Wind turbine main-bearing lubrication - Part 1: An introductory review of elastohydrodynamic lubrication theory

Edward Hart[1], Elisha de Mello[2], and Rob Dwyer-Joyce[2]

[1]Wind Energy and Control Centre, Department of Electronic and Electrical Engineering, The University of Strathclyde, Glasgow, UK
[2]Leonardo Centre for Tribology, Department of Mechanical Engineering, The University of Sheffield, UK

Correspondence: Edward Hart (edward.hart@strath.ac.uk)

Abstract. This paper is the first in a two-part study on lubrication in wind turbine main-bearings. Elastohydrodynamic lubrication is a complex field, the formulas and results from which should not be applied blindly, but with proper awareness and consideration of their context, validity and limitations in any given case. The current paper, "Part 1", therefore presents an introductory review of elastohydrodynamic lubrication theory in order to provide this necessary background and context in an accessible form, promoting cross-disciplinary understanding. Fundamental concepts, derivations and formulas are presented, followed by the more advanced topics of: starvation, dynamic effects, surface roughness interactions and grease lubrication. "Part 2" applies the presented material in order to analyse wind turbine main-bearing lubrication in the context of available film thickness formulas and related results from lubrication theory. Aside from the main-bearing, the material presented here is also applicable to other lubricated non-conformal contacts in wind turbines, including pitch and yaw bearings and gear-teeth.

1 Introduction

Wind turbine main-bearings have come under increased research scrutiny of late, due to higher than expected failures rates and failure mechanisms which are yet to be fully understood (Hart et al., 2019, 2020; Hart, 2020; Guo et al., 2021). Integral to main-bearing function and performance is the fact that it is a rolling bearing, tasked with allowing low-friction, free rotation of the shaft while also supporting the turbine rotor. Lubrication of the main-bearing is therefore necessary to prevent rapid wear and damage propagation from taking place. As such, the lubricant and lubrication mechanisms acting within this component must be accounted for as part of any attempt to fully characterize and understand main-bearing internal operational conditions, failure mechanisms and reliability. This two-part study seeks to begin this process.

Lubrication, and elastohydrodynamic lubrication (EHL) in particular, is a complex, nuanced and rapidly evolving field. For example, while simplified film thickness formulas have been developed, they should not be applied without careful consideration of their validity and possible limitations in any given case. Furthermore, additional effects may be present where operational conditions vary rapidly or where grease lubrication is used. For the benefit of non-EHL-specialists, it is argued that there is a need for an introductory review of this field which presents the reader with a comprehensive overview of EHL theory, including: fundamental equations and the problem formulation, the approximations being applied, numerical solution methods, general characteristics of solutions, simplified film equations and their validity, and, additional effects from dynamic

25 conditions, starvation, surface roughness and grease behaviour. To the best of the authors' knowledge no single paper in the literature covers all of these listed topics, including discussions of their more subtle nuances, while remaining accessible to non-experts. This paper, "Part 1" of the main-bearing lubrication study, therefore presents such an introductory review for the elastohydrodynamic lubrication of non-conformal (roller bearings, gear-teeth *etc.*) contacts. While main-bearings form the focus of the overall study, the material presented in Part 1 applies equally to other lubricated non-conformal contacts in wind

30 turbines, including pitch and yaw bearings and gear-teeth. To aid the reader, a table of symbols is provided in Appendix A.

2 Surface separation and lubrication regimes

Fluid film lubrication exists when two machine surfaces are completely separated by a layer of lubricant. In such circumstances, forces are carried via pressures generated within the lubricant and frictional/wear conditions are greatly improved. The presence of an adequate lubricant film is therefore critical to the reliability and longevity of machine components. In the context of wind

35 energy, one normally encounters non-conformal contacts (roller bearings, gear-teeth *etc.*) operating in the elastohydrodynamic regime[1], in which significant elastic deformations of lubricated surfaces occur.

For completely smooth surfaces a lubricant film would always be present, even if vanishingly small. However, in reality material surfaces are not perfectly smooth and exhibit roughness, geometrical variations, on the order of 0.01-10 μm (Hamrock et al., 2004). It transpires that typical lubricating film thicknesses also sit somewhere within this range, meaning film thickness

40 needs to be considered relative to surface roughness in order to determine whether a separation of surfaces has been achieved. The appropriate relative quantity is the film parameter,

$$\Lambda = \frac{h_m}{\sigma}, \tag{1}$$

which relates the minimum film thickness, h_m, to the combined (root-mean-square) roughness of contacting surfaces, $\sigma = \sqrt{\sigma_{\text{I}}^2 + \sigma_{\text{II}}^2}$ (Hamrock et al., 2004). While delineations between lubrication regimes are difficult to make exactly, the following

45 rough estimates indicate film parameter values associated with each (Hamrock et al., 2004):

– Hydrodynamic lubrication, $5 < \Lambda < 100$

– Elastohydrodynamic lubrication, $3 < \Lambda < 10$

– Mixed lubrication, $1 < \Lambda < 5$

– Boundary lubrication, $\Lambda < 1$.

50 Hydrodynamic lubrication is generally associated with conformal surfaces (e.g. journal bearings) and negligible elastic deformations. Boundary lubrication occurs when surfaces are no longer separated by a lubricant film and there is significant surface-surface contact. Mixed lubrication represents an intermediate state in which some penetration of the lubricant film has

---

[1]Although a novel (conformal) journal bearing design in this space is being developed (Rolink et al., 2020, 2021).
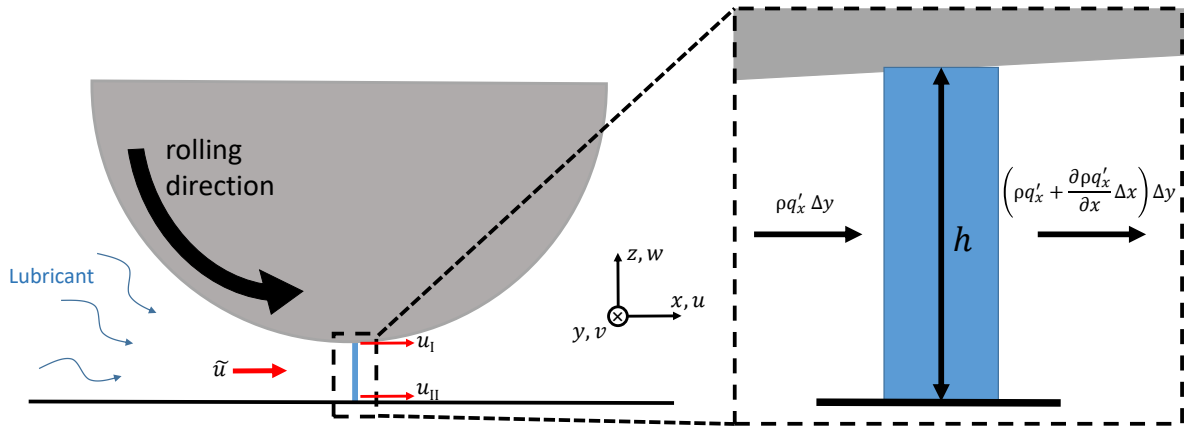
Figure 1. A control volume of fluid lying between moving bearing surfaces with surface tangential velocities u_{I} and u_{II}.

occurred, such that the load is shared between asperity contacts and fluid pressures. For non-conformal contacts fully elastohydrodynamic lubrication is aspired to, with mixed and boundary cases representing increasing levels of friction and a heightened
55   risk of wear related damage.

3   Reynolds equation and the elastohydrodynamic lubrication problem

Lubricated conjunctions can support applied loads as a result of pressure distributions generated through fluid film interactions. The differential equation governing these interactions is known as the Reynolds equation. While derivations and applications of this equation are commonplace, a proper discussion of nuances occurring in the current problem requires a more detailed
60   understanding of the equation's origin and underlying terms. As such, key elements from the derivation based on the laws of viscous flow and mass conservation will be presented, for more detailed considerations see Hamrock et al. (2004) and Dowson (1962). The full EHL problem is then defined. Note, a table of symbols is provided in Appendix A. The derivation of Reynolds equation begins by considering the rectangular control volume shown in Figure 1 of height h, width Δx and length (into the page) Δy. The mass of lubricant contained within this control volume at any point in time is $\rho h \Delta x \Delta y$, where ρ is the lubricant
65   density. The rate at which this mass changes over time is determined by the difference between mass flowing into and out of the control volume. From Figure 1, mass flow differences in x- and y-directions are given by,

$$-\frac{\partial \rho q_x'}{\partial x}\Delta x \Delta y \ \text{ and } \ -\frac{\partial \rho q_y'}{\partial y}\Delta y \Delta x, \tag{2}$$

where,

$$q_x' = \int_o^h u \, dz \ \text{ and } \ q_y' = \int_o^h v \, dz, \tag{3}$$

are volume flow rates per unit length/width. Conservation of mass requires that the rate at which mass is accumulated in the control volume equals the total difference between mass flowing into and out of the volume. Thus,

$$-\frac{\partial \rho q'_x}{\partial x}\Delta x \Delta y - \frac{\partial \rho q'_y}{\partial y}\Delta y \Delta x = \frac{\partial}{\partial t}\left(\rho h \Delta x \Delta y\right), \tag{4}$$

from which Δ terms cancel such that,

$$-\frac{\partial \rho q'_x}{\partial x} - \frac{\partial \rho q'_y}{\partial y} = \frac{\partial}{\partial t}\left(\rho h\right). \tag{5}$$

Applying zero-slip boundary conditions at lubricant-solid interfaces, and performing a number of integrations[2], it can be shown that volume flow rate expressions take the form (Hamrock et al., 2004),

$$q'_x = \underbrace{-\frac{h^3}{12\mu}\frac{\partial p}{\partial x}}_{\text{Poiseuille flow}} + \underbrace{\frac{u_\mathrm{I} + u_\mathrm{II}}{2}h}_{\text{Couette flow}} \tag{6}$$

$$q'_y = -\frac{h^3}{12\mu}\frac{\partial p}{\partial y} + \frac{v_\mathrm{I} + v_\mathrm{II}}{2}h. \tag{7}$$

μ is the lubricant dynamic viscosity and p is pressure. As indicated, these flow rates contain Poiseuille and Couette contributions. Poiseuille flow is that driven by pressure gradients in the fluid, whereas Couette flow is induced by surface velocities u_I and u_II (shown in Figure 1); more specifically by shear stresses resulting from a viscous fluid interacting with moving boundary surfaces.

