in the section "Piston-Cylinder Conjunctions" in this article.

Surface Topography

Surface roughness is usually measured through the use of various forms of profilometer or spectrometer, either using a contacting stylus running over the rough surface or through imaging of the same. In all cases, the results are magnified and analytical calculations are made to represent the surface by a series of statistical parameters. Traditionally, arithmetic average roughness (R_a) has been used in industry to quantify the roughness of a surface:

$$R_a = \frac{1}{L} \int_0^L |z(x)| dx = \frac{1}{n} \sum_{i=0}^n |z_i| \quad (\text{Eq 37})$$

where, in continuous description, L is the length of the measured surface, and $z(x)$ is the measured profile as a function of length at any point, x . In discrete description, n is the number of measured data points, and z_i is the measured surface height at any given point along the surface. However, the arithmetic average or centerline average is not a good representation of the actual surface roughness. Advances in surface engineering have resulted in the use of various more-pertinent parameters, describing various features of a rough surface. It is better to use the geometric average roughness (R_q , or

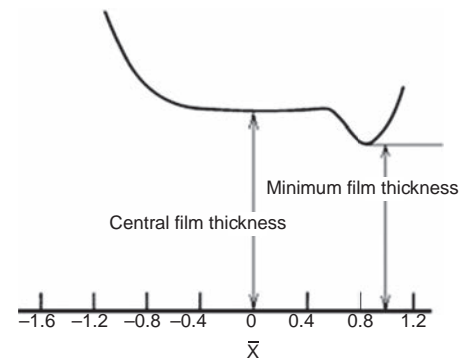


Fig. 10 Lubricant film shape in the direction of rolling

root-mean-square roughness), which is more sensitive to the variations in roughness peaks and valleys:

$$R_q = \sqrt{\frac{1}{L} \int_0^L z^2(x) dx} = \sqrt{\frac{1}{n} \sum_{i=0}^n z_i^2} \quad (\text{Eq 38})$$

As a rule of thumb, for a given surface, usually R_q should be 11% higher than its R_a value.

An improved description of the surface roughness for various engineering applications includes parameters such as core roughness, R_k , reduced peak height, R_{pk} , and reduced valley height, R_{vk} . These parameters are obtained through Abbott-Firestone or bearing area curves and describe the surface texture or roughness structure. With the use of these sets of parameters, one can differentiate between the core roughness (the plateau), the peaks above the plateau, and the valleys between the plateaus, such as for the case of cross-hatched honed surfaces of cylinder liners (Fig. 11). These parameters are suitable for monitoring of running-in wear process (Ref 27, 28). During the process, R_{pk} usually changes dramatically until the roughness profile settles. In addition, it is desirable to reduce the R_{pk} of the surface with harder materials to reduce abrasive friction. The R_{vk} is an estimate of the valleys with the potential to act as reservoirs of trapped lubricant.

Other roughness parameters are also used to determine the nature of the surface roughness distribution, such as skewness and kurtosis.

Skewness determines the symmetry of roughness distribution. Positive values of skewness indicate that the peaks are predominant, while negative values indicate dominance of the valleys:

$$R_{sk} = \frac{1}{R_q^3} \left(\frac{1}{n} \sum_{i=0}^n z_i^3 \right) \quad (\text{Eq 39})$$

Kurtosis is a measure of deviation of the roughness profile from an ideal Gaussian distribution. Kurtosis values higher than 3 indicate that there are excessive high peaks or valleys in the measured profile, while values lower than 3 show a lack of these:

$$R_{ku} = \frac{1}{R_q^4} \left(\frac{1}{n} \sum_{i=0}^n z_i^4 \right) \quad (\text{Eq 40})$$

A surface with Gaussian symmetric roughness distribution has a skewness of 0 and a kurtosis value of 3.

Mathematical models such as the Greenwood and Tripp model (Ref 29) used to estimate boundary friction in mixed or boundary regimes of lubrication are based on the contact of opposing rough surfaces with a Gaussian distribution of roughness peaks. These models make use of other topographical parameters, such as the average asperity peak radius, β , and the area density of asperity peaks at certain separations of the contacting surfaces, κ , as well as the composite surface roughness, σ_s (Eq 11). For σ_s , the arithmetic average or plateau core roughness, R_q or R_k , may be used for run-in surfaces, and during running-in wear with new surfaces, R_{pk} is found to be the most appropriate parameter (Ref 28).

The parameters β and κ are evaluated using spectral moments of the measured roughness profile (Ref 30). For anisotropic surfaces, the spectral moments should be calculated in two preferably perpendicular directions and then combined to provide a representative spectral moment for the surface, based on which values of β and κ can be determined (Ref 31). For isotropic surfaces, the area density of the asperity peaks is calculated as (Ref 32):

$$\kappa = \frac{1}{32.65} \frac{\sum_{i=0}^n (d^2z/dx^2)^2}{\sum_{i=0}^n (dz/dx)^2} \quad (\text{Eq 41})$$

and the average asperity radius as (Ref 33):

$$\beta = \frac{0.665}{\sqrt{\frac{1}{n} \sum_{i=0}^n (d^2z/dx^2)^2}} \quad (\text{Eq 42})$$

All of these parameters can readily be obtained from most postprocessing software installed on modern surface-measuring optical metrology devices. Thus, for two rough contacting surfaces, the composite value becomes:

$$\kappa_c = \frac{1}{2} (\kappa_1 + \kappa_2) \quad (\text{Eq 43})$$

and

$$\frac{1}{\beta_c} = \frac{1}{\beta_1} + \frac{1}{\beta_2} \quad (\text{Eq 44})$$

In practice, the roughness of surfaces is dependent on the material and the processes carried out, such as turning, grinding, honing, and so on. Representative values are provided in Ref 8.

Friction and Power Loss

Viscous Friction. Friction in lubricated contacts is generated due to the shear of a thin lubricant film. The shear stress is obtained as:

$$\tau = \frac{h dp}{2 dx} \pm \frac{\eta \Delta U}{h} \quad (\text{Eq 45})$$

where the first term is shear caused through application of the pressure gradient (lubricant entering into the contact wedge). This is called Poiseuille shear, after the French physiologist who first defined it. This component of shear is dominant at the inlet to a contact and becomes negligible relative to the second term, which is the shear of the lubricant film due to the relative motion of the contiguous surfaces, termed Couette shear, after the French physicist who first defined it. ΔU is the sliding speed (the relative speed of the contacting surfaces, $\Delta U = |(u_1 - u_2)|$). The shear stress in Eq 45 is most appropriate for Newtonian behavior of the lubricant. As the shear rate, $\dot{\gamma} = \frac{\Delta U}{h}$, increases, a limit is reached where any further increase in shear rate has no effect on the shear stress. At this limit, the proportional linearity between shear stress and shear strain rate is lost. The shear stress at this limit is defined as the Eyring shear stress (Ref 15) (see the section "Shear Characteristics of Lubricants" in this article). The shear behavior of the lubricant thereafter is referred to as non-Newtonian (Ref 8). Under non-Newtonian shear, the behavior of the lubricant is visco-elastic, meaning that shear stress is not only dependent on the shear strain rate but also on the shear strain itself (note: for elastic material in shear, $\tau \propto \gamma$, and for viscous media, $\tau \propto \dot{\gamma}$). Thus, $\tau \propto (\gamma, \dot{\gamma})$. Thereafter, with further shear the lubricant enters into a visco-plastic behavior, with a diminutive surface-adsorbed film (Fig. 12). Evans and Johnson (Ref 34) provide charts for shear behavior

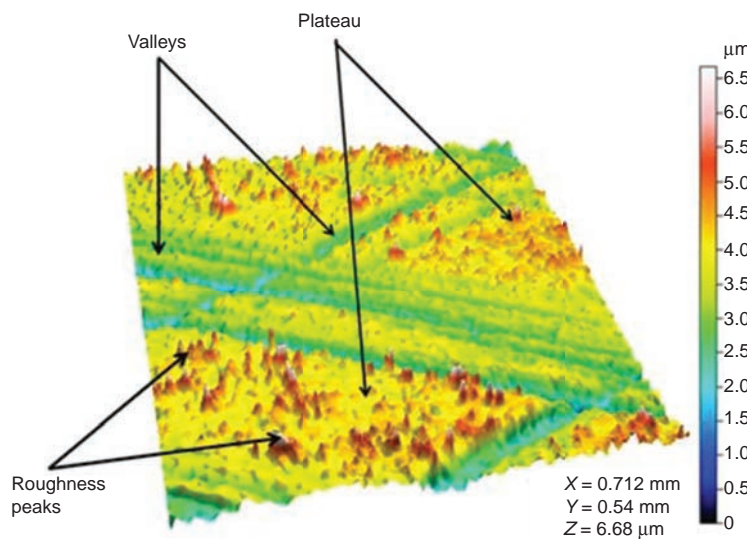


Fig. 11 Typical image of a cross-hatched honed specimen surface using white light interferometry

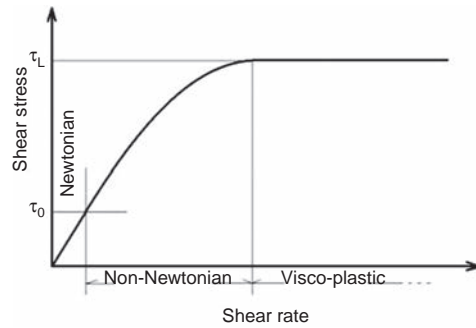


Fig. 12 Typical shear rate characteristics of lubricants

of some lubricants. Because such films are not coherent (continuous in nature), a limiting shear stress is reached as:

$$\tau_L = \tau_{L0} - \varepsilon p_m \quad (\text{Eq 46})$$

where clearly the limiting shear stress is a function of the average or mean pressure, p_m (Pascal pressure as contact load over the area); ε is the pressure-induced shear coefficient, usually in the range of 0.05 to 0.3; and τ_{L0} is the limiting shear stress of the lubricant at atmospheric pressure.

The mathematical definition of friction is:

$$f = \iint \tau dx dy \quad (\text{Eq 47})$$

where the contact occurs in the xy -plane. Therefore, for Newtonian shear, viscous friction becomes (Eq 45):

$$f_v = \iint \left(\frac{h dp}{2 dx} \pm \frac{\eta \Delta U}{h} \right) dx dy \quad (\text{Eq 48})$$

Note that in the contact:

$$\frac{h dp}{2 dx} \ll \frac{\eta \Delta U}{h}$$

With the film thickness, h , obtained from the appropriate equation in the section "Predicting Lubricant Film Thickness" in this article, and the viscosity of the lubricant determined at a given temperature and pressure from the section "Lubricant Rheology," the viscous friction can be obtained if the contact area is known. For the case of nonconforming contacts, these are readily available in Table 1, where for a circular point contact footprint, $A = \pi a^2$; for an elliptical contact footprint, $A = \pi ab$; and for a finite line contact as a rectangular band, $A = 2aL$. For journal bearings, see the section "Piston-Cylinder Junctions" in this article.