Substituting Equations 6 and 7 into Equation 5 and defining the mean entrainment velocities,

$$\tilde{u} = \frac{u_\mathrm{I} + u_\mathrm{II}}{2}, \quad \tilde{v} = \frac{v_\mathrm{I} + v_\mathrm{II}}{2}, \tag{8}$$

the general Reynolds equation, which governs the pressure distribution in fluid film lubrication, is obtained;

$$\frac{\partial}{\partial x}\left(\frac{\rho h^3}{12\mu}\frac{\partial p}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\rho h^3}{12\mu}\frac{\partial p}{\partial y}\right) = \frac{\partial}{\partial x}\left(\rho h \tilde{u}\right) + \frac{\partial}{\partial y}\left(\rho h \tilde{v}\right) + \frac{\partial}{\partial t}\left(\rho h\right). \tag{9}$$

In a bearing context, the full EHL problem consists of finding a solution to the Reynolds equation which also satisfies the following conditions of *total film thickness* and *load balance*,

$$h = h_m + h_g + h_e, \tag{10}$$

$$\iint p(x,y,t)\,dxdy = w(t). \tag{11}$$

Equation 10 stipulates that, at each location, total film thickness (h) must equal the sum of surface separation components resulting from: minimum separation (h_m), un-deformed surface geometry (h_g) and local elastic deformations (h_e) caused by

---

[2]For example, this includes integrations: $\frac{\partial}{\partial z}\left(\mu \frac{\partial u}{\partial z}\right) \rightarrow \frac{\partial u}{\partial z} \rightarrow u \rightarrow \int u$

resulting pressures in the system. Equation 11 stipulates that the pressure distribution must balance the applied force, $w(t)$, at
each point in time. Evaluation of elastic deformations is discussed further in Section 3.1, below.

As presented, the Reynolds equation is valid for variable-viscosity compressible flows, with changes in viscosity and density
driven primarily by pressure (under isothermal conditions). Variations in lubricant properties are usually captured via empirical
equations in pressure; for example, the *Barus law* and *Roelands equation* (Hamrock et al., 2004),

$$\mu = \mu_0 e^{\alpha p}, \quad \mu = \mu_0 10^{-[1.2 + \log_{10}(\mu_0)]\left[1 - (1 + p/2000)^{Z_1}\right]}, \tag{12}$$

(respectively) are both commonly used under isothermal conditions to characterise changes in viscosity. μ_0 is the lubricant
dynamic viscosity at the inlet temperature and for (gauge-pressure) $p = 0$, α is the pressure-viscosity coefficient and Z_1 a
dimensionless pressure-viscosity index. Viscosity also varies strongly with temperature. In practise, viscosity information is
normally provided at two reference temperatures, with interpolation allowing an appropriate value for μ_0 (at the inlet tem-
perature) to be identified (ASTM, 2020). For non-isothermal computational EHL modelling, empirical equations along the
lines of Equation 12 that also include temperature have been developed (Hamrock et al., 2004). The Barus law, while easily
implemented and useful for gaining an intuitive understanding of pressure-viscosity effects, is known to provide a poor ap-
proximation of real lubricant viscosity variations with pressure. Similar equations are used to describe density variations with
pressure, for example,

$$\rho = \rho_0 \left(1 + \frac{0.6p}{1 + 1.7p}\right), \tag{13}$$

for some mineral oils, in which a roughly linear initial increase in density with pressure levels off to a maximum of $+35.3\%$ as
$p \to \infty$. In Equation 13, different coefficients may be used depending on the lubricant. More accurate representations have also
been developed which rely on greater numbers of coefficients. Empirical equations of the above types are developed within a
specific range of conditions, hence, the validity of applied empirical relationships should be considered when looking to solve
any given EHL problem.

3.1 Appromixations in EHL modelling

The outlined derivation and EHL problem definition together provide an accessible justification of Equations 9-11, however,
it should be noted that approximations are present for which proper consideration requires a first-principles derivation, and
associated discussions, starting from the Navier-Stokes and continuity equations. Similarly, elastic deformations of bearing
surfaces are usually resolved by approximating each body as an elastic half-space. As with any model, it is important to
understand the approximations being made and the conditions under which these approximations are valid. Therefore, these
aspects of numerical EHL models will be briefly outlined.

First, the fluid context is considered. EHL of machine components generally results in film thickness values which are
small compared to the length of the conjunction through which the lubricant is flowing (see Section 6, below). Denoting
typical surface separation in the conjunction by h_0 and typical conjunction length (over which changes in separation occur)
by L, in the limit $h_0/L \to 0$ the inertial terms in the Navier-Stokes momentum equations disappear and one also obtains

$\partial p/\partial z = 0$, *i.e.* pressure becomes constant across the film. Volume flow-rate expressions (Equations 6 and 7) are then obtained via integration of the simplified, and now quasi-steady, momentum equations. Reynolds equation (Equation 9) follows by applying an integrated form of the continuity equation, ensuring mass-flow conservation. The approximation, $h_o/L \approx 0$, is known as the "lubrication approximation" and is valid in cases where $h_o/L \ll 1$. See Panton (2013) and Hamrock et al. (2004) for more details. Note, the derivation outlined in Section 3 implicitly uses this same approximation.

With respect to conjunction shape and deformation, undeformed surface geometries are generally approximated as being parabolic, resulting in,

$$h_g = \frac{x^2}{2R_x} + \frac{y^2}{2R_y}, \tag{14}$$

for R_x and R_y the reduced radii of curvature in x and y directions (see Appendix A). Local deformations are most commonly evaluated by treating each body as an elastic half-space, to which is applied the same pressure distribution as exists in the lubricated conjunction. This simplification allows deformations to be evaluated using relatively simple integral formulas (Johnson, 1987; Evans and Hughes, 2000). The magnitude of deflections resulting from an applied distribution of pressure is governed by the contacting materials' elastic moduli (E) and Poisson ratios (ν), which may all be combined into a single reduced modulus of elasticity, E' (see Appendix A). The half-space approximation is valid only if: surface geometries close to the contact region roughly approximate a plane surface, strains within the contact region are small enough to be evaluated using linear elasticity theory, and, stress fields resulting from pressures in the conjunction are not strongly influenced by body boundaries. These requirements are satisfied if the significant dimensions of the contact region are small with respect to the dimensions of the contacting bodies and the relative radii of curvature of the surfaces (Johnson, 1987). The same conditions on dimensions and curvature also ensure validity of the parabolic geometries approximation. Hertzian contact theory, which concerns the (dry) contact of non-conforming elastic solids, is closely related and also of relevance to lubrication problems (e.g. see Section 4). Hertzian theory assumes contacting surfaces (no longer separated by a lubricant layer) are frictionless, that the contact patch is elliptical, and applies the same approximations as outlined above for the evaluation of geometries and deflections (Johnson, 1987). While the EHL problem requires numerical modelling to solve, Hertzian analysis of dry contact yields elegant analytical formulae describing load-deflection relationships.

4 Line and point contacts

The current section deals with instances of dry contact. When an elastic solid is acted on by a load, deformation will occur. Cases of two contacting and loaded solids result in the formation of a contact patch, with the geometry of contacting solids determining the shape of the contact patch. Components which initially contact along a line (e.g. a cylindrical roller contacting a raceway), referred to as *line contacts*, lead to a rectangular contact patch with a semi-cylindrical surface normal-stress distribution which remains identical along its length[3] (Harris and Kotzalas, 2006b). For an applied load per unit length of

---

[3]Ignoring end-effects and roller crowning.

roller, w_l, this takes the form,

$$s_l(x) = \frac{2w_l}{\pi b} \left(1 - \left(\frac{x}{b} \right)^2 \right)^{1/2} . \tag{15}$$

Line contacts can be considered as the limiting case of a long (elliptical) point contact, see below, or the problem can be reduced to that of line-loading on a two-dimensional elastic half-space (Johnson, 1987). Components which initially contact

160 at a single point (e.g. a ball or spherical roller bearing contacting a raceway), referred to as *point contacts*, lead to an elliptical contact patch and semi-ellipsoidal surface normal-stress distribution,

$$s_p(x,y) = \frac{3w}{2\pi ab} \left(1 - \left(\frac{x}{b} \right)^2 - \left(\frac{y}{a} \right)^2 \right)^{1/2} , \tag{16}$$

for w the applied load and with maximum normal-stress located at the contact centre (Harris and Kotzalas, 2006b). With respect to pressures acting within a contact patch note that, under Hertzian contact, the pressure distribution being applied

165 across each surface (by the other) must equal the surface normal-stress distribution which results (Equations 15 and 16). Therefore, Equations 15 and 16 also describe the pressure distributions acting within the contacts, commonly referred to as the "Hertzian pressure distributions".

Contact patch geometry is captured by the *ellipticity* parameter,

$$k = \frac{a}{b}, \tag{17}$$

170 for a and b elliptical semi-major and semi-minor axes respectively[4]. For a given Hertzian elliptical contact, as load is applied, the axes expand proportionately to each other. Hence, k remains constant and is a function of undeformed surface geometries only. The value of k in a particular contact case is determined by an implicit equation, involving elliptic integrals, which requires iterative solving. Approximate formulas have been developed to allow for fast evaluation of the ellipticity parameter and associated elliptic integrals (Brewe and Hamrock, 1977; Antoine et al., 2006). The EHL problem outlined above applies

175 equally to both contact types, with the line contact case often simplified to a 2D axis-symmetric problem in which side-leakage (y, v and $\partial/\partial y$ terms) is neglected and applied load at time t becomes the applied load per unit length of roller, $w_l(t)$.

Given simplifications associated with line contact EHL, it has proved useful in some analyses to consider the concept of an *equivalent line contact* representation of a point contact. Taking a point contact with applied load w and patch dimensions a and b, a 2D line contact representation is sought which shares its rolling direction patch width (b), geometry (R_x) and

180 centreline ($y=0$) stress distribution under dry contact. These conditions can be shown to hold for the distributed load $w_l = 3w/4a$ applied to a line contact whose x-direction geometry matches that of the point contact, but, which has the adjusted reduced modulus $\tilde{E}' = \left(1 + \frac{R_x}{R_y} \right) \mathcal{E}^{-1} E'$. $\mathcal{E}$ is an elliptic integral of the second kind whose value depends on k (*e.g.* see Hart (2020)). Full details of this equivalent line contact formulation are provided in Appendix B. Other approaches to these types of equivalence have also been taken in the literature.

---

[4]Note, no single convention holds for the allocation of axis labels (x/y) and contact patch dimensions (a/b). In the current work, x is taken to be the direction of rolling and a the semi-major axis of the contact patch, with line contacts treated as long elliptical contacts in this context. For the types of rollers/contacts seen in wind turbine main-bearings (*i.e.* where the semi-major axis of contact lies transverse to the direction of rolling) this allocation results in the normal-stress distributions shown in Equations 15 and 16.

5 Full EHL solutions

The first complete solution to an EHL problem was presented by Dowson and Higginson in 1959 (Dowson and Higginson, 1959) for four line contact cases. All calculations were carried out by hand, using mathematical tables and mechanical calculators (Hooke, 2009). Much important work followed this initial breakthrough, but it wasn't until almost two decades later that computing power became sufficient to allow EHL solutions in the point contact case to be obtained (Lubrecht et al., 2009). Since then, a plethora of significant advances have followed regarding numerical solvers for EHL problems, including: development of advanced multilevel (equivalently multigrid) solvers (Venner, 1994; Venner and Lubrecht, 2000a, b); full coupling of elastic and hydrodynamic equations - the *differential deflection method* (Evans and Hughes, 2000; Hughes et al., 2000; Holmes et al., 2003) - which enhances algorithmic efficiency and stability; and computational fluid dynamics implementations. These listed solvers apply the elastic half-space approximation for evaluation of deflections. A *full system approach* has also been developed which incorporates full-body elasticity using finite element methods (Habchi et al., 2008; Lugt and Morales-Espejel, 2011; Habchi, 2018). Implementation of EHL in multibody dynamics software modelling has also been considered (Dlugoš and Novotnỳ, 2021). For an overview of recent developments in numerical EHL modelling see (Meng et al., 2020). Accurate EHL solutions are now generated routinely and fairly easily even for complex cases such as those involving dynamically varying loads and speed, moving surface roughness or mixed lubrication conditions. However, it should be noted that this only holds where solver code and relevant expertise are available since setting up such solvers is highly non-trivial.

Simplified formulations of the EHL problem have also been developed which allow for analytical and semi-analytical solutions to be obtained (Morales-Espejel and Wemekamp, 2008; Greenwood, 1972). Such formulations, while approximate, are highly efficient and provide important insights into EHL behaviour and conditions, even proving useful when implementing full numerical solvers (e.g. they can help with identification of appropriate mesh dimensions, as well as supporting interpretation and sense-checking of results).

6 General characteristics of EHL contacts

Figure 2 details characteristic features which tend to be present in EHL contacts[5], as has been confirmed extensively using numerical modelling and experimental investigations (Hamrock et al., 2004; Spikes, 2006; Albahrani et al., 2016; Foord et al., 1969; Wymer and Cameron, 1974). Significant elastic deformation can be see to have taken place, resulting in a near parallel channel throughout most of the contact conjunction. Pressure at the inlet can be seen to rise rapidly to meet the Hertzian (dry-contact) pressure curve (see Section 4), which is then tracked through much of the conjunction. Extremely high pressures develop within the contact gap, resulting in dramatic increases in lubricant viscosity and so dominance of the shear (Couette) driven terms of Equations 6 and 7. Prior to the outlet, a constriction occurs in the oil film, immediately after a sudden spike in pressure. These features are coupled, with the pressure spike driven by the abrupt reduction in film height. The constriction itself is a consequence of mass-flow continuity as follows: ignoring flow in the y-direction (side-leakage) and focusing on the

---

[5]More specifically, Figure 2 shows a 'slice' through the contact in the direction of rolling, x.
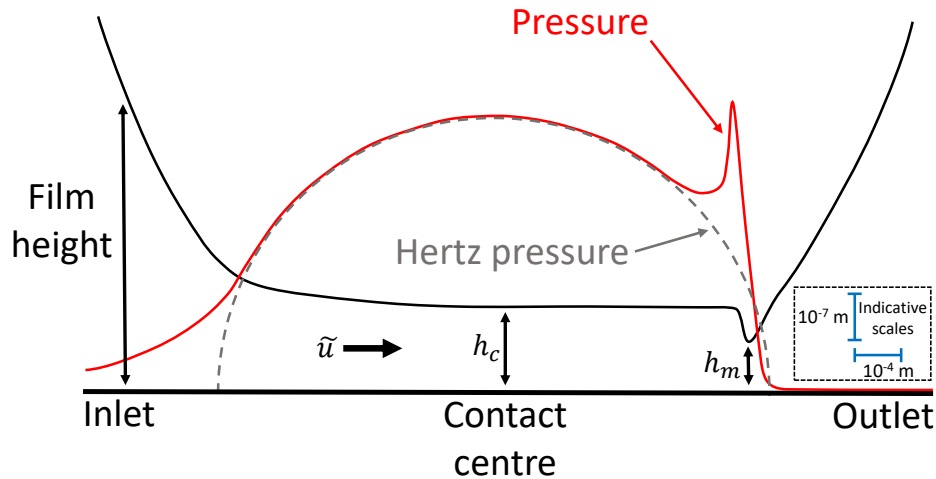
Figure 2. General characteristics of EHL contacts, including indicative orders of magnitude for vertical and horizontal scales. Also shown are the central and minimum film thickness values, h_c and h_m respectively. Note, the depicted features would be expected to occur within a narrow central section of the conjunction shown in Figure 1.

non-transient case, $\partial (\rho h) / \partial t = 0$, mass-flow continuity (Equation 5) requires,

$$\rho q'_x = \text{constant},\qquad(18)$$

throughout. Considering Equation 6 at different points in the conjunction: at the entrance $\partial p / \partial x > 0$ and so the Poisuille term will act against the Couette flow, in the centre of the contact it has already been indicated that Poisuelle flow is minimal and

220   Couette flow dominant (so mass flow $\sim \rho \tilde{u} h$); however, in the exit region, decreasing pressures ($\partial p / \partial x < 0$) lead to Poisuelle and Couette terms acting in the same direction while, simultaneously, rapid reductions in viscosity are taking place - increasing the Poisuelle term magnitude. From Equations 6 and 18 it is clear that in order to avoid flow discontinuity, a reduction in h or ρ (or both) must occur. In practise, it has been found that a marked reduction[6] in film thickness occurs close to the outlet, as shown in Figure 2, in both incompressible and compressible cases. In the latter case the drop in film thickness ($h_c \rightarrow h_m$) is

225   slightly reduced while the pressure spike magnitude tends to dramatically reduce, indicating variations in h and ρ both play a role (Venner and Bos, 1994).

For line contacts the features shown in Figure 2 are present along much of the roller length, with distortions known to occur at roller ends (Wymer and Cameron, 1974). For point contacts the same features are present, but arranged in a *horseshoe* which tracks the elliptical contact patch boundary. Central film thickness, representative of much of the conjunction, is still at the

230   contact centre; while minimum film thickness tends to occur at two side-lobes, away from the centreline (Foord et al., 1969). Graphical depictions of typical film thickness variations in point and line contacts are shown in Figure 3. The horseshoe and side-lobe characteristics of point contacts arise due to the presence of Poiseuille-term driven lateral flow (side-leakage). In such

---

[6]$(h_c - h_m)/h_c \times 100\%$ values of between 17 and 70% have been reported in the literature (Chaomleffel et al., 2007), with operating conditions being a strong driver.
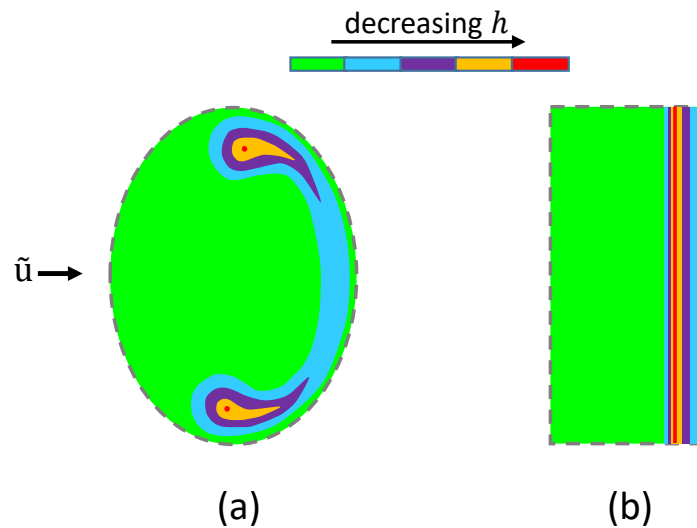
Figure 3. Typical film thickness variations in (a) point and (b) line contacts, end effects are not shown in the latter case. Note, these depictions are purely conceptual, film height values in one should not be interpreted as necessarily being equal to those in the other.

cases, mass-flow continuity[7] again indicates the presence of a constricted band aligned perpendicular to conjunction outflow, but, with outflow velocities now vector values containing u and v components (not to be confused with $\tilde{u}$ and $\tilde{v}$, see Figure

235   1). The perspective of relative (Poiseuille vs Couette) flow contributions, and implications for u and v values seen at different points, is important when interpreting effects of load, speed and ellipticity on described contact features; an excellent account of such analysis can be found in Wheeler et al. (2016a).