For the non-Newtonian fluid film regime of lubrication in the visco-elastic (Eyring) region, an analytical approach is reported by Evans and Johnson (Ref 34), where the coefficient of friction, μ , is given as:

$$\mu = 0.87 \alpha \tau_0 + 1.74 \frac{\tau_0}{p_m} \ln \left[\frac{1.2}{\tau_0 h_0} \left(\frac{2K \eta_0}{1 + 9.6 \xi} \right)^{\frac{1}{2}} \right] \quad (\text{Eq 49})$$

where values for the pressure-viscosity coefficient, α , lubricant thermal conductivity, K , and viscosity, η , for the correct operating conditions can be calculated using the expressions in the section "Regimes of Lubrication" in this article. Also:

$$\xi = \frac{4}{\pi} \frac{K}{h_0 / R_{zx}} \left(\frac{p_m}{E^* R_{zx} K' \rho' c' U} \right)^{1/2} \quad (\text{Eq 50})$$

which is in terms of lubricant thermal conductivity and that of solid surfaces, where ρ' is the density, c' is the specific heat capacity, and K' is the thermal conductivity.

For the visco-plastic region, where interactions occur in the glassy amorphous state:

$$\mu = \frac{\tau_L}{p_m} \quad (\text{Eq 51})$$

Therefore, the viscous component of friction becomes:

$$f_v = \mu W \quad (\text{Eq 52})$$

where W is the contact load.

Prior to any calculation of friction, it is important to establish the regime of traction (shear behavior of the lubricant: Newtonian, non-Newtonian visco-elastic Eyring, or non-Newtonian visco-plastic). For this purpose, Reiner (Ref 35) defined a number, called Deborah's number, De , which is the ratio of relaxation time for a material to adjust to an applied state of stress to a characteristic time of the investigation. In rheology/lubrication, this characteristic time is the time of passage of the lubricant through a given contact conjunction. The relaxation time of a lubricant in shear is $\frac{\eta}{G}$, where G is the shear modulus of the lubricant, which only becomes significant at high pressures (typical of visco-elastic and visco-plastic traction), experienced in elastohydrodynamic contacts (pressures in the gigapascal range). The value of G depends on the maximum generated pressure (Table 1) as (Ref 8):

$$G = 0.1 + 3p_0 \quad (\text{Eq 53})$$

The unit for this equation is gigapascals, and the value of 0.1 GPa represents the onset of lubricant solidification to an incompressible amorphous solid, taken to be 100 MPa (14.5 ksi) in this case for most engine lubricants. p_0 is the maximum Hertzian pressure (Table 1).

The time of passage of the lubricant through a contact depends on the contact footprint dimension in the direction of flow, usually $2a$ (Table 1), and the speed of entrainment of the lubricant into the contact, $U = u_1 + u_2$, or, in other words, $\frac{2a}{U}$. Therefore, the Deborah number becomes:

$$De = \frac{\eta U}{2aG} \quad (\text{Eq 54})$$

Viscosity should be adjusted for pressure, using the Roelands equation (Ref 18) (see the section "Lubricant Rheology" in this article). It can be seen that at high viscosities or high lubricant entrainment speeds, a large Deborah number would result, indicating visco-elastic behavior as the lubricant elastically deforms/shears in flow through the contact. Conversely, at lower viscosities and entraining speeds, the Deborah number is reduced, such as in Newtonian shear, when $De \leq 1$. With more excessive pressures, the lubricant acts as a glassy solid and, depending on the transient value of G , undergoes plastic deformation as well. Like a solid, the stress-strain relationship governs the conditions rather than the stress-strain rate, as predominates for liquids. The limiting shear stress becomes:

$$\tau_L = \gamma G \quad (\text{Eq 55})$$

which can be compared with Eq 46 to get an idea of shear strain under any prevailing conditions, using the data already provided and the value of τ_{L0} for a particular lubricant from its manufacturer.

It should be noted that as either of the contact pressures or the speed of entraining motion increases, the generated temperatures reduce the lubricant viscosity; thus, the lubricant shear stress falls in all the stated regimes of traction.

Boundary Friction. When the thickness of the lubricant film is comparable to the roughness of the contacting surfaces, interactions occur between the asperities of the rough counterface surfaces. The regime of lubrication is no longer fluid film lubrication but mixed or boundary regimes of lubrication (see the section "Regimes of Lubrication" in this article). The result is increased friction through boundary interactions (i.e., boundary friction). There are two main mechanisms for this form of friction. One is welding of the opposing asperities under pressure into junctures that must be broken for the continuance of relative motion of the contacting surfaces. The resisting mechanism is called adhesive friction. In the process, globules of the softer of the two surfaces may be broken and stick to the harder one or contaminate the lubricant, resulting in adhesive wear. There is also friction due to tribochemical reaction layers formed on the contiguous surfaces in contact. The other mechanism of boundary friction is the oblique contact of opposing asperities with those of the harder surface, deforming those of the softer counterface. The resistance introduced by these asperities in deformation (elastically or plastically) accounts for a mechanism referred to as deformation or plowing friction. This can lead to a host of wear mechanisms, mainly abrasive wear and scuffing. Similar counterface materials should be avoided due to solid solubility of similar materials and bonding under pressure. As a guideline, a mixed regime of lubrication occurs when $1 < \lambda < 3$, and a boundary regime of lubrication occurs when $\lambda \leq 1$.

When mixed or boundary regimes of lubrication occur, the Greenwood and Tripp (Ref 29) model is often used to calculate boundary friction. The method assumes a Gaussian distribution of surface asperity heights that introduces some limitations along with real rough engineering surfaces (see the section "Surface Topography" in this article) (Ref 36).

Under a mixed regime of lubrication, a portion of the contact load is carried by the opposing asperities of the counterface surfaces as a function of their separation (Ref 29):

$$W_a = \frac{16\sqrt{2}}{15} \pi (\kappa\beta\sigma_s)^2 \sqrt{\frac{\sigma_s}{\beta}} E^* A F_{5/2}(\lambda) \quad (\text{Eq 56})$$

where β is the average asperity tip radius, σ_s is the composite root-mean-square surface roughness, and κ is the asperity distribution per unit scanned/measured area. The product $\kappa\beta\sigma_s$ is known as the roughness parameter, usually in the range of 0.03 to 0.07, and $\sigma/\beta = 10^{-4}$ to 10^{-2} (a measure of asperity slope) for many engineering surfaces. This parameter affects the adhesion of surfaces, as described by Fuller and Tabor (Ref 37). $F_{5/2}(\lambda)$, a statistical function, can be represented by a polynomial fit for ease of application:

(see Eq 57 at bottom of page)

The load-carrying capacity of the lubricant film is obtained as $W_h = \iint p dx dy$, which is usually obtained through numerical solution of the Reynolds equation to find the pressure distribution, $p(x, y)$. Analytical solutions are available for simple cases, many of which are provided in Ref 8. Therefore, the applied load, W , is carried by the lubricant film as well as a portion (usually small) of directly contacting opposing asperities; hence, $W = W_h + W_a$.

Similarly, the asperity contact area at any given separation of surfaces, represented by the Stribeck's oil film parameter, λ , is (Ref 29):

$$A_a = \pi^2 (\xi\beta\sigma)^2 A F_2(\lambda) \quad (\text{Eq 58})$$

The statistical function $F_2(\lambda)$ is:

(see Eq 59 at bottom of page)

Boundary friction can be evaluated as:

$$f_b = \tau_L A_a + \zeta W_a \quad (\text{Eq 60})$$

where ζ is analogous to the asperity-scale adhesive coefficient of friction. It is also assumed that pockets of lubricant trapped in the valleys between asperity peaks and any ultrathin adsorbed film at the summit of asperities would shear at the limiting shear stress. Therefore, with measured values of roughness parameter and asperity slope, as well as using an

appropriate film thickness equation (see the section "Predicting Lubricant Film Thickness" in this article), boundary friction can be predicted.

Total Friction and Power Loss. The total generated contact friction is the result of combined viscous and boundary friction components, thus:

$$f = f_v + f_b \quad (\text{Eq 61})$$

The frictional power loss is obtained as:

$$P_f = f \Delta U \quad (\text{Eq 62})$$

where ΔU is the sliding velocity (relative speed of the contacting surfaces).

It is important to evaluate the frictional power loss to ascertain the efficiency of the power train, such as engine subsystems, engine bearings, the piston-cylinder system, and the valve train system.

Piston-Cylinder Conjunctions

The ring pack usually uses three to four rings, comprising one to two compression rings, an oil control ring, and a scraper ring. In general, the losses are nearly equally shared between the piston skirt and the ring pack. Among the piston rings, the tribology of the top compression ring is the most complex, because its main function is to seal the combustion chamber, thereby guarding against flow of combustion gases and soot into the crankcase. This phenomenon is called blow-by. Therefore, good sealing action of the compression ring(s) is required to mitigate blow-by and any loss of power. As a result, there is a diminished gap between the compression ring contact face and the liner or bore surface, promoting increased friction. In an engine cycle comprising in-take, compression, combustion (or power), and exhaust strokes (for a four-stroke engine), the compression ring is subjected to a transient regime of lubrication, which usually includes boundary, mixed, and hydrodynamic conditions. Hydrodynamics often dominate in all the engine strokes, except at piston reversals at the top and bottom dead centers of piston motion. Here, because of a change in the direction of sliding velocity, there is momentary cessation of entraining motion of the lubricant, $U = 0$; thus, no lubrication occurs, except for any entrapped pockets of lubricant on the rough surface topography or lubrication through squeeze-film motion. Mixed or boundary regimes of lubrication are prevalent in these parts of the cycle and also near the top dead center in the compression stroke and in a large part of the power stroke because of the higher contact

pressures. It is rare that the contact pressures (usually of the order of several megapascals) would cause any local deformation of contacting surfaces or significantly cause piezo-viscous action of the lubricant. Thus, elastohydrodynamic conditions are not usually encountered, although some have claimed otherwise in the literature.

Due to the complexity of the ring-liner conjunction, numerical solutions of the Reynolds equation or even more general solutions of the Navier-Stokes equations are used. Readers can refer to such solutions in Ref 38 to 42. However, there have been some analytical solutions that, although idealized to a certain extent, provide good estimates of prevailing conditions. Furthermore, for a handbook with a broad readership, such solutions are more appropriate. The piston-skirt contact with the cylinder liner is usually subject to an elastohydrodynamic regime of lubrication, with thermoelastic deformation of the skirt being a function of skirt structure and contact profile. No accurate analytical solution is possible. Readers are referred to Ref 43 to 45.

To obtain any analytical solution, certain simplifying assumptions are made. In this case, it is assumed that the ring conforms perfectly to the surface of the liner, which is considered to be an ideal right circular cylinder. In practice, the ring is subject to modal deformation when subjected to varying contact loads and friction during an engine cycle. It deforms in its radial plane (in-plane motion) as well as in the axial direction of the piston (out-of-plane motions, including twist and flutter motions) (Ref 46–48). Lubricant film thickness is formed along the ring contacting face in the axial direction of motion. The film thickness is generally quite thin (tenths of a micrometer to a few micrometers). Therefore, one can assume no side leakage of the lubricant from its axial flow into the peripheral direction of the cylinder. These assumptions mean that a one-dimensional solution of the Reynolds equation can be undertaken, which lends itself to an analytical approach:

$$\frac{\partial}{\partial x} \left(h^3 \frac{\partial p}{\partial x} \right) = 6\eta_0 U \frac{\partial h}{\partial x} + 12\eta_0 \frac{\partial h}{\partial t} \quad (\text{Eq 63})$$

where η_0 is the dynamic viscosity of the lubricant at atmospheric pressure, assumed here because, as shown in the section "Variation of Lubricant Viscosity with Pressure and Temperature" in this article, with low-to-medium hydrodynamic pressures, $\eta \sim \eta_0$. However, η_0 should be adjusted for the contact temperature to better represent the prevailing conditions. In the case of ring-cylinder liner contact, Morris et al. (Ref 40) have shown that liner temperature can be safely used for this

$$F_{5/2} = \begin{cases} -0.004\lambda^5 - 0.057\lambda^4 - 0.29\lambda^3 - 0.784\lambda^2 - 0.784\lambda - 0.617 & \text{for } \lambda < 2.5 \\ 0; & \text{for } \lambda \geq 2.5 \end{cases} \quad (\text{Eq 57})$$

$$F_2(\lambda) = \begin{cases} -0.002\lambda^5 - 0.028\lambda^4 - 0.173\lambda^3 + 0.526\lambda^2 - 0.804\lambda - 0.500 & \text{for } \lambda < 2.5 \\ 0; & \text{for } \lambda \geq 2.5 \end{cases} \quad (\text{Eq 59})$$

purpose. The film thickness, h , is required for the solution of the Reynolds equation (Eq 63), to obtain the pressure distribution, p . The film thickness is a function of the ring contact face profile and the nominal minimum gap (designed clearance), h_0 , as:

$$h = h_0 + S(x) \quad (\text{Eq 64})$$

where $S(x)$ is the axial profile of the ring, as shown in Fig. 13.

A parabolic ring face profile is shown in Fig. 13. However, in practice there are many forms of ring face profile, some fairly flat with chamfered edges. In most cases, the profile can be approximated with the following expression:

$$S(x) = C \left(\frac{x \pm o}{b/2 + o} \right)^m \quad (\text{Eq 65})$$

where $m = 2$ yields a parabolic profile, C is the ring crown height, o is any crown offset from the center, with “+” shifting the vertex of the parabola toward its left side and “-” shifting the vertex of the parabola toward its right side, as shown in Fig. 13. The crown height and the parabolic shape alter with running-in and gradual wear, as shown by Rahmani et al. (Ref 39). The profile of a run-in ring was measured, showing a complex mix of various orders of m . Morris et al. (Ref 27) provide analysis for rings with different values of m .

For a parabolic contact face profile:

$$S(x) = \frac{1}{2R} (x \pm o)^2 \quad (\text{Eq 66})$$

where the radius of curvature of the ring profile is:

$$R = \frac{1}{2C} (b/2 + o)^2 \quad (\text{Eq 67})$$

It is now possible to solve the Reynolds equation (Eq 63) analytically by integrating it twice with respect to x . Two boundary conditions are required to find the integration constants. These boundary conditions are based on the assumed inlet and outlet of the contact. They also depend on the sense of motion of the ring: upstroke toward the top dead center or downstroke toward the bottom dead center. In the case of the former, the pressure at the inlet is the combustion chamber pressure, and in the case of the

latter, the inlet pressure is the interring pressure (between the compression and the oil control rings). For simplicity, this is often taken to be the crankcase pressure at approximately the atmospheric pressure.

The lubricant is drawn into the contact in the opposite sense to the ring sliding motion. This means that in the upstroke sense of the ring, the inlet meniscus would be in the ring converging profile adjacent to the combustion chamber, where one would expect starved conditions. In the downstroke sense of the ring, the inlet is fed by any volume of surface lubricant between the oil control ring and the lower side of the compression ring. Most analytical solutions assume a fully flooded inlet. This means that the inlet distance to the center of the contact can be assumed to be theoretically at infinity, $x_i \rightarrow -\infty$; the negative sign is assigned to the inlet, where the center of the contact is at $x = 0$. In practice, the fully flooded inlet is at a distance of 11.298 times the half-width of the sliding contact (in the case of a ring of face width b , the half-width can be taken as $b/2$). The general case of rolling and sliding contacts is shown by Birkhoff and Hays (Ref 49). Therefore, at least for most of the ring upstroke motion, the inlet to the contact would be partially starved. Tipei (Ref 50) shows that even with an assumed drowned inlet, some of the inward flow into the converging wedge of the contact is subject to counter and swirl flows (observed experimentally in Ref 49). Under these conditions, the actual inlet occurs at the stagnation point, with no reverse and swirl flows thereafter, where:

$$\left. \frac{\partial p}{\partial x} \right|_{x=x_i} = 2 \frac{\eta_0 U}{h_i^2} \quad (\text{Eq 68})$$

Integrating the Reynolds condition (Eq 63) once and using Eq 67, the position of the inlet meniscus can be determined. Tipei (Ref 50) favored Reynolds (Ref 51) or Swift (Ref 52)-Stieber (Ref 53) boundary conditions, based on the observation of the inlet flow field prior to the stagnation point, where, in addition to condition (Eq 68), $p = p_c$. On the other hand, the zero reverse flow inlet boundary conditions of Prandtl-Hopkins (Ref 54) require zero lubricant film velocity as well as its gradient to determine the stagnation boundary. It is shown that the pressure gradient at the actual inlet in this case becomes (Ref 55):

$$\left. \frac{\partial p}{\partial x} \right|_{x=x_i} = 4 \frac{\eta_0 U}{h_i^2} \quad (\text{Eq 69})$$

Using this condition in conjunction with the Reynolds equation determines the inlet boundary to be at (Ref 55):

$$x_i = -2.623 \sqrt{R h_0} \quad (\text{Eq 70})$$

For outlet (exit) boundary conditions, the most widely used are the Swift-Stieber conditions (Ref 53), where:

$$p = p_c, \quad \left. \frac{\partial p}{\partial x} \right|_{x=x_c} = 0 \quad (\text{Eq 71})$$

where x_c is the lubricant film rupture point, beyond which cavitation occurs, and p_c is the cavitation vaporization pressure, assumed to be the atmospheric pressure in most applications. The inlet rupture point is then obtained as:

$$x_c = 0.672 \sqrt{R h_0} \quad (\text{Eq 72})$$

With solution of the Reynolds equation (Eq 63), using any assumed inlet and outlet boundary conditions, the pressure distribution is obtained, together with the inlet and outlet distances from the center of the contact, x_i and x_c . Gohar (Ref 55) provides a solution that, with the use of Eq 72 and with zero inlet pressure for a fully flooded inlet for a sliding contact, becomes:

$$p = \frac{\eta_0 \Delta U r_0 \sqrt{R}}{h_0^{3/2}} \left[\frac{8.1724}{\left(1 + \frac{x^2}{2R h_0}\right)} + 1.5133 \frac{x}{\sqrt{R h_0}} \tan^{-1} \left(\frac{x}{\sqrt{2R h_0}} \right) + 3.3617 \right] \quad (\text{Eq 73})$$

where r_0 is the bore radius. To account for the inlet and outlet pressures, the result of Eq 73 can be supplemented by the average hydrostatic pressure caused by these pressures:

$$p_{av} = \frac{P_i + P_o}{2} \quad (\text{Eq 74})$$

The sliding speed of the ring is ΔU . This is a function of the engine stroke/compression ratio and rotational speed of the crankshaft and its higher harmonics (Ref 56). For contributions up to second engine order (first harmonic of engine speed):

$$\Delta U \approx \omega r_c \left(\sin \omega t + \frac{r_c}{l} \sin 2 \omega t \right) \quad (\text{Eq 75})$$

where l is the connecting rod length, r_c is the crank-pin radius, and ω is the engine speed in rad/s.

The outlet boundary conditions determine the position of lubricant film rupture (start of the cavitation zone). The most commonly used one is the Reynolds or Swift-Stieber boundary condition. Now, the hydrodynamic load-carrying capacity can be obtained as:

$$W_h = 2\pi r_0 \int_{-x_i}^{x_c} p dx \quad (\text{Eq 76})$$

Using a fully flooded inlet and Swift-Stieber exit boundary conditions, the hydrodynamic reaction becomes (Ref 55):

$$W_h = 15.394 \frac{\eta_0 r_0 R U}{h_0} \quad (\text{Eq 77})$$

This should equate the applied contact force due to gas pressure acting behind the inner

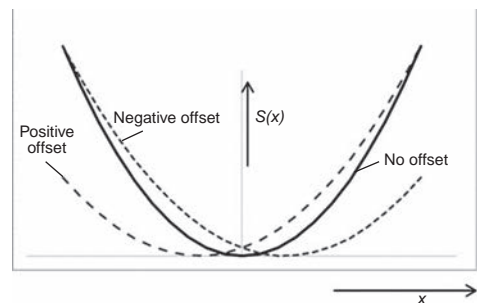


Fig. 13 Parabolic ring contact face profile

rim of the ring and the ring tension force, both of which strive to push the ring against the surface of the liner. For a ring of contact face width b , the ring tension force is:

$$F_t = p_e r_0 b \quad (\text{Eq 78})$$

where the elastic pressure, p_e , is:

$$p_e = \frac{GEJ}{3\pi b a^4} \quad (\text{Eq 79})$$

where a is the ring thickness, G is the free end gap of the ring, and EJ is its flexural rigidity. Note that the ring is incomplete with a gap, which in its free state (unfitted) is usually 2 to 8 mm (0.08 to 0.32 in.). When fitted in situ, the end gap is of the order of a few tenths of a millimeter, which, under applied gas pressure, reduces to a few micrometers.

There are various empirical and analytical methods for determining the gas pressure loading of the compression ring. Assuming that 100% of the gas pressure acts behind the fully conformed ring to the surface of an idealized right circular cylindrical liner, the gas pressure loading becomes:

$$F_g = 2\pi r_0 b P_G \quad (\text{Eq 80})$$

where P_G is the combustion chamber pressure, which is measured using a pressure sensor inserted into the chamber.

Thus, any load carried by the direct interaction of surfaces can be obtained, without the need to determine the film thickness and the Stribeck's oil film thickness parameter:

$$W_a = F_t + F_g - W_h \quad (\text{Eq 81})$$

Gohar (Ref 55) also shows that:

(see Eq 82 at bottom of page)

Therefore, knowing the value of x_c , viscous friction can be obtained easily. For starved conditions, the lower limit, $-b/2$, can be replaced by x_f from Eq 70.