Additional relevant characteristics of EHL contacts (both point and line) are as follows:

1. Both minimum, h_m, and central, h_c, film thickness values are important for understating conditions within the contact

240   conjunction. The former, combined with surface roughness information, indicates the degree to which separation of surface asperities has been achieved; while the latter allows good representation of traction/friction conditions throughout most of the almost parallel gap.

2. Entrainment velocity ($\tilde{u}$) is known to be the main driver of lubricant film thickness. The relative effect of load is significantly smaller, attributable to the fact that load changes coincide with an expansion or contraction of the contact patch.

245   Material properties are not insignificant, but in practise only a narrow range of values will apply in any given rolling bearing situation.

3. As load increases and/or entrainment velocity decreases, surface geometries and pressures converge to those of dry Hertzian contact. The pressure spike also reduces such that maximum pressure occurs at the contact centre and equals that of dry contact.

---

[7]Following a similar argument to that outlined above.

250   4. As the ellipticity of point contact geometry increases[8], the elliptical conjunction (Figure 3a) tends asymptotically to a line
contact conjunction (Figure 3b). This effect can be understood in the context of relative flow, since $v \to 0$ as ellipticity
increases.

7   Dimensionless groupings and film thickness equations

When modelling a physical system via a set of equations, such as for EHL, it is possible to re-express the problem in an
255   equivalent dimensionless form which generally depends on a reduced number of, also dimensionless, parameters. These di-
mensionless parameters are constructed as appropriate products, powers and ratios of dimensional quantities appearing in the
original set of equations. In reduced and dimensionless form the problem is simplified and generalised, with effects of interact-
ing physical phenomena elucidated. To illustrate this last point, consider a system for which solutions depends on parameters
q_1 and q_2 with physical units in common. It may be the case that system response (e.g. flow rate, wave height etc.) is propor-
260   tionately increased by q_1, but decreased by q_2. In such a scenario, ultimate response is driven by the dimensionless quantity
q_1/q_2, rather than each value independently. The relevant number of parameters to characterise response is therefore reduced,
and the interaction between effects associated with q_1 and q_2 made clear. *Dimensional analysis* or *similarity analysis* are the
names given to the study and application of such ideas and associated methods.

In EHL, the following parameters comprise the most common set of dimensionless groupings used to describe lubrication
265   conditions in line contact conjunctions:

$$W_l = \frac{w_l}{E' R_x} \qquad \text{(load)} \tag{19}$$

$$U = \frac{\mu_0 \tilde{u}}{E' R_x} \qquad \text{(speed)} \tag{20}$$

$$G = \alpha^* E' \qquad \text{(material)}. \tag{21}$$

Recall that w_l is load per unit length. α^* is the lubricants inverse asymptotic isoviscous pressure coefficient, $1/\alpha^* = \int_0^\infty \frac{\mu_0}{\mu} dp$,
270   a quantity which can be directly determined for a given lubricant using high-pressure viscometer measurements.

---

[8]Elongating Figure 3a vertically.

$\underline{\alpha^* \text{ vs } \alpha}$

The pressure-viscosity coefficient α, generally obtained via curve-fitting to measurements, has commonly been used instead of α^* for EHL analyses. Indeed, for a lubricant that really did follow a Barus law pressure-viscosity curve (see Equation 12) it is easily shown that $\alpha^* = \alpha$, with the substitution therefore appropriate in this case. However, in practise the Barus law does not provide a good characterisation of viscosity variations with pressure. The coefficient α^*, first proposed by Blok in 1963, characterises a lubricants pressure-viscosity behaviour as a single parameter, but without assuming any particular functional form for $\mu(p)$. Other candidate coefficients have also been proposed in the literature, but, for the purposes of estimating EHL film thickness (using existing formulas) it has been demonstrated that the inverse asymptotic isoviscous pressure coefficient, α^*, is the one which should be used (Vergne and Bair, 2014). All such coefficients are temperature dependent, with the α^* used in film thickness equations necessarily being the value corresponding to the lubricant inlet temperature, T. For further discussion of these aspects of EHL modelling, see Bair (1993); Vergne and Bair (2014).

In the case of point contacts, speed and material parameters remain unchanged whereas dimensionless load becomes,

$$W = \frac{w}{E' R_x^2}, \tag{22}$$

where w is the applied load. Point contact geometry is captured in dimensionless form by the ellipticity parameter, k, or equivalently by $D = R_y / R_x$. Equivalence follows from the fact that k can be expressed as a function of D only (Masjedi and Khonsari, 2015). Non-dimensional film thickness is given by $H = h / R_x$.

Dimensionless groupings are commonly identified using a combination of intuition, experience and trial-and-error. However, systematic processes exist by which a minimal, or "optimal", set of dimensionless quantities can be identified that fully describe system behaviour (Moes, 1992; Hsiao, 2001). It should be emphasised that identification processes, and the resulting optimal dimensionless sets, depend on the governing equations and boundary conditions of a problem only. Note, also, that minimal sets of dimensionless quantities for a given problem tend not to be unique. While the number of elements in each minimal set will be the same, alternative choices for the groupings of dimensional variables are generally present. All alternative groupings which form a minimal set can be identified if required (Hsiao, 2001).

For the EHL problem in line and points contacts, optimal parameter analysis reduces the three parameters above (load, speed and material) to just two, a load parameter,

$$M_l = w_l \left(\frac{1}{\mu_0 \tilde{u} E' R_x} \right)^{1/2} = W_l (2U)^{-1/2} \quad \text{(line contact)} \tag{23}$$

$$M = w \left(\frac{1}{\mu_0^3 \tilde{u}^3 E' R_x^5} \right)^{1/4} = W (2U)^{-3/4} \quad \text{(elliptical contact)}, \tag{24}$$

and viscosity parameter,

$$L = \alpha^* \left(\frac{\tilde{u} \mu_0 E'^3}{R_x} \right)^{1/4} = G (2U)^{1/4}. \tag{25}$$

290 The factors of 2 are present to coincide with Moes (1992), in which entrainment speed is taken to be the sum, rather than mean, of surface velocities. While this is the form generally used (Marian et al., 2020; Wheeler et al., 2016b), the parameters are equally valid with the factors of 2 removed. Both forms have appeared in the literature (Moes, 1992; Hsiao, 2001), hence, it is important to ascertain which has been applied if comparing operating conditions or applying related film thickness equations. Non-dimensional film thickness in the optimal parameter case takes the form $H = (h/R_x)(2U)^{-1/2}$. The same

295 analysis identifies D^{-1}, introduced above, as the parameter representing contact patch geometry in the elliptical case (Hsiao, 2001). Additional parameters are required to capture more complex viscosity and density characteristics, such as those given by the Roelands equation and Equation 13 (Hsiao, 2001). Despite the proven reduction in the number of parameters required to characterize EHL conditions, use of W_l (or W), U and G persists, although film thickness equations utilising reduced sets of parameters have been developed (Marian et al., 2020). The parameters M_l (or M) and L are, however, commonly used when

300 plotting operating regions and results, since visualisation becomes clearer and easier with a reduced number of variables.

Having identified a set of dimensionless parameters (optimal or otherwise) which determine the response of the system defined by governing equations, it follows that features of interest (e.g. h_m and h_c) will also be determined by these same parameters. If such relationships can be sufficiently well approximated by analytical equations, fast evaluation and analysis of key features becomes possible without requiring complex numerical solvers to be implemented in every case. Film thickness

305 formulas have therefore been developed by performing least-squares curve fits between outputs of full EHL solvers and analytical expressions containing dimensionless parameters. Equations 26 and 27 present two of the earlier equations identified this way for estimating minimum film thickness (h_m) in line and point contacts, respectively,

Dowson 1967 (line contacts) (Dowson, 1967):

$$\frac{h_m}{R_x} = 2.65 \frac{U^{0.70} \, G^{0.54}}{W_l^{0.13}} \tag{26}$$

310 **Hamrock and Dowson 1977 (point contacts)** (Hamrock and Dowson, 1977):

$$\frac{h_m}{R_x} = 3.63 \frac{U^{0.68} \, G^{0.49}}{W^{0.073}} \left(1 - e^{-0.68k}\right) \tag{27}$$

Despite the early stage at which they were developed, these equations provide remarkably accurate estimates and are still used today (Harris and Kotzalas, 2006a). Quite a number of subsequent refinements have been undertaken using larger datasets generated by more advanced solvers (Marian et al., 2020). Some of the most extensive fitting was undertaken for line and point

315 contacts in Masjedi and Khonsari (2012) and Masjedi and Khonsari (2015), respectively, resulting in the following equations:

Masjedi and Khonsari 2012 (line contacts) (Masjedi and Khonsari, 2012):

$$\frac{h_m}{R_x} = 1.652 \frac{U^{0.716} G^{0.695}}{W_l^{0.077}} \tag{28}$$

Masjedi and Khonsari 2015 (point contacts) (Masjedi and Khonsari, 2015):

$$\frac{h_m}{R_x} = 1.637 \frac{U^{0.711k^{-0.023}} \, G^{0.65k^{-0.045}}}{W^{0.09k^{-0.15}}} \left(1 - 0.974e^{-0.676k}\right) \tag{29}$$

320 Unfortunately, the full range of dimensionless parameter values over which these formulas were fitted appears to have been misrepresented in the literature. In Wheeler et al. (2016b), and then reproduced in Marian et al. (2020), parameter limits for the point contact h_m equation are given (approximately) as $15 \leq M \leq 10^4$ and $5 \leq L \leq 20$. While these limits are those of the cases shown in the results tables of Masjedi and Khonsari (2015), it is explicitly stated that only a subset of the full analysis is reproduced there. Moreover, the full range of dimensionless parameter values used for curve fitting are also given. Taking the

325 stated limiting values in Masjedi and Khonsari (2015), the domain across which these equations were fitted is in fact bounded by $2.82 \leq M \leq 1.47 \times 10^5$ and $2.97 \leq L \leq 28.20$. The true domain of validity for these equations is therefore significantly larger than has been reported. For completeness, parameter limits for the line contact equation (Masjedi and Khonsari, 2012) are also stated; limits in L match those of the point contact case and for M_l, $2 \leq M_l \leq 353.55$. With respect to ellipticity, Equation 29 was developed across the range $1 \leq k \leq 8$. At $k = 8$, film thickness predictions match closely those of the line

330 contact formula (Equation 28) applied to an "equivalent line contact" representation of the conjunction (see Section 4). It was therefore recommended that the equivalent line contact approach is taken for cases where $k > 8$ (Masjedi and Khonsari, 2015). Similar formulas to those presented in the current section have also been developed for estimating central film thickness, h_c.

7.1 Accuracy of film thickness equations

It is important to appreciate that analytical film thickness equations provide estimated values based on curve fitting within a

335 specific range of dimensionless parameter values, *i.e.* operating conditions. Applicability limits for any given equation must therefore be checked and respected for each case being analysed. Furthermore, isothermal conditions and Newtonian fluid behaviour are also often assumed. Correction factors for such effects have been proposed (Marian et al., 2020), these are subject to the same limitations as outlined for the film thickness equations themselves. In the context of curve-fitting, derived formulas are able to recreate the numerical results on which they are based to a high degree of accuracy. For example, comparisons

340 between Equation 29 and numerical fitting data result in a mean error of 3.27% and a maximum error of 9.79% (Masjedi and Khonsari, 2015); note, it is not clear whether these numbers relate to the full dataset or only a subset. Similarly, for the line contact h_m equation (Equation 28) a maximum error of 10.41% is reported (Masjedi and Khonsari, 2012). This is certainly promising, but, investigating accuracy at points not included in the fitting set is crucial to forming a full picture of equation performance. In Wheeler et al. (2016b) a range of point contact analytical film thickness equations for h_m and

345 h_c were compared in this way. Maximum observed errors across tested h_m equations occurred for the circular contact case ($k = 1$), reaching about 90%. In general, h_m was found to be overestimated by analytical equations, with the severity of over-estimation increasing as load increases or entrainment speed decreases. When ellipticity was increased to $k = 2.92$, errors in h_m predictions reduced to satisfactory levels (about 6% on average). The study concluded that Equation 29 (along with its companion h_c equation) should be used for cases of wide elliptical contacts ($k > 1$). Other equations are recommended for

350 circular contacts ($k = 1$), while the slender contact case ($k < 1$) remains an open problem (Wheeler et al., 2016b). The major conclusion of Wheeler et al. (2016b) is that current analytical equations must be considered as providing qualitative, rather than truly quantitative, estimates of film thickness. Note, consistent with the discussion concerning viscosity coefficients (above), Wheeler et al. (2016b) use the inverse asymptotic isoviscous pressure coefficient, α^*, throughout their analysis.

8 Starvation

355 Film thickness equations presented in the previous section assume fully-flooded conditions, in which an adequate supply of lubricant is available to the contact. Starvation is the term given to cases in which lubricant supply is insufficient. Starvation acts to reduce film thickness, with the severity of reduction linked to the distance of the lubricant inlet *meniscus* (the final point for which a continuous vertical column of lubricant is present between the surfaces) from the centre of the contact. Various formulas, based on numerical and experimental results, have been developed which allow film thickness effects of starvation

360 to be estimated (Harris and Kotzalas, 2006a; Poll et al., 2019), however, this tends to require knowledge of the meniscus distance. In general this quantity, and even the degree of starvation, will not be known. While important research in this area continues, an early contribution by Dowson and others provides a practical rule of thumb for when the degree of starvation is unknown. Through a range of experimental analyses for nominal line contacts it was demonstrated that EHL conjunctions, initially supplied with a fixed amount of lubricant, appear to tend towards an equilibrium state in which the inlet meniscus sits at

365 the point upstream of the contact centre where reverse flow ceases. Reverse flow is known to occur within contact inlets, away from bounding surfaces, driven by high pressure-gradients of the type shown in Figure 2. See Mohammadpour et al. (2014) for a depiction of such flow. For EHL of pure rolling line contacts under starvation levels consistent with the *zero-reverse-flow* condition, minimum film thickness values have been shown to reduce to about 71% of their fully flooded values (Hamrock and Dowson, 1981), reported as 70% in Harris and Kotzalas (2006a). This reduction factor is useful for when starvation levels are

370 unknown, since it has a physical basis. However, it should be interpreted as providing a ball-park estimate of starvation effects only since, in reality, film reductions due to starvation depends on bearing operating parameters, especially speed (Poll et al., 2019). Correction factors have also been developed to account for effects of starvation when estimating film thickness (Marian et al., 2020), but, additional information is often required that is not easily obtained outside of a laboratory setting. Starvation is known to occur commonly in grease lubricated roller bearings, hence, starvation will be revisited in this context in Section

375 11, below.

9 Dynamic effects in EHL

Many EHL contacts operate under non-steady conditions in which load, speed and even contact curvature (the latter being the case during gear meshing) change with time. The critical timescale, t_c, determining the impact of such variations is the time it takes for a particle of lubricant to pass through the contact (Venner and Wijnant, 2005). When the time taken for conditions

380 to vary is large, compared to t_c, non-steady effects become negligible and a quasi-static analysis is sufficient for characterising film variations over time. For example, this is generally the case for roller bearings operating under constant load, in which each roller sees a continuous variation in applied force as it traverses the loaded zone. In the absence of additional effects, the film thickness variations are well captured by applying steady-state equations (Section 7) at each time step. However, if variations occur rapidly, meaning over timescales similar to or shorter than t_c, local surface deformations are induced at the inlet which

385 then propagate through the contact. Dynamic EHL phenomena have been studied numerically, including via semi-analytical models, and experimentally (Hooke, 2003; Venner and Wijnant, 2005; Morales-Espejel, 2008; Zhang et al., 2020). Steady-state

formulae do not capture variations in film thickness and pressure within the contact resulting from dynamics effects, their use in such cases therefore risks over-estimating minimum film thickness values (Hooke, 2003). Piezoviscous behaviour plays an important role in dynamic EHL. Recall, as discussed in Section 6, that very high viscosities and Couette flow dominance hold

390 in the central region of a loaded EHL contact. As the lubricant film passes into the central region, it therefore becomes very "stiff" and moves through the conjunction, at approximately $\tilde{u}$, as a shear flow that is independent (for the most part) of the contact inlet. In the case that entrainment speed is suddenly increased, increased film thicknesses occur at the contact inlet (via local surface deformations) which are then carried through the conjunction at the new entrainment speed (Hooke, 2003). The new, increased, film thickness values are only seen across the whole contact once the original values have passed out

395 of the conjunction. While the increased entrainment speed ($\tilde{u}$) is seen simultaneously across the whole contact, a greater film thickness requires an increased volume of lubricant in the conjunction. For a shear dominant flow, this can only happen through the propagation process described above, irrespective of pressure changes at the inlet. This explains why film changes from rapid speed adjustments aren't uniform across the contact. While in steady-state conditions film thickness sensitivity to load is known to be small, the effects of rapid load variations can be dramatic (Venner and Wijnant, 2005; Hooke, 2003). As load

400 increases, the width (b) of the contact also increases, meaning the contact edge is moving rapidly in the opposite direction to entrainment. The result is an augmented "effective" entrainment speed,

$$\tilde{u}_{\text{eff}} = \tilde{u} + \frac{db}{dt}, \tag{30}$$

for the inlet. Due to the stiff nature of the central film, this results in the formation of a dome-shaped entrapment of lubricant at the inlet which is then carried through the conjunction. Similarly, a rapid reduction in load results in negative values of

405 db/dt, and so a reduced effective entrainment speed at the inlet. In this case, a local drop in film thickness forms at the inlet and moves through the contact (Venner and Wijnant, 2005). Somewhat counter intuitively, it is therefore the case that rapid increases in load can temporarily increase the minimum film thickness, while rapid load reductions can temporarily decrease the minimum film thickness. Periodic load variations have also been considered in the literature. Film height variations through the contact are accompanied by local variations in pressure and material stress (Hooke, 2003). For time-varying loads, the rate

410 of expansion or contraction, db/dt, relative to entrainment speed, $\tilde{u}$, has been proposed as a criteria for determining whether a quasi-static analysis may be used (Hooke, 2003). Dynamic effects were found to become important where,

$$\frac{|db/dt|}{\tilde{u}} (\times 100\%) \geq 25\%. \tag{31}$$

While effective entrainment speed plays an important role in dynamic EHL, squeeze film effects ($\partial(\rho h)/\partial t$ in Equation 9) also strongly influence the resulting film thickness variations over time (Zhang et al., 2020). Both effects must therefore be

415 accounted for.