Film thickness can be calculated using Eq 64. Then, the procedure in the sections "Boundary Friction" and "Total Friction and Power Loss" in this article can be followed to find the boundary friction contribution, f_b , and thus the total generated friction.

The rather simple procedure described here yields quite acceptable predictions. Gore et al. (Ref 57) measured the friction of compression ring-cylinder liner contact under motorized condition for a high-performance single-cylinder motocross motorbike engine, using a floating liner. The predictions of their analytical model, quite similar to the one reported here,

agreed well with the measurements, within an error of 5 to 10%.

Engine Bearings

Journal bearings are used in many areas of engine and drive train systems. They are used as connecting rod bearings, crankshaft and camshaft support bearings, and in several other applications. Figure 14 shows a typical circular bore journal bearing. This is typical of crankshaft support bearings. Connecting rod (big-end) bearings are not usually of circular configuration. They mostly have an elliptical bore/bushing, commonly referred to as "lemon-shaped" bearings. In fact, bushing or shell shapes of higher-order out-of-roundness also exist. The reason for these out-of-round bushings can be appreciated by noting that the eccentric position of the journal center within the bushing, along the instantaneous line of centers (LOC), e , creates a pressurized wedge (region), shown by the film thickness, h , in Fig. 14. Directly opposite to the high-pressure region, a divergent gap emerges and therefore is often unloaded with no coherent film of lubricant. This is the cavitation area, with a mix of lubricant, air, and vapor. The angle, ϕ , is measured from the LOC. The minimum clearance occurs at $\phi = 180^\circ$ and the maximum clearance at $\phi = 0^\circ$. The film thickness around the journal is:

$$h = c + e \cos\phi \quad (\text{Eq 83})$$

Therefore, the minimum and maximum film thickness values are:

$$h_{\min} = c - e \text{ and } h_{\max} = c + e \quad (\text{Eq 84})$$

where c is the nominal designed clearance and is usually a fraction of the journal radius (termed the clearance ratio), typically in the range of $\frac{1}{1000} \leq \frac{c}{R} \leq \frac{1}{100}$, where $R = \frac{D}{2}$. The upper limit corresponds to most commercial and mass-manufactured road vehicles, and the lower limit is that of high-performance racing vehicles, where higher pressures are generated in smaller clearances with higher applied loads.

The clearance, c , is an important design parameter, as well as the eccentricity, e . Therefore, the film thickness can be expressed as $h = c(1 + \epsilon \cos\phi)$, where the eccentricity ratio is:

$$\epsilon = \frac{e}{c} \quad (\text{Eq 85})$$

The minimum film thickness becomes $h_{\min} = c(1 - \epsilon)$, and when $\epsilon \rightarrow 1$, direct

contact of surfaces takes place. Therefore, determination of ϵ in a design process is essential. Conversely, when $\epsilon \rightarrow 0$, no high-pressure wedge would be formed. The loss of lubricant reaction means that under certain rotational conditions the phenomenon known as journal whirl would occur, leading to impact of the journal with the bearing bushing.

Referring to Fig. 14, if the journal was stationary with no applied load, it would fall under its own weight, W , onto the bushing surface until the lubricant film reaction, W_z , equilibrates its weight, with the LOC being vertical. The lubricant film would be sustained under pure squeeze condition, with the surfaces finally coming into contact. With journal rotation, the LOC becomes inclined to the vertical, because any applied load is supplemented by the effect of generated friction torque. There are now two components of lubricant reaction: W_z along the LOC and W_x perpendicular to it, with W being their resultant, making an angle, ψ (attitude angle), with the LOC, which is also a function of ϵ as:

$$\psi = \tan^{-1} \left[\frac{\pi}{4} \left(\frac{\sqrt{1 - \epsilon^2}}{\epsilon} \right) \right] \quad (\text{Eq 86})$$

Therefore, it is important to determine ϵ in any design or analysis. To do so in a generic manner, use should be made of a chart, known as the Sommerfeld chart (Fig. 15).

The Sommerfeld number is defined as:

$$\Delta = \frac{W}{\eta \omega R L} \left(\frac{c}{R} \right)^2 \quad (\text{Eq 87})$$

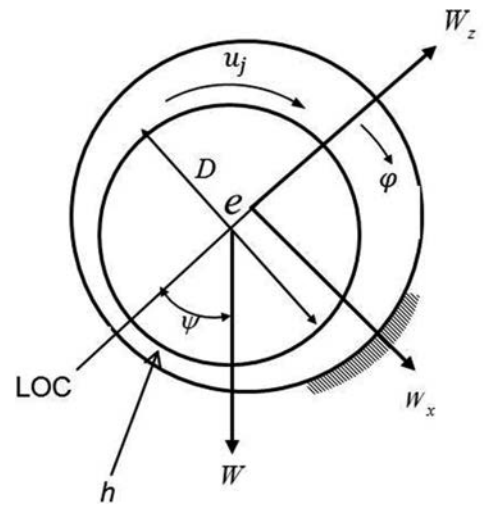


Fig. 14 Circular bore journal bearing. LOC, line of center

$$f_v = \begin{cases} \frac{\eta_0 \Delta U}{h_0} \left[\begin{array}{l} 2\sqrt{2Rh_0} \tan^{-1} \left(\frac{x}{\sqrt{2Rh_0}} \right) - 1.839\sqrt{2Rh_0} \tan^{-1} \left(\frac{x}{\sqrt{2Rh_0}} \right) \\ -1.839(2h_0R) \frac{x}{(2Rh_0 + x^2)} \end{array} \right]_{-b/2}^{x_c} & \text{for: } -\frac{b}{2} \leq x < x_c \\ \frac{\eta_0 \Delta U}{h_0} \left[\sqrt{2Rh_0} \tan^{-1} \left(\frac{x}{\sqrt{2Rh_0}} \right) \right]_{x_c}^{b/2} & \text{for: } x_c \leq x \leq \frac{b}{2} \end{cases} \quad (\text{Eq 82})$$

The Sommerfeld number is another important design parameter based on the intended operating conditions: load, W ; clearance ratio; journal speed, $\omega = 2\pi N$ (N in rev/s); and lubricant dynamic viscosity, η , which should be adjusted for temperature (see the section “Variation of Lubricant Viscosity with Pressure and Temperature” in this article). L is the width of the bearing, sometimes referred to as its length in engine parlance, because it is along the axial direction of the crankshaft.

The various curves in Fig. 15 correspond to the ratio L/D . When $L/D \leq 0.5$, the bearing is termed a short-width bearing, and long-width bearing is when $L/D \geq 2$. Most crankshaft engine bearings fall between the two, and those in the camshaft system are near enough to short-width configuration. Thus, the Sommerfeld chart provided here suffices for the purpose of any engine bearing analysis. Therefore, for given operating conditions—load, speed, and temperature—for a given bearing geometry and lubricant used, the Sommerfeld number can be calculated, and for a given L/D ratio, the eccentricity ratio is obtained using the chart in Fig. 15. Film thickness can then be obtained using Eq 85 as well as the attitude angle from Eq 86. Then, the following relationships provide viscous friction, f_v , friction torque, T_f , power loss, P_1 , and minimum required flow rate to guard against starvation, Q , as (Ref 8):

$$f_v = \pm \frac{c\varepsilon}{2R} W \sin \psi + 2\pi\eta\omega R^2 \left(\frac{L}{c}\right) \frac{1}{\sqrt{1-\varepsilon^2}} \quad (\text{Eq 88})$$

$$T_f = \pm f_v R \quad (\text{Eq 89})$$

$$P_1 = T_f \omega \quad (\text{Eq 90})$$

$$Q_1 = \omega R L c \varepsilon \quad (\text{Eq 91})$$

The negative signs in the aforementioned equations correspond only to the direction of friction on the opposing surfaces. A typical design evaluation process would involve calculation of the Sommerfeld number for a range of loads and engine speeds for a given bearing geometry and lubricant choice. In each case, the eccentricity ratio is obtained for various values of the Sommerfeld number. For good design, an

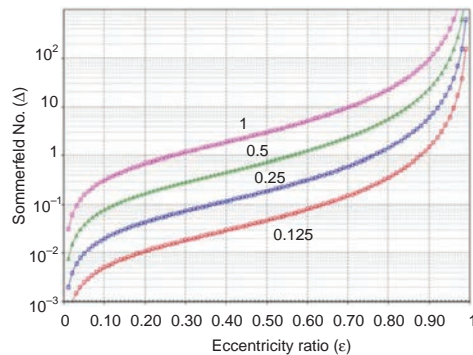


Fig. 15 Sommerfeld chart

eccentricity ratio in the range of $0.65 \leq \varepsilon \leq 0.85$ is sought. The film thickness is then calculated using Eq 85, which ideally should be $\lambda = \frac{h}{\sigma_s} \geq 3$ (Eq 11) to attain hydrodynamic lubrication. If the film thickness falls below the stated condition, then some degree of surface interactions would occur, leading to their wear. The section “Boundary Friction” in this article deals with the general case of boundary friction. In the case of journal bearings, the surface roughness can either be regarded as isotropic/uniform, transversal (along the bearing circumference), or longitudinal (along the bearing width). Then, the film thickness is expected to comprise a nominal value, h (Eq 85), and a stochastic contribution, h_s , based on the surface roughness pattern. The Reynolds equation is then solved numerically to include the effect of roughness (Ref 58). Solutions for elliptic bore big-end bearings are provided in Ref 58 and 59. To mitigate direct contact of surfaces, and thus wear, it is important to enhance the load-carrying capacity of the contact. One approach is to pump pressure into the contact, causing additional hydrostatic effect. Another approach is to use bearings with elliptic or higher-order bores (bushing), where instead of one high-pressure zone, several high-pressure wedges are formed, where the film thickness for the usual elliptic bore connecting rod bearing is of the form:

$$h = c(1 + G \cos^2 \varphi + \varepsilon \cos \varphi) \quad (\text{Eq 92})$$

where $G = \frac{a-b}{c}$, where a and b are the elliptic bore semimajor and semiminor half-widths, and $0 \leq G \leq 1$ is the degree of noncircularity, with $G = 0$ representing a circular bore bearing. The effect of G on improving film thickness can be seen in Fig. 16.

Another approach is to use an overlay on the bearing bushing that comprises low elastic modulus layers (e.g., copper, Babbitt), which would locally elastically deform, thus creating an additional gap for the lubricant to run into as well as a thin, hard layer (e.g., indium, bismuth) to resist wear when surfaces come into

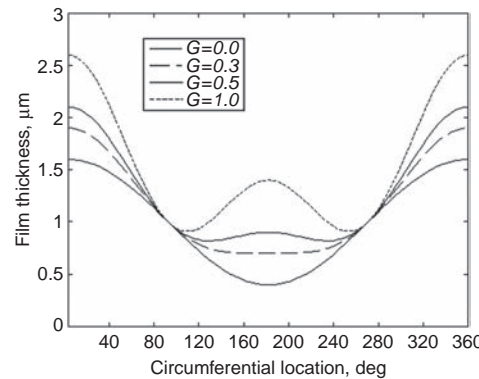


Fig. 16 Improved film thickness with elliptic bore bearings. Source: Ref 58

contact (e.g., under start-up conditions). These issues are discussed in Ref 8 and 11.

Now, returning to the issue of journal whirl, the lubricant film acts like a mass-spring system. Its natural frequency is of the form (Ref 8):

$$\omega_0 = n \sqrt{\frac{g}{c}} \quad (\text{Eq 93})$$

where g is the gravitational acceleration (9.81 m/s^2), and c is the bearing clearance. The constant $n \approx 1.3$ for $0.1 \geq \varepsilon \geq 0.65$ and $n \approx 2$ for $\varepsilon > 0.7$. Clearly, when the speed of crankshaft rotation $\Omega \rightarrow \omega_0$, resonant conditions occur, known as synchronous whirl, where the center of the rotating journal with speed Ω commences to rotate around the center of the bushing at speed ω_0 . This problem can be overcome by quickly speeding up or down through this condition. Another condition is when $\Omega \approx 2\omega_0$ for a horizontal bearing (such as the crankshaft support bearings), whereupon the lubricant entraining speed into the contact ceases; thus, there is no lubricant reaction. The unloaded journal can spin and strike the bushing, causing catastrophic failure (Ref 8). To avoid this condition, a series of numerical analyses are carried out to obtain a whirl stability chart, shown in Fig. 17. The ordinate in the figure is the dimensionless stability factor $\frac{g}{c\Omega^2}$. It can be seen that as the speed decreases, the value of the factor increases, affording better stability for a given bearing clearance, c . The area under the curve includes all conditions that lead to the diminution of lubricant entrainment into the convergent high-pressure zone, ascertained through numerical solution of the Reynolds equation. An increasing ε corresponds to increased loading, thus higher pressures and improved load-carrying capacity. It is clear that journal bearings operate in a more stable manner at higher eccentricity ratios, $\varepsilon \geq 0.7$, for a broader range of engine speeds, Ω (above and to the right of the demarcation line). However, as the eccentricity ratio increases, there is a tendency for thinner films, which can lead to direct contact of surfaces, causing wear, scuffing, and scoring due to excess heat generation. Therefore, for given operating conditions, the

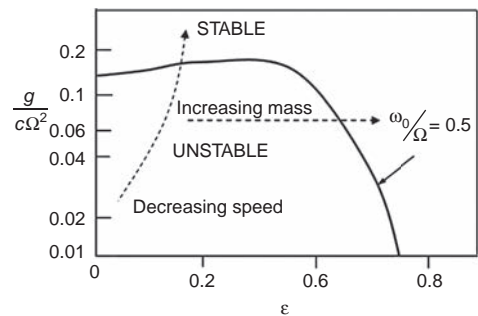


Fig. 17 Journal bearing whirl stability chart. Source: Ref 8

eccentricity ratio can be obtained from the Sommerfeld chart in Fig. 15 and used in Fig. 17 to ascertain the state of system stability.

Cam-Tappet Contact

Unlike journal bearings (conforming) and ring-liner contact (partially conforming), the cam-follower contact is nonconforming (counterformal) and subject to high loads and generated pressures (see the section “Contact Configuration” in this article). The regime of lubrication is mostly elastohydrodynamics (see the section “Regimes of Lubrication”); thus, the section “Contact Fatigue” is directly applicable in this case. Furthermore, for all intents and purposes, Hertzian conditions can be used to obtain an estimate of generated pressures, contact footprint dimensions, and elastic deflection of contacting surfaces in the medium-to-highly loaded parts of the cam-follower cycle (Table 1). For most cam-follower contacts, the medium-to-high contact loads occur between the valve opening and closing points in cam traverse through the cam nose position. However, an elastohydrodynamic analysis is required to determine the lubricant film thickness (Hertzian theory is for dry contacts), friction, and power loss.

Figure 18 shows a typical automotive polynomial cam. It comprises a base circle with a slight clearance, so that the valve remains closed between the valve closing and its opening point, marked on the cam profile. The valve lift profile from the start of the flank to the cam nose position is shown in the inset to the figure. The polynomial lift function, s , is of the form (Ref 8):

$$s = s_{\max} + C_2\bar{\theta}^2 + C_m\bar{\theta}^m + C_p\bar{\theta}^p + C_q\bar{\theta}^q + C_r\bar{\theta}^r; \quad \bar{\theta} = \theta/\hat{\theta} \quad (\text{Eq 94})$$

where s_{\max} is the maximum lift at the cam nose; θ is the cam-lift angle, commencing at the start of the flank; and $\hat{\theta}$ is the maximum cam-lift angle at the position of the cam nose. The exponents of θ are the polynomial powers, shown for three different polynomial cams in the key to the inset in Fig. 18. The polynomial coefficients are obtained as:

(see Eq 95 at bottom of page) where m , p , q , and r are the polynomial coefficients in a polynomial of the form given in Eq 94.

The flank profile for valve closure in most cases is quite similar to the opening one. However, this need not be the case. The profile

is essentially designed in line with engine thermodynamic requirements. Therefore, as far as tribological assessment is concerned, the cam profile is a given, unless the preliminary design, based on engine thermodynamics, results in excess friction and wear and thus poor reliability of the valve train system.

There are many types of valve train systems. Here, the case of an overhead cam-tappet (a flat follower) is considered. Figure 19 shows a direct overhead cam valve train system, with an automotive cam-flat tappet arrangement.

The contact film thickness, h , is usually of the order of a few tenths of a micrometer to a few micrometers (utmost), formed due to the localized elastic deformation of contacting surfaces, δ . The valve is subjected to lift, s , velocity, v , and acceleration, a , as the cam rotates. The valve lift profile is so chosen to comply with the required engine breathing and, at the same time, results in a continuous (nonjerky) acceleration profile. Figure 19 also shows the typical variations of s , v , and a for a polynomial cam, where s is given by Eq 94 and 95 for the polynomial cam. The valve velocity is:

$$v = \frac{\partial s}{\partial t} = \dot{s} \quad (\text{Eq 96})$$

As the cam rotates with an angular velocity, ω , there is an accelerative curvature change, termed geometrical acceleration:

$$j_0 = \frac{\partial^2 s}{\partial \theta^2} = \frac{d^2 s}{d(\omega t)^2} = \frac{1}{\omega^2} \frac{d^2 s}{dt^2} = \frac{1}{\omega^2} a \quad (\text{Eq 97})$$

Thus, valve acceleration becomes:

$$a = \omega^2 j_0 \quad (\text{Eq 98})$$

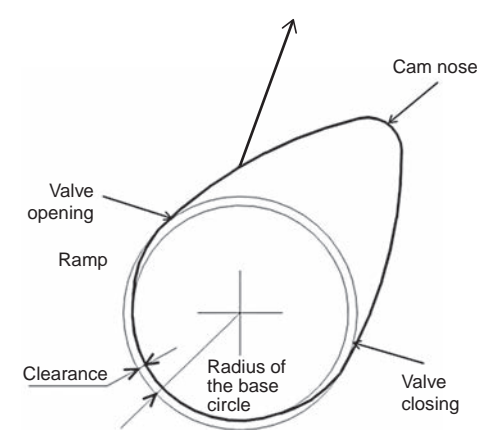
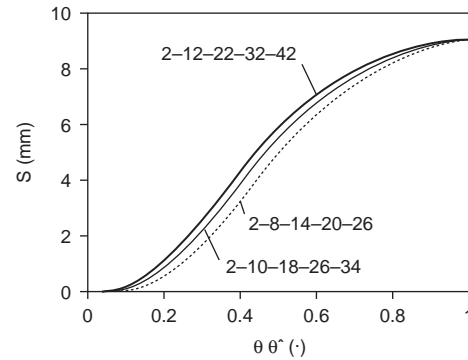


Fig. 18 Automotive polynomial cam

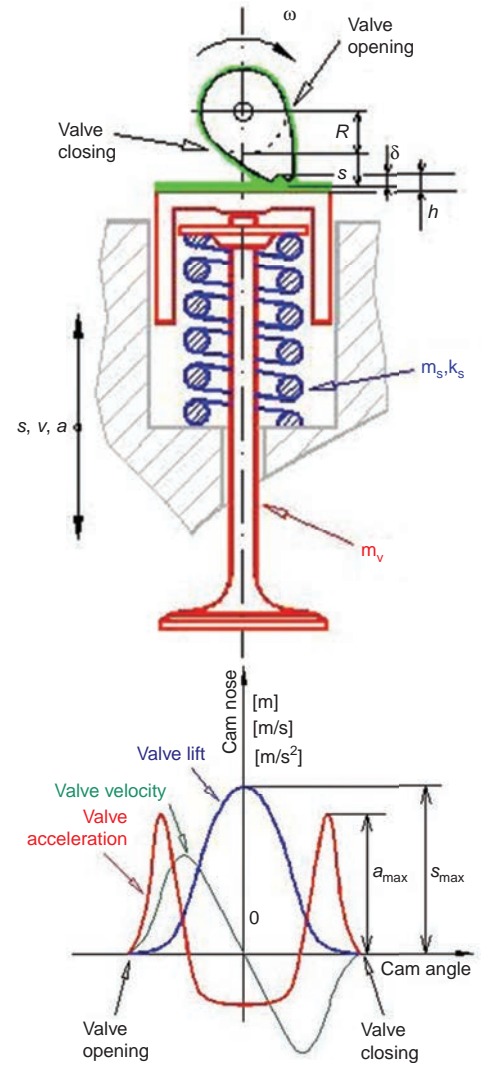


Fig. 19 Direct overhead cam-tappet valve train system

$$\begin{cases} C_2 = -s_{\max} \cdot m \cdot p \cdot q \cdot r / [(m-2)(p-2)(q-2)(r-2)] \\ C_m = -s_{\max} \cdot 2 \cdot p \cdot q \cdot r / [(2-m)(p-m)(q-m)(r-m)] \\ C_p = -s_{\max} \cdot 2 \cdot m \cdot q \cdot r / [(2-p)(m-p)(q-p)(r-p)] \\ C_q = -s_{\max} \cdot 2 \cdot m \cdot p \cdot r / [(2-q)(m-q)(p-q)(r-q)] \\ C_r = -s_{\max} \cdot 2 \cdot m \cdot p \cdot q / [(2-r)(m-r)(p-r)(q-r)] \end{cases} \quad (\text{Eq 95})$$

The applied load is the result of two forces. One is the valve spring elastic force, F_e , which is the result of any applied preload, F_p , and the instantaneous spring force, F_s , caused by the valve lift, s . Thus:

$$F_e = F_p + F_s = k_s(s + s_p) \quad (\text{Eq 99})$$

where s_p is any precompression of the valve spring, often carried out to guard against any loss of contact that may occur with a surge in inertial force, F_i (Fig. 20). Therefore, there is ideally always a net contact force, W .