10 Surface roughness interactions

As discussed in Section 2, film thickness must be considered relative to surface roughness in order for the lubrication regime to be known. However, it is also the case that rough surface micro-geometry will interact with the lubricant flow and deform

elastically, with both effects influencing surface separation and lubrication conditions. Much work has been undertaken over
the years in what is now known as micro-EHL, leading to significant advances in understanding and modelling capabilities. An
excellent overview is provided by Morales-Espejel (2014). The presence of roughness results in part of the load being carried
by surface asperities (Masjedi and Khonsari, 2015), as opposed to being carried purely hydrodynamically. Such interactions
are important when considering micropitting of machine elements (Morales-Espejel, 2014). With respect to conditions in
the lubricated conjunction, roughness has been shown to result in both "mean" and "local effects". Mean effects are overall
modifications to surface separation and pressure, relative to an equivalent smooth contact. Local effects are in the form of
local variations in film height and pressure which move through the contact and, hence, are dynamic in nature. The mean effect
resulting from the presence of homogeneous surface roughness is an increase in film thickness, but, by an amount that is smaller
than the change in surface σ (Masjedi and Khonsari, 2015). Therefore when roughness increases, h_m increases, but Λ decreases.
Note, more structured roughness can have a different effect (Morales-Espejel, 2014). Film thickness equations presented in
Section 7 are those for smooth surfaces. Additional multiplicative factors have been identified, also via curve fitting, which
account for surface roughness effects. For the Masjedi and Khonsari line and point contact minimum film thickness equations
these, respectively, take the form (Masjedi and Khonsari, 2015, 2012):

$$\Phi_l \quad = \quad 1 + 0.026 \left(\frac{s_{\text{std}}}{R_x}\right)^{1.120} V^{0.185} W_l^{-0.312} U^{-0.809} G^{-0.977} \tag{32}$$

$$\Phi_p \quad = \quad 1 + 0.141 \left(\frac{s_{\text{std}}}{R_x}\right)^{1.073} V^{0.149} W^{-0.044} U^{-0.828} G^{-0.954} k^{-0.395}, \tag{33}$$

where s_{std} is the standard deviation of surface heights (assuming normally distributed roughness, $s_{\text{std}} = \sigma$) and $V = v_h/E'$ is
a dimensionless hardness number, with v_h the surface Vickers hardness. From analysis across standard ranges of operating
parameters, line and point contact results for dimensionless roughness levels, s_{std}/R_x, of around $1 \times 10^{-6} - 5 \times 10^{-6}$ or
less have been shown to be well approximated by smooth surface results (Masjedi and Khonsari, 2015, 2012). Note, the above
modification factors may be applied where $\Lambda > 0.5$, but are no longer valid if the film parameter falls below this value (Masjedi
and Khonsari, 2015, 2012).

11   Grease lubrication

The vast majority of rolling element bearings are grease lubricated, where "grease" may be defined as a dispersion of a
thickening agent in a liquid lubricant (Lugt, 2009). Lubricant base oil is held inside sponge-like structures of thickener fibres
through a combination of Van der Waal and capillary forces. The resulting semi-solid consistency is beneficial due to its ease
of use, good sealing action and corrosion resistance. However, this same consistency generally leads to starved lubrication
conditions, since grease will not reflow (at a macroscopic level) back to the rolling track after being swept out by the passage
of rolling elements. The total quantity of grease directly participating in the separation of contacting surfaces is therefore
reduced. This initial phase of grease redistribution, the "churning phase", occurs within the first ~10 hours of operation after a
bearing has been freshly charged with grease (Cen and Lugt, 2020). Once this initial grease flow has ceased, the bearing enters
the "bleeding phase" in which swept grease reservoirs are generally only able supply lubricant to the contacts by releasing

oil through phase separation (Lugt, 2016). Understanding, modelling and predicting grease lubrication is difficult. This is because thickener and base oil interactions result in nonlinear shear stress - shear rate behaviour, even at low shear rates and pressures. The apparent viscosity of grease also decreases continuously over time while being sheared, then recovers once shearing stops (thixotropy) (Lugt, 2009, 2013). Over longer timescales, grease properties change as the thickener structure

455   deteriorates due to being mechanically worked (Cen et al., 2014). Oxidation also slowly degrades grease performance. Due to these complexities, there is as yet no complete theory which allows film thickness in grease lubricated conjunctions to be accurately and consistently predicted. The difficulties associated with numerically modelling grease behaviour has meant that grease lubrication has mostly been investigated experimentally. While a complete theory of grease lubrication is not yet established, significant advances have been made regarding the key mechanisms and interactions at work. A summary of

460   pertinent results in this field will therefore be outlined:

1) *The thickener contributes to film thickness*: At higher speeds the film thickness observed in fully-flooded grease lubricated contacts coincides with that of oil lubrication (with the grease base-oil) (Kanazawa et al., 2017). Hence, at higher speeds fully-flooded grease lubrication is entirely determined by the viscosity of the base oil. As speed is reduced, film thickness initially reduces inline with the behaviour predicted by standard film thickness equations (See Section 7), but, eventually a "transition

465   speed" is reached after which further reductions in speed result in increasing film thickness values. The rate of increase in this region, as speed is decreased, is similar to that seen for oil lubrication as speed is increased, meaning these low speed grease effects are significant with respect to resulting film thicknesses. Speed versus film thickness plots for fully-flooded grease lubrication therefore exhibit a characteristic 'V' shape (Cen et al., 2014; Kanazawa et al., 2017). The described behaviour at low speeds results from entrained thickener fibers becoming the dominant driver of surface separation, further evidenced by the

470   fact that in this region the film thickness, for a given grease, is independent of base oil viscosity and temperature (Kanazawa et al., 2017). As the grease is mechanically worked over time, the thickener structure degrades and constituent particles become smaller. This reduces film thickness values seen in the low speed region, whereas the higher speed region is unaffected (Cen et al., 2014). The degradation process has been found to primarily occur within the first 100h of operation. The transition from low-speed to higher-speed behaviour has been found to be dependent on film thickness. More specifically, experimental

475   findings indicate that the described "low-speed" effects occur for film thicknesses $h < c\mathcal{D}$, for some constant c and where $\mathcal{D}$ is the diameter of thickener fibres or possibly that of entangled fibre networks (Kanazawa et al., 2017). Surface deposits of thickener material, which contribute to overall film thickness and may be important in cases of starved lubrication (Cann, 1996), have been shown to be present in operating regions where thickener effects dominate (Lugt, 2009; Kanazawa et al., 2017). Impacts of grease lubrication on friction have also been considered (Kanazawa et al., 2017).

480   2) *Grease lubrication is non-deterministic*: Grease lubricated bearings operating under steady conditions have been shown to experience film thickness fluctuations (also referred to as "events") within otherwise stable film thickness time-histories. Such events are caused by new grease/bled-oil entering the contacts as a result of frictional heating, creep flow, vibrations or cage interactions (Lugt, 2009; Cen and Lugt, 2019). These effects are more pronounced at high speeds.

3) *Contact replenishment occurs in grease lubricated bearings, but as a strictly local phenomenon*: While grease lubricated
bearings generally operate under starved conditions, the level of starvation (and hence the film thickness) has been shown
to stabilise during continuous operation (Cen and Lugt, 2019). The observed balance between "feed and loss" mechanisms
requires significant replenishment of oil to the contacts. Importantly, starvation stabilisation cannot be accounted for by grease
models which assume the oil available to a contact is that left by the preceding contact, combined with additional oil from
track replenishment between ball/roller passes (Cen and Lugt, 2020). Similarly, rates of bearing grease oil-loss observed in
practise confirm that grease bleed is too slow to be the source of the significant replenishment taking place. Recent work has
explored ball-bearing contact replenishment in real bearings (Cen and Lugt, 2020). For bearings with different numbers of
balls (hence different timescales between contact passes) and different cage geometries, the normalised film thickness (relative
to fully-flooded conditions) was found to be a function of the parameter *speed × viscosity × half contact-width*. Since these
quantities are local to the contact and independent of the time between ball passes, it follows that contact replenishment must
therefore be a purely local phenomenon. Local replenishment likely takes place under the action of capillary forces and/or
ball spin among other possible mechanisms (Cen and Lugt, 2020), with replenishment rates determined by viscosity. For a
full bearing operating under starvation it has been found that increases in speed and/or viscosity lead to increased levels of
starvation (Cen and Lugt, 2019, 2020) and reduced film thickness values (Cen and Lugt, 2020). This is in stark contrast to
the fully flooded case. Interestingly, overall sensitivity of film thickness to speed and viscosity is much reduced under starved
conditions in a real bearing (Cen and Lugt, 2020). This implies that base-oil viscosity may be less important than might be
expected when selecting grease for a given application. This recently described work brings closer the prospect of a general
formula for predicting film thickness in grease lubricated bearings, however, further work remains before this goal may be
realised.

In addition to the above, a recent experimental study considered dynamic effects in grease lubricated starved contacts, showing
that behaviour similar to that outlined in Section 9 can also be observed for grease lubrication (Zhang and Glovnea, 2020).