The inertial force is:

$$F_i = \left(m_v + \frac{1}{3}m_s\right)a \quad (\text{Eq 100})$$

where m_v is the mass of the valve, and m_s is the mass of the valve spring. The valve spring is an unequal-pitched nonlinear spring of non-negligible mass. It can be shown that a third of its mass contributes to the valve dynamics, as stated in Eq 100. In reality, its stiffness alters as a function of its compression. Therefore, the approach here, using a constant stiffness, is idealized. A better representation of valve spring stiffness is shown in Ref 60. The variations of the elastic force, F_e , and the inertial force, F_i , are shown in Fig. 20. The net result is the contact force, W :

$$W = F_e - F_i \quad (\text{Eq 101})$$

Clearly, if $F_i \geq F_e$, there would be no contact force. This leads to the valve spring surge effect, loss of contact, and subsequent cam-tappet impact. In some cases, the valve spring surge can lead to spring coil clash as well. To mitigate these damaging effects, a suitable preload is necessary, as shown in the figure. Also note that a is a function of engine speed, ω (Eq 98), and because the inertial force significantly alters post valve opening and prior to valve closure, the contact force changes accordingly.

With the contact force determined, the maximum Hertzian contact pressure can be obtained by using Table 1 for the case of elastic line contact footprint geometry in the case of the cam-flat tappet arrangement. For cam-roller configuration, the contact footprint area is elliptical when using the relations in Table 1. To find the lubricant film thickness, the speed of entraining motion of the lubricant into the contact should be determined first. The tappet spins as the cam rotates. This is to avoid a repetitive path of sliding contact, which tends to scuff the tappet surface. Therefore, the cam is designed to approach the tappet with a predefined eccentricity that encourages the tappet to spin, thus reducing the chance of scuffing (Ref 61). Ignoring the effect of tappet spin, the relative sliding speed of the cam surface with respect to an assumed stationary tappet becomes:

$$v_c = \omega r \quad (\text{Eq 102})$$

where v_c is the instantaneous cam surface velocity, and r is the effective contact radius of the cam at the point of contact:

$$r = R + s + j_0 \quad (\text{Eq 103})$$

where R is the base circle radius. Therefore, for an assumed stationary tappet, the speed of entraining motion of the lubricant into the contact becomes (average speed of the two contacting surfaces):

$$u = \frac{1}{2}\omega\left(R + s + \frac{a}{\omega^2}\right) \quad (\text{Eq 104})$$

Now, an extrapolated oil film thickness equation, such as those in the section "Predicting Lubricant Film Thickness" in this article, for the line contact geometry can be used to obtain the lubricant film thickness (Ref 8, 25). With a known film thickness, viscous and boundary friction contributions can be obtained through the procedures highlighted in the section "Friction and Power Loss."

Figure 21 shows that at either side of the cam nose, the speed of entraining motion reverses. This means that the inlet to the contact conjunction reverses in position. Therefore, there is a short period of time with zero entrainment velocity, and, as a result, the lubricant film diminishes. Any quite thin film of lubricant is sustained by pure squeeze-film motion or through entrapment in the roughness of the contiguous surfaces. Therefore, there is a chance of wear in these positions. Figure 21 shows the film thickness at high and low cam rotational speeds, with little change in the inlet reversal positions. The figure also shows the film thickness in a polar representation in a cam cycle.

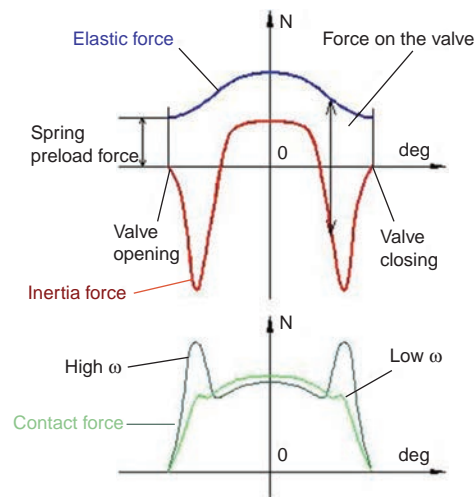


Fig. 20 Forces acting in a direct overhead valve train system

Transmission and Differential Gearing Systems

Aside from the engine conjunctions, the drive train system also contributes to frictional power loss. This occurs mainly in the gear pair contacts in the transmission and the differential unit. There are also losses due to bearing supports (usually ball, rolling-element, and taper roller bearings) of gear shafts, such as the transmission input and output shafts, the pinion shaft of the differential, and the output axle half-shafts of the differential ring gear. These systems are fairly efficient compared with journal bearings and piston-cylinder conjunctions, because they are generally subjected to the elasto-hydrodynamic regime of lubrication, with the least generated friction (Fig. 5). The transmission efficiency for helical gears of the usual automotive transmission systems and hypoid or bevel gears of the differential unit is usually in the range of 93 to 98% per meshing pair. This means that the frictional losses, being the main source of inefficiency, account for 2 to 7% of the input power. Nevertheless, any reduction of these would improve fuel efficiency and, by implication, the emissions from the power train system. With elasto-hydrodynamic conditions, the main areas of concern in gearing systems as well as their bearing supports are contact fatigue with excessive loads and poor noise and vibration performance with loss of contact and/or any preload. There are a large number of gear pair configurations. Here, the emphasis is on cylindrical gears—spur and helical—because these are most commonly

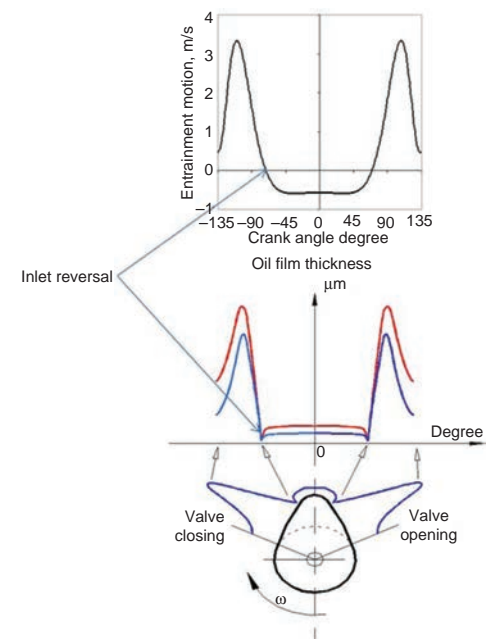


Fig. 21 Speed of lubricant entrainment and film thickness in a cam cycle

used in transmission systems. For spur gears, readers can also refer to Ref 62 and 63; for helical gears of transmission systems, analytical and numerical solutions are provided by De la Cruz et al. in Ref 64 and 65; and for hypoid gear pairs, analytical solutions are provided in Ref 66 and 67. More comprehensive numerical solutions are reported in Ref 68 and 69.

For all cases of gearing systems, in an analytical approach, lubricant film thickness can be obtained using the equations in the section "Predicting Lubricant Film Thickness" in this article. Note: for spur and helical gears, use the line contact equation; for spiral gear teeth, use any elliptical point contact equation; and for bevel and hypoid gears, use an elliptical point contact equation with angled flow lubricant entrainment, such as Eq 35 (Ref 26). For an analytical solution, viscous friction can be obtained using Eq 49, 50, and 52, because gears generally operate under non-Newtonian viscous shear of the lubricant. Boundary friction contribution can be obtained by using the procedure set out in the section "Boundary Friction" in this article, if the Stribeck oil film parameter, $\lambda < 3$ (in the section "Regimes of Lubrication"). To ascertain the potential for fatigue of contacting surfaces, use the approach used in the section "Contact Fatigue." A more comprehensive approach is highlighted by Paouris et al. (Ref 70) for the case of hypoid gears, which is applicable to all gear pairs.

Inlet Starvation

Most gear teeth contacts are usually partially starved. However, a sufficient inlet meniscus of lubricant exists to form a thin elastohydrodynamic film. As noted in the section "Piston-Cylinder Conjunctions" in this article, some of the inward flow into the contact is subject to swirl and counter (or reverse) flow (Fig. 22). Thus, only a portion of lubricant available at the entrance to the contact is actually entrained into it. This means that many contacts, including those of gear teeth, are subject to starvation with thinner films than that predicted, usually assuming a fully flooded or drowned inlet.

Based on numerical analyses, Hamrock and Dowson (Ref 71) provided a dimensionless parameter, m^* , which defines the inlet boundary at the onset of starvation as:

$$m^* = 1 + 3.06 \left[\left(\frac{R_{zx}}{b} \right)^2 \frac{h_0}{R_{zx}} \right]^{0.58} \quad (\text{Eq 105})$$

where R_{zx} is the equivalent radius in the principal plane of contact zx , with x being the direction of entraining motion. For a pair of gear teeth, R_{zx} is the instantaneous equivalent radius of the contacting teeth pair during meshing (each approximated by a cylinder of a given radius), and b is the half-width of the rectangular strip for a narrow band line contact or the semiminor half-width of an elliptical contact footprint (see the section "Contact

Configuration" in this article). The actual inlet parameter, m , is a function of the surface speed of the contacting surfaces, $k = \frac{u_1}{u_2}$. Therefore, m is the inlet parameter for the condition of zero reverse flow, at an angle θ_i , to the centerline of the contact. This was obtained through potential flow analysis by Tipei (Ref 50). The inlet distance to the zero reverse boundary is:

$$m = \frac{x_i}{b} = \frac{R_{zx} \sin \theta_i}{b} \quad (\text{Eq 106})$$

where x_i is the inlet distance, and

$$\theta_i = \cos^{-1} \left(1 - \frac{h_{0,s}}{R_{zx}} (\cosh \theta_i - 1) \right) \quad (\text{Eq 107})$$

The inlet position angle, θ_i as a function of $\cosh \theta_i$, varies with the value of k . For pure rolling condition, $k = 1$, and the value of $\cosh \theta_i = 11.28$ (Ref 50). For other k values, refer to Ref 50. h_0 is the central film thickness, for example, using Eq 35. The relationship between m and m^* is given in Ref 71 as:

$$\left(\frac{h_{0,s}}{h_0} \right) = \left(\frac{m-1}{m^*-1} \right)^{0.29} \quad (\text{Eq 108})$$

This relationship can be used to obtain the lubricant film thickness with contact starvation, $h_{0,s}$, which is reduced from h_0 , thus affecting both viscous friction as well as friction contribution due to direct boundary interactions. The approach here is validated by a combined numerical and experimental study using optical interferometry as well as measurement of contact pressure distribution for the case of circular point contacts (Ref 72) and applied to hypoid gear pairs in Ref 73.

The variables θ_e and x_e mark the outlet position from the conjunction (Ref 50).

Cylindrical Spur Gears

To use the procedures highlighted previously, the teeth-pair contact load, speed of lubricant entraining motion, and the contact sliding velocity are required at any instant of time during the meshing cycle of a pair of teeth,

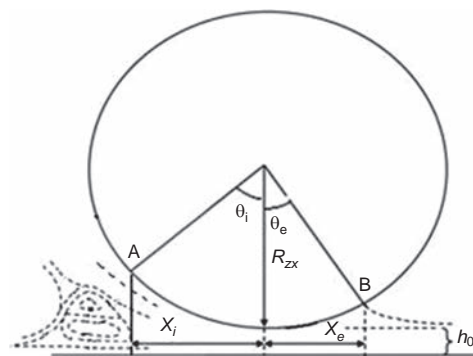


Fig. 22 Inlet reverse and swirl flows causing contact starvation

all of which are functions of system dynamics as well as contact geometry (gear teeth geometry). During a meshing period, one, two, or sometimes three teeth pairs are in contact. Figure 23 shows the length of a meshing cycle, g_x , where X is any instant of contact from the start of a meshing cycle. Two teeth pairs (a leading and an intermediate pair) are in simultaneous contact. As the leading pair begins to depart, a trailing pair enters into contact. Therefore, in the demonstrated case here, there are instances of two simultaneous teeth pairs in contact, and for a proportion of a meshing cycle, represented in the figure by the base pitch, p_e , a single teeth pair carries all the load.

The instantaneous equivalent radius of curvature of contact for a pair of cylindrical spur gears at any instant during the meshing cycle becomes:

$$R_{zx}(X) = \frac{(r_{b1} \tan \varphi + X)(r_{b2} \tan \varphi - X)}{(r_{b1} + r_{b2}) \tan \varphi} \quad (\text{Eq 109})$$

where r_{b1} and r_{b2} are the base circle radii of the gear pairs, and φ is the pressure angle. For gears other than spur gears, an analytical expression such as that in Eq 109 is not easily obtained. In these cases, tooth contact analysis is usually carried out, providing both the geometrical information, such as the principal radii of contact, as well as contact kinematics and deformation information (Ref 74).

The transmitted torque determines the teeth pair contact load, $W \approx \frac{T}{p}$, where the torque, T , is the input torque, resident on the transmission input shaft. When more than one pair of teeth is in contact, one can assume that this torque is equally shared between them.

The speed of entraining motion of the lubricant into the contact is obtained as:

$$U(X) = \omega r_p \left\{ X \cos(\varphi) \left(\frac{1}{r_{b1}} - \frac{1}{r_{b2}} \right) + 2 \sin(\varphi) \right\} \quad (\text{Eq 110})$$

The sliding velocity is obtained as:

$$\Delta U(X) = \omega r_p X \cos(\varphi) \left(\frac{1}{r_{b1}} - \frac{1}{r_{b2}} \right) \quad (\text{Eq 111})$$

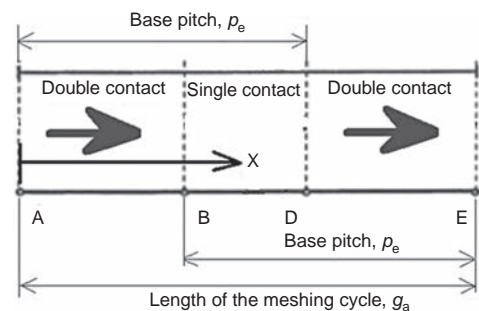


Fig. 23 Meshing cycle with simultaneous teeth-pair contacts

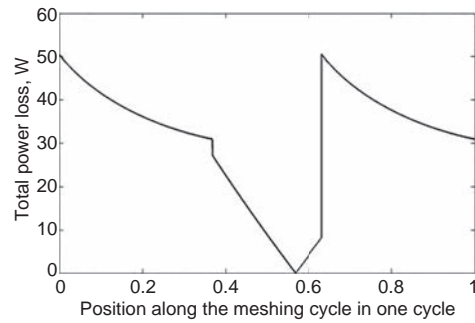


Fig. 24 Power loss for one meshing cycle

Now the relationships for prediction of film thickness, friction, and power loss can be used for all the teeth pairs in simultaneous mesh during a meshing cycle, $0 \leq X \leq g_z$, and summed up for each value of X . Figure 24 shows the total power loss for a meshing cycle. The abscissa value of zero indicates the beginning of a meshing cycle, and the value of unity indicates the end of the meshing cycle. For a spur gear pair, the sliding velocity diminishes at the pitch point contact, as shown in the figure.

REFERENCES

- H. Hertz, On the Contact of Elastic Solids, *J. Reine Angew. Math.*, Vol 92 (No. 110), 1881, p 156–171
- P.M. Johns and R. Gohar, Roller Bearings under Radial and Eccentric Loads, *Tribol. Int.*, Vol 14 (No. 3), 1981, p 131–136
- P.M. Johns-Rahnejat and R. Gohar, "Point Contact Elastohydrodynamic Pressure Distribution and Sub-Surface Stress Field," Tri-Annual Conference on Multi-Body Dynamics: Monitoring and Simulation Techniques (Bradford, U.K.), 1997
- M. Teodorescu, H. Rahnejat, R. Gohar, and D. Dowson, Harmonic Decomposition Analysis of Contact Mechanics of Bonded Layered Elastic Solids, *Appl. Math. Model.*, Vol 33 (No. 1), 2009, p 467–485
- M.T. Huber, Zur Theorie der Berührung Fester Elastischer Körper, *Ann. Phys.*, Vol 14, 1904
- E. Broszeit and O. Zwirlein, Internal Stresses and Their Influence on Material Stresses in Hertzian Contacts—Calculations with Different Stress Hypotheses, *Trans. ASME, J. Tribol.*, Vol 108, 1986, p 387
- E. Ioannides and T.A. Harris, A New Fatigue Life Model for Rolling Bearings, *Trans. ASME, J. Tribol.*, Vol 107 (No. 3), 1985, p 367–378
- R. Gohar and H. Rahnejat, *Fundamentals of Tribology*, Imperial College Press, London, 2008
- K.L. Johnson, *Contact Mechanics*, Cambridge University Press, Cambridge, 1985
- R. Stribeck, Die Wesentliechen ichen Eigenschaften Gleit und Rollen Lager (Ball Bearings for Various Loads), *Trans. ASME*, Vol 29, 1907, p 420–463
- S. Balakrishnan, C. McMin, C.E. Baker, and H. Rahnejat, *Fundamentals of Crank and Camshaft Support Journal Bearings, Tribology and Dynamics of Engine and Powertrain*, Woodhead Publishing, Cambridge, U.K., 2010
- S. Theodossiades, O. Tangasawi, and H. Rahnejat, Gear Teeth Impacts in Hydrodynamic Conjunctions Promoting Idle Gear Rattle, *J. Sound Vib.*, Vol 303 (No. 3), 2007, p 632–658
- A.M. Sharaf, H. Rahnejat, and P.D. King, Analysis of Handling Characteristics of All-Wheel-Drive Off-Road Vehicles, *Int. J. Heavy Vehicle Syst.*, Vol 15 (No. 1), 2008, p 89–106
- A.M. Sharaf, G. Mavros, H. Rahnejat, P.D. King, and S.K. Mohan, Optimisation of AWD Off-Road Vehicle Performance Using Visco-Lock Devices, *Int. J. Heavy Vehicle Syst.*, Vol 15 (No. 2–4), 2008, p 188–207
- H. Eyring, Viscosity, Plasticity, and Diffusion as Examples of Absolute Reaction Rates, *J. Chem. Phys.*, Vol 4 (No. 4), 1936, p 283–291
- M.M. Cross, Rheology of Non-Newtonian Fluids: A New Flow Equation for Pseudoplastic Systems, *J. Colloid Sci.*, Vol 20 (No. 5), 1965, p 417–437
- C. Barus, Isothermal Isopietics and Isometrics in Relation to Viscosity, *Am. J. Sci.*, 3rd Series, Vol 45, 1893, p 87–96
- C.J.A. Roelands, "Correlation Aspects of the Viscosity-Temperature-Pressure Relationships of Lubricating Oils," V.R.B. Druck, Kleine der A3-4, Groningen, 1966
- L. Houpert, New Results of Traction Force Calculations in Elastohydrodynamic Contacts, *Trans. ASME*, Vol 107, 1985, p 241–248
- H. Vogel, The Law of the Relation between the Viscosity of Liquids and the Temperature, *Phys. Z.*, Vol 22, 1921, p 645–646
- D. Dowson and G.R. Higginson, *Elastohydrodynamic Lubrication: The Fundamentals of Roller and Gear Lubrication*, Pergamon Press, Oxford, 1966
- P. Yang, J. Cui, Z.M. Jin, and D. Dowson, Transient Elastohydrodynamic Analysis of Elliptical Contacts, Part 2: Thermal and Newtonian Lubricant Solution, *Proc. Inst. Mech. Eng. J, J. Eng. Tribol.*, Vol 219, 2005, p 187–200
- R. Larsson and O. Andersson, Lubricant Thermal Conductivity and Heat Capacity under High Pressure, *Proc. Inst. Mech. Eng. J, J. Eng. Tribol.*, Vol 214 (No. 4), 2000, p 337–342
- D.G. Wymer and A. Cameron, Elastohydrodynamic Lubrication of a Line Contact, *Proc. Inst. Mech. Eng.*, Vol 188 (No. 1), 1974, p 221–238
- H. Rahnejat, Computational Modelling of Problems in Contact Dynamics, *Eng. Analy.*, Vol 2 (No. 4), 1985, p 192–197
- R.J. Chittenden, D. Dowson, J.F. Dunn, and C.M. Taylor, A Theoretical Analysis of the Isothermal Elastohydrodynamic Lubrication of Concentrated Contacts, Part I: Direction of Lubricant Entrainment Coincident with the Major Axis of the Hertzian Contact Ellipse, *Proc. R. Soc. A: Math., Phys., Eng. Sci.*, Vol 397 (No. 1813), 1985, p 245–269
- N. Morris, R. Rahmani, H. Rahnejat, P.D. King, and B. Fitzsimons, The Influence of Piston Ring Geometry and Topography on Friction, *Proc. Inst. Mech. Eng. J, J. Eng. Tribol.*, Vol 227 (No. 2), 2013, p 141–153
- M. Gore, M. Perera, G. Styles, P.D. King, and H. Rahnejat, Wear Characteristics of Advanced Honed and Cross-Hatched Coated Cylinder Liners, *Proc. of the 66th Annual Meeting and Exhibition of the STLE*, 2011, p 73
- J.A. Greenwood and J.H. Tripp, The Contact of Two Nominally Flat Rough Surfaces, *Proc. Inst. Mech. Eng.*, Vol 185 (No. 1), 1970, p 625–633
- J.I. McCool, Relating Profile Instrument Measurements to the Functional Performance of Rough Surfaces, *Trans. ASME, J. Tribol.*, Vol 109, 1987, p 264–270
- R.S. Sayles and T.R. Thomas, Thermal Conductance of Rough Elastic Contact, *Appl. Energy*, Vol 2, 1976, p 249–267
- M.S. Louquet-Higgings, The Statistical Analysis of a Random Moving Surface, *Philos. Trans. R. Soc. (London) A*, Vol 249, 1957, p 321–387
- A.W. Bush, R.D. Gibson, and G.P. Keogh, The Limit of Elastic Deformation in the Contact of Rough Surfaces, *Mech. Res. Commun.*, Vol 3, 1976, p 169–174
- C.R. Evans and K.L. Johnson, The Rheological Properties of Elastohydrodynamic Lubricants, *Proc. Inst. Mech. Eng. C, J. Mech. Eng. Sci.*, Vol 200 (No. 5), 1986, p 303–312
- M. Reiner, The Deborah Number, *Phys. Today*, Vol 17 (No. 1), 1964, p 62
- M. Leighton, N. Morris, M. Gore, R. Rahmani, H. Rahnejat, and P.D. King, Boundary Interactions of Rough Non-Gaussian Surfaces, *Proc. Inst. Mech. Eng. J, J. Eng. Tribol.*, June 30, 2016
- K.N. Fuller and D.F. Tabor, The Effect of Surface Roughness on the Adhesion of Elastic Solids, *Proc. R. Soc. (London) A, Math., Phys., Eng. Sci.*, Vol 345 (No. 1642), 1975, p 327–342
- M.-T. Ma, I. Sherrington, E.H. Smith, and N. Grice, Development of a Detailed Model for Piston-Ring Lubrication in IC Engines with Circular and Non-Circular Cylinder Bores, *Tribol. Int.*, Vol 30 (No. 11), 1997, p 779–788
- R. Rahmani, S. Theodossiades, H. Rahnejat, and B. Fitzsimons, Transient Elastohydrodynamic Lubrication of Rough New or Worn Piston Compression Ring