As previously stated, understanding of grease lubrication has advanced significantly. It should be noted that much of the
recent progress, outlined here, has been developed using ball bearings either as single contacts or within full bearing test rigs
(although much important work has also been undertaken for roller bearings (Lugt, 2009, 2013)). The relative importance of
described effects may therefore differ in practise, depending on the roller type, the characteristics of applied loading and the
distribution of load within the bearing *etc*. While general film thickness formulas for grease lubrication are not yet available,
there is a consensus that grease lubricated bearings are usually operating under starved conditions. One study proposed that
grease lubricated film thickness can be estimated as 70% of the fully flooded value (under oil lubrication), assuming a viscosity
equal to that of the base-oil (Lugt, 2009). This coincides with the film reduction proposed for starved oil lubrication under
conditions of zero-reverse-flow (see Section 8). However, given the levels of film reduction (relative to the fully flooded case)
recently observed in real bearings (Cen and Lugt, 2019, 2020), the 70% factor may lead to overestimation of the film thickness.
It should therefore be treated as a best-case-scenario estimate only.

12 Discussion

EHL theory has come an exceptionally long way since its inception around 1950. Incredibly complex EHL problems are now solved routinely, with experimental comparisons demonstrating the effectiveness of currently available models. However,
520 important further work remains. It is now understood that some of the simplifying assumptions applied in classical EHL, and still used today, must be challenged. Much of this is centred around the need for more realistic modelling of lubricant rheological behaviour (Bair et al., 2016; Bair, 2019). It is argued that these aspects of EHL must be properly accounted for before it can be considered a truly quantitative discipline (Vergne and Bair, 2014; Bair et al., 2016; Bair, 2019). Film thickness estimates from existing formulas must therefore necessarily be interpreted as being qualitative in nature (Wheeler et al., 2016b)
525 and, as has been highlighted, their applicability limits in each case must be checked and respected. Understanding of grease behaviour and grease lubrication has also advanced significantly but, given the additional complexities present in this case, further work is needed before even qualitative estimates of film thickness in general grease lubricated contacts can be made. As stated in Section 1, EHL is therefore (and is likely to remain) a complex, nuanced and rapidly evolving field.

13 Conclusions

530 This paper has presented an accessible introductory review of elastohydrodynamic lubrication theory, including both fundamental concepts and more advanced topics. As far as possible, the nuances and complexities of existing theory was discussed. It is hoped that this review will promote improved cross-disciplinary understanding with respect to elastohydrodynamic lubrication analysis. The theory presented here is applied in "Part 2" of the study in order to consider lubrication in a wind turbine main-bearing.

535 Appendix A: Table of symbols

a	semi-major contact dimension, assumed here to lie transverse (y) to the rolling direction (m)
b	semi-minor contact dimension, assumed here to lie in (x) the rolling direction (m)
$E_\mathrm{I}, E_\mathrm{II}$	Youngs modulii of solids I and II (Pa)
E'	reduced modulus of elasticity (Pa), $2/E' = (1 - \nu_\mathrm{I}^2)/E_\mathrm{I} + (1 - \nu_\mathrm{II}^2)/E_\mathrm{II}$
G	dimensionless material parameter (-), $G = \alpha^* E'$
h, h_m, h_c	film thickness, minimum film thickness and central films thickness, respectively (m)
k	ellipticity parameter (-), $k = a/b$
L	dimensionless viscosity parameter (Moes) (-), $L = G(2U)^{1/4}$
M_l, M	dimensionless load parameter (Moes) for line and point contacts respectively (-), $M_l = W_l(2U)^{-1/2}, M = W(2U)^{-3/4}$
p	pressure (Pa)
$r_{Ix}, r_{IIy}, \ldots$	radius of curvature of surface I/II in the x/y direction, a strictly positive quantity (m)
R_x	reduced radius of curvature in the entrainment direction (m), $\frac{1}{R_x} = \frac{sgn(\mathrm{I}x)}{r_{\mathrm{I}x}} + \frac{sgn(\mathrm{II}x)}{r_{\mathrm{II}x}}$
R_y	reduced radius of curvature transverse to the entrainment direction (m), $\frac{1}{R_y} = \frac{sgn(\mathrm{I}y)}{r_{\mathrm{I}y}} + \frac{sgn(\mathrm{II}y)}{r_{\mathrm{II}y}}$
$sgn(\cdot)$	$sgn(\mathrm{I}x)$ is 1 if surface I is convex in the x-direction and -1 if it is concave in the x-direction (similarly for II and/or y)
s_l, s_p	surface stress in line and point contacts, respectively (Pa)
s_std	standard deviation of surface heights (assuming normally distributed roughness, $s_\mathrm{std} = \sigma$) (m)
T	lubricant inlet temperature (°C)
$u_\mathrm{I}, u_\mathrm{II}$	tangential velocities, in the entrainment direction (x), of surfaces I and II at the contact location (m/s)
$\tilde{u}$	mean entrainment velocity (m/s), $\tilde{u} = (u_\mathrm{I} + u_\mathrm{II})/2$
U	dimensionless speed parameter (-), $U = \mu_0 \tilde{u}/(E' R_x)$
$v_\mathrm{I}, v_\mathrm{II}, \tilde{v}$	similar to "u" terms, but transverse (y) to entrainment direction (m/s)
V	dimensionless hardness number, $V = v_h/E'$, for v_h the surface Vickers hardness (-)
w	normal load in point contact (N)
w_l	normal load per unit length in line contact (N/m)
W_l, W	dimensionless load parameter for line and point contacts respectively (-), $W_l = w_l/(E' R_x), W = w/(E' R_x^2)$
α	pressure-viscosity coefficient of the lubricant (at the inlet temperature, T) (Pa^{-1})
α^*	inverse asymptotic isoviscous pressure coefficient (at the inlet temperature, T) (Pa^{-1}), $1/\alpha^* = \int_0^\infty \frac{\mu_0}{\mu} dp$
Λ	lubrication film parameter (-), $\Lambda = h_m/\sigma$
μ	lubricant dynamic viscosity (Pa·s)
μ_0	lubricant dynamic viscosity at the inlet temperature, T, and for (gauge-pressure) $p = 0$ (Pa·s)
$\nu_\mathrm{I}, \nu_\mathrm{II}$	Poisson's ratios of solids I and II (-)
ρ	lubricant density (kg·m^{-3})
ρ_0	lubricant density at the inlet temperature, T, and for (gauge-pressure) $p = 0$ (kg·m^{-3})
$\sigma_\mathrm{I}, \sigma_\mathrm{II}$	surface roughness, in the form of root-mean-square deviations, for surfaces I and II respectively (m)
σ	combined roughness of contacting surfaces (m), $\sigma = \sqrt{\sigma_\mathrm{I}^2 + \sigma_\mathrm{II}^2}$
Φ_l, Φ_p	film thickness modification factors accounting for surface roughness effects, line and point contacts respectively (-)

Appendix B: Equivalent line contact formulation

Starting from Equations 15 and 16 and reducing s_e to the case $y = 0$, the resulting stress distributions can only be identical if
contact widths (b) are the same and if $w_l = 3w/4a$. Therefore, it only remains to ensure that the contact width for the equivalent
line contact, under distributed load $w_l = 3w/4a$, matches that of the elliptical contact under applied load w. Elliptical contact
semi-minor and semi-major axes, for the applied load w, take the following forms,

$$b = \left(\frac{6w\mathcal{E}}{\pi k \Sigma_\rho E'} \right)^{1/3}, \tag{B1}$$

$$a = \left(\frac{6k^2 w\mathcal{E}}{\pi \Sigma_\rho E'} \right)^{1/3}. \tag{B2}$$

$\Sigma_\rho = 1/R_x + 1/R_y$ is the point contact curvature sum. The line contact with equivalent x-direction geometry has curvature
sum $\Sigma_{\rho,\text{line}} = 1/R_x$. The line contact with this geometry and reduced modulus $\tilde{E}'$ sees the following contact width, under
distributed load w_l,

$$b = \left(\frac{8 w_l R_x}{\pi \tilde{E}'} \right)^{1/2}. \tag{B3}$$

Note, Equations B1-B3 all constitute standard formulae in the Hertzian contact theory. Equating Equations B1 and B3, having
substituted $w_l = 3w/4a$ and with a given by Equation B2, it follows that $\tilde{E}' = \left(1 + \frac{R_x}{R_y} \right) \mathcal{E}^{-1} E'$.

Competing interests. The authors declare they have no competing interests.

References

Albahrani, S., Philippon, D., Vergne, P., and Bluet, J.: A review of in situ methodologies for studying elastohydrodynamic lubrication, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 230, 86–110, 2016.

Antoine, J.-F., Visa, C., Sauvey, C., and Abba, G.: Approximate Analytical Model for Hertzian Elliptical Contact Problems, Journal of
555   Tribology, 128, 660–664, https://doi.org/10.1115/1.2197850, 2006.

ASTM, 2020: D341 - 20e1 Standard Practice for Viscosity-Temperature Equations and Charts for Liquid Petroleum or Hydrocarbon Products, Standard, ASTM International, West Conshohocken, PA, 2020.

Bair, S.: An experimental verification of the significance of the reciprocal asymptotic isoviscous pressure for EHD lubricants, Tribology transactions, 36, 153–162, 1993.

560 Bair, S.: The rheological assumptions of classical EHL: what went wrong?, Tribology International, 131, 45–50, 2019.

Bair, S., Martinie, L., and Vergne, P.: Classical EHL versus quantitative EHL: a perspective part II—super-Arrhenius piezoviscosity, an essential component of elastohydrodynamic friction missing from classical EHL, Tribology Letters, 63, 1–10, 2016.

Brewe, D. and Hamrock, B.: Simplified Solution for Elliptical-Contact Deformation Between Two Elastic Solids, J. of Lubrication Tech., 99, 485–487, 1977.

565 Cann, P.: Starvation and reflow in a grease-lubricated elastohydrodynamic contact, Tribology Transactions, 39, 698–704, 1996.