- Conjunction with an Out-of-Round Cylinder Bore, *Proc. Inst. Mech. Eng. J, J. Eng. Tribol.*, Vol 226 (No. 4), 2012, p 284–305
40. N. Morris, R. Rahmani, H. Rahnejat, P.D. King, and B. Fitzsimons, Tribology of Piston Compression Ring Conjunction under Transient Thermal Mixed Regime of Lubrication, *Tribol. Int.*, Vol 59, 2013, p 248–258
 41. N.W. Bolander, B.D. Steenwyk, F. Sadeghi, and G.R. Gerber, Lubrication Regime Transitions at the Piston Ring-Cylinder Liner Interface, *Proc. Inst. Mech. Eng. J, J. Eng. Tribol.*, Vol 219 (No. 1), 2005, p 19–31
 42. H. Shahmohamadi, R. Rahmani, H. Rahnejat, C.P. Garner, and P.D. King, Thermo-Mixed Hydrodynamics of Piston Compression Ring Conjunction, *Tribol. Lett.*, Vol 51 (No. 3), 2013, p 323–340
 43. B. Littlefair, M. De la Cruz, S. Theodossiades, R. Mills, S. Howell-Smith, H. Rahnejat, and R.S. Dwyer-Joyce, Transient Tribo-Dynamics of Thermo-Elastic Compliant High-Performance Piston Skirts, *Tribol. Lett.*, Vol 53 (No. 1), 2014, p 51–70
 44. F. McClure, “Numerical Modelling of Piston Secondary Motion and Skirt Lubrication in Internal Combustion Engines,” Ph.D. thesis, Massachusetts Institute of Technology, 2007
 45. S. Howell-Smith, H. Rahnejat, P.D. King, and D. Dowson, Reducing In-Cylinder Parasitic Losses through Surface Modification and Coating, *Proc. Inst. Mech. Eng. D, J. Automob. Eng.*, Vol 228 (No. 4), 2014, p 391–402
 46. T. Tian, Dynamic Behaviours of Piston Rings and Their Practical Impact, Part 1: Ring Flutter and Ring Collapse and Their Effects on Gas Flow and Oil Transport, *Proc. Inst. Mech. Eng. J, J. Eng. Tribol.*, Vol 216, 2002, p 209–227
 47. C.E. Baker, S. Theodossiades, H. Rahnejat, and B. Fitzsimons, Influence of In-Plane Dynamics of Thin Compression Rings on Friction in Internal Combustion Engines, *Trans. ASME, J. Eng. Gas Turb. Power*, Vol 134 (No. 9), 2012, p 092801
 48. C. Baker, R. Rahmani, S. Theodossiades, H. Rahnejat, and B. Fitzsimons, On the Effect of Transient In-Plane Dynamics of the Compression Ring upon Its Tribological Performance, *Trans. ASME, J. Eng. Gas Turb. Power*, Vol 137 (No. 3), 2015, p 032512
 49. G. Birkhoff and D.F. Hays, Free Boundaries in Partial Lubrication, *J. Math. Phys.*, Vol 42 (No. 1), 1963, p 126–138
 50. N. Tipei, Boundary Conditions of a Viscous Flow between Surfaces with Rolling and Sliding Motion, *Trans. ASME, J. Lubr. Technol.*, Vol 90 (No. 1), 1968, p 254–261
 51. O. Reynolds, On the Theory of Lubrication and Its Application to Mr. Beauchamp Tower’s Experiments, Including an Experimental Determination of the Viscosity of Olive Oil, *Proc. R. Soc. (London)*, Vol 40 (No. 242–245), 1886, p 191–203
 52. H.W. Swift, The Stability of Lubricating Films in Journal Bearings, *Proc. Inst. Civ. Eng.*, Vol 233, 1932, p 267–288
 53. W. Stieber, *Des Schwimmlager: Hydrodynamische Theorie des Gleitlagers*, VDI-Verlag, 1933
 54. L. Prandtl, On Boundary Layers in Three-Dimensional Flow, *Rep. Trans.*, Vol 64, 1946
 55. R. Gohar, *Elastohydrodynamics*, 2nd ed., Imperial College Press, London, 2001
 56. H. Rahnejat, *Multi-Body Dynamics: Vehicles, Machines and Mechanisms*, Professional Engineering Publishing, Bury St. Edmunds, U.K., 1998
 57. M. Gore, R. Rahmani, H. Rahnejat, and P.D. King, Assessment of Friction from Compression Ring Conjunction of a High-Performance Internal Combustion Engine: A Combined Numerical and Experimental Study, *Proc. Inst. Mech. Eng. C, J. Mech. Eng. Sci.*, Vol 230 (No. 12), 2016, p 2073–2085
 58. P.C. Mishra and H. Rahnejat, Tribology of Big-End-Bearings, *Tribology and Dynamics of Engine and Powertrain*, Woodhead Publishing, Cambridge, U.K., 2010
 59. M. Mohammadpour, R. Rahmani, and H. Rahnejat, Effect of Cylinder Deactivation on the Tribo-Dynamics and Acoustic Emission of Overlay Big End Bearings, *Proc. Inst. Mech. Eng. K, J. Multi-Body Dynam.*, Vol 228 (No. 2), 2014, p 138–151
 60. M. Kushwaha, H. Rahnejat, and Z.M. Jin, Valve-Train Dynamics: A Simplified Tribo-Elasto-Multi-Body Analysis, *Proc. Inst. Mech. Eng. K, J. Multi-Body Dynam.*, Vol 214 (No. 2), 2000, p 95–110
 61. M. Teodorescu, M. Kushwaha, H. Rahnejat, and S.J. Rothberg, Multi-Physics Analysis of Valve Train Systems: From System Level to Microscale Interactions, *Proc. Inst. Mech. Eng. K, J. Multi-Body Dynam.*, Vol 221 (No. 3), 2007, p 349–361
 62. T.T. Petry-Johnson, A. Kahraman, N.E. Anderson, and D.R. Chase, An Experimental Investigation of Spur Gear Efficiency, *Trans. ASME, J. Mech. Des.*, Vol 130 (No. 6), 2008, p 062601
 63. S. Li and A. Kahraman, A Transient Mixed Elastohydrodynamic Lubrication Model for Spur Gear Pairs, *Trans. ASME, J. Tribol.*, Vol 132 (No. 1), 2010, p 011501
 64. M. De la Cruz, S. Theodossiades, and H. Rahnejat, An Investigation of Manual Transmission Drive Rattle, *Proc. Inst. Mech. Eng. K, J. Multi-Body Dynam.*, Vol 224 (No. 2), 2010, p 167–181
 65. M. De la Cruz, W.W.F. Chong, M. Teodorescu, S. Theodossiades, and H. Rahnejat, Transient Mixed Thermo-Elastohydrodynamic Lubrication in Multi-Speed Transmissions, *Tribol. Int.*, Vol 49, 2012, p 17–29
 66. I. Karagiannis, S. Theodossiades, and H. Rahnejat, On the Dynamics of Lubricated Hypoid Gears, *Mech. Mach. Theory*, Vol 48, 2012, p 94–120
 67. M. Mohammadpour, S. Theodossiades, H. Rahnejat, and P. Kelly, Transmission Efficiency and Noise, Vibration and Harshness Refinement of Differential Hypoid Gear Pairs, *Proc. Inst. Mech. Eng. K, J. Multi-Body Dynam.*, Vol 228 (No. 1), 2014, p 19–33
 68. H. Xu and A. Kahraman, Prediction of Friction-Related Power Losses of Hypoid Gear Pairs, *Proc. Inst. Mech. Eng. K, J. Multi-Body Dynam.*, Vol 221 (No. 3), 2007, p 387–400
 69. M. Mohammadpour, S. Theodossiades, H. Rahnejat, and T. Saunders, Non-Newtonian Mixed Elastohydrodynamics of Differential Hypoid Gears at High Loads, *Meccanica*, Vol 49 (No. 5), 2014, p 1115–1138
 70. L. Paouris, S. Theodossiades, M. De la Cruz, H. Rahnejat, A. Kidson, G. Hunt, and W. Barton, Lubrication Analysis and Sub-Surface Stress Field of an Automotive Differential Hypoid Gear Pair under Dynamic Loading, *Proc. Inst. Mech. Eng. C, J. Mech. Eng. Sci.*, Vol 230 (No. 7–8), 2016, p 1183–1197
 71. B.J. Hamrock and D. Dowson, Isothermal Elastohydrodynamic Lubrication of Point Contacts, Part IV: Starvation Results, *Trans. ASME, J. Lubr. Technol.*, Vol 99 (No. 1), 1977, p 15–23
 72. M. Mohammadpour, P.M. Johns-Rahnejat, H. Rahnejat, and R. Gohar, Boundary Conditions for Elastohydrodynamics of Circular Point Contacts, *Tribol. Lett.*, Vol 53 (No. 1), 2014, p 107–118
 73. M. Mohammadpour, S. Theodossiades, and H. Rahnejat, Transient Mixed Non-Newtonian Thermo-Elastohydrodynamics of Vehicle Differential Hypoid Gears with Starved Partial Counter-Flow Inlet Boundary, *Proc. Inst. Mech. Eng. J, J. Eng. Tribol.*, Vol 228 (No. 10), 2014, p 1159–1173
 74. F.L. Litvin and A. Fuentes, *Gear Geometry and Applied Theory*, 2nd ed., Cambridge University Press, New York, 2004