Cen, H. and Lugt, P.: Film thickness in a grease lubricated ball bearing, Tribology international, 134, 26–35, 2019.

Cen, H. and Lugt, P.: Replenishment of the EHL contacts in a grease lubricated ball bearing, Tribology International, 146, 106 064, 2020.

Cen, H., Lugt, P., and Morales-Espejel, G.: Film thickness of mechanically worked lubricating grease at very low speeds, Tribology Transactions, 57, 1066–1071, 2014.

570 Chaomleffel, J.-P., Dalmaz, G., and Vergne, P.: Experimental results and analytical film thickness predictions in EHD rolling point contacts, Tribology International, 40, 1543–1552, 2007.

Dlugoš, J. and Novotný, P.: Effective Implementation of Elastohydrodynamic Lubrication of Rough Surfaces into Multibody Dynamics Software, Applied Sciences, 11, 1488, 2021.

Dowson, D.: A generalized Reynolds equation for fluid-film lubrication, International Journal of Mechanical Sciences, 4, 159–170, 1962.

575 Dowson, D.: Elastohydrodynamics, in: Proceedings of the Institution of Mechanical Engineers, Conference Proceedings, vol. 182, pp. 151–167, 1967.

Dowson, D. and Higginson, G.: A numerical solution to the elasto-hydrodynamic problem, Journal of mechanical engineering science, 1, 6–15, 1959.

Evans, H. and Hughes, T.: Evaluation of deflection in semi-infinite bodies by a differential method, Proceedings of the Institution of Me-
580   chanical Engineers, Part C: Journal of Mechanical Engineering Science, 214, 563–584, 2000.

Foord, C., Wedeven, L., Westlake, F., and Cameron, A.: Optical elastohydrodynamics, Proceedings of the Institution of Mechanical Engineers, 184, 487–505, 1969.

Greenwood, J.: An extension of the Grubin theory of elastohydrodynamic lubrication, Journal of Physics D: Applied Physics, 5, 2195, 1972.

Guo, Y., Bankestrom, O., Bergua, R., Keller, J., and Dunn, M.: Investigation of main bearing operating conditions in a three-Point mount
585   wind turbine drivetrain, Forschung im Ingenieurwesen, 85, 405–415, 2021.

Habchi, W.: Finite element modeling of elastohydrodynamic lubrication problems, John Wiley & Sons Ltd., 2018.

Habchi, W., Eyheramendy, D., Vergne, P., and Morales-Espejel, G.: A full-system approach of the elastohydrodynamic line/point contact problem, Journal of tribology, 130, 2008.

Hamrock, B. and Dowson, D.: Isothermal Elastohydrodynamic Lubrication of Point Contacts: Part III—Fully Flooded Results, J. of Lubrication Tech., 99, 264–275, 1977.

Hamrock, B. and Dowson, D.: Ball bearing lubrication - the elastohydrodynamics of elliptical contacts, John Wiley and Sons, 1981.

Hamrock, B., Schmid, B., and Jacobson, B.: Fundamentals of fluid film lubrication, vol. 169, CRC press, 2004.

Harris, T. and Kotzalas, M.: Advanced Concepts of Bearing Technology: Rolling Bearing Analysis, CRC Press, 2006a.

Harris, T. and Kotzalas, M.: Essential concepts of bearing technology, CRC press, 2006b.

Hart, E.: Developing a systematic approach to the analysis of time-varying main bearing loads for wind turbines, Wind Energy, 23, 2150–2165, 2020.

Hart, E., Turnbull, A., Feuchtwang, J., McMillan, D., Golysheva, E., and Elliott, R.: Wind turbine main-bearing loading and wind field characteristics, Wiley Wind Energy, 22, 1534–1547, https://doi.org/10.1002/we.2386, 2019.

Hart, E., Clarke, B., Nicholas, G., Kazemi Amiri, A., Stirling, J., Carroll, J., Dwyer-Joyce, R., McDonald, A., and Long, H.: A review of wind turbine main bearings: design, operation, modelling, damage mechanisms and fault detection, Wind Energy Science, 5, 105–124, https://doi.org/10.5194/wes-5-105-2020, 2020.

Holmes, M., Evans, H., Hughes, T., and Snidle, R.: Transient elastohydrodynamic point contact analysis using a new coupled differential deflection method part 1: theory and validation, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 217, 289–304, 2003.

Hooke, C.: Dynamic effects in EHL contacts, in: Tribology Series, vol. 41, pp. 69–78, Elsevier, 2003.

Hooke, C.: A review of the paper 'A numerical solution to the elastohydrodynamic problem' by D. Dowson and GR Higginson, Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 223, 49–63, 2009.

Hsiao, H.-S. S.: Optimum dimensional analysis with applications to elastohydrodynamic lubrication, J. Trib., 123, 822–827, 2001.

Hughes, T., Elcoate, C., and Evans, H.: Coupled solution of the elastohydrodynamic line contact problem using a differential deflection method, Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 214, 585–598, 2000.

Johnson, K.: Contact mechanics, Cambridge University Press, 1987.

Kanazawa, Y., Sayles, R., and Kadiric, A.: Film formation and friction in grease lubricated rolling-sliding non-conformal contacts, Tribology International, 109, 505–518, 2017.

Lubrecht, A., Venner, C. H., and Colin, F.: Film thickness calculation in elasto-hydrodynamic lubricated line and elliptical contacts: the Dowson, Higginson, Hamrock contribution, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 223, 511–515, 2009.

Lugt, P.: A review on grease lubrication in rolling bearings, Tribology Transactions, 52, 470–480, 2009.

Lugt, P.: Modern advancements in lubricating grease technology, Tribology international, 97, 467–477, 2016.

Lugt, P. and Morales-Espejel, G.: A review of elasto-hydrodynamic lubrication theory, Tribology Transactions, 54, 470–496, 2011.

Lugt, P. M.: Grease lubrication in rolling bearings, John Wiley & Sons, 2013.

Marian, M., Bartz, M., Wartzack, S., and Rosenkranz, A.: Non-Dimensional Groups, Film Thickness Equations and Correction Factors for Elastohydrodynamic Lubrication: A Review, Lubricants, 8, 95, 2020.

Masjedi, M. and Khonsari, M.: Film thickness and asperity load formulas for line-contact elastohydrodynamic lubrication with provision for surface roughness, Journal of tribology, 134, 2012.

625   Masjedi, M. and Khonsari, M.: On the effect of surface roughness in point-contact EHL: Formulas for film thickness and asperity load, Tribology International, 82, 228–244, 2015.

Meng, Y., Xu, J., Jin, Z., Prakash, B., and Hu, Y.: A review of recent advances in tribology, Friction, 8, 221–300, 2020.

Moes, H.: Optimum similarity analysis with applications to elastohydrodynamic lubrication, Wear, 159, 57–66, 1992.

Mohammadpour, M., Johns-Rahnejat, P., Rahnejat, H., and Gohar, R.: Boundary conditions for elastohydrodynamics of circular point con-

630   tacts, Tribology Letters, 53, 107–118, 2014.

Morales-Espejel, G.: Central film thickness in time-varying normal approach of rolling elastohydrodynamically lubricated contacts, Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 222, 1271–1280, 2008.

Morales-Espejel, G.: Surface roughness effects in elastohydrodynamic lubrication: a review with contributions, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 228, 1217–1242, 2014.

635   Morales-Espejel, G. and Wemekamp, A.: Ertel—Grubin methods in elastohydrodynamic lubrication–a review, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 222, 15–34, 2008.

Panton, R.: Incompressible flow (Chapter 22 - Lubrication approximation), John Wiley & Sons, 2013.

Poll, G., Li, X., Bader, N., and Guo, F.: Starved lubrication in rolling contacts-a review, Bearing World J., 4, 69–81, 2019.

Rolink, A., Schröder, T., Jacobs, G., Bosse, D., Hölzl, J., and Bergmann, P.: Feasibility study for the use of hydrodynamic plain bearings

640   with balancing support characteristics as main bearing in wind turbines, in: Journal of Physics: Conference Series, vol. 1618, p. 052002, IOP Publishing, 2020.

Rolink, A., Jacobs, G., Schröder, T., Keller, D., Jakobs, T., Bosse, D., Lang, J., and Knoll, G.: Methodology for the systematic design of conical plain bearings for use as main bearings in wind turbines, Forschung im Ingenieurwesen, 85, 629–637, 2021.

Spikes, H.: Sixty years of EHL, Lubrication Science, 18, 265–291, 2006.

645   Venner, C.: Higher-Order Multilevel Solvers for the EHL Line and Point Contact Problem, J. Tribol., 4, 741–750, https://doi.org/10.1115/1.2927328, 1994.

Venner, C. and Bos, J.: Effects of lubricant compressibility on the film thickness in EHL line and circular contacts, Wear, 173, 151–165, 1994.

Venner, C. and Lubrecht, A.: Multigrid techniques: A fast and efficient method for the numerical simulation of elastohydrodynamically

650   lubricated point contact problems, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 214, 43–62, https://doi.org/10.1243/1350650001543007, 2000a.

Venner, C. and Lubrecht, A.: Multi-level methods in lubrication, Elsevier, 2000b.

Venner, C. and Wijnant, Y.: Validation of EHL contact predictions under time varying load, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 219, 249–261, 2005.

655   Vergne, P. and Bair, S.: Classical EHL versus quantitative EHL: a perspective part I—real viscosity-pressure dependence and the viscosity-pressure coefficient for predicting film thickness, Tribology Letters, 54, 1–12, 2014.

Wheeler, J., Fillot, N., Vergne, P., Philippon, D., and Morales-Espejel, G.: On the crucial role of ellipticity on elastohydrodynamic film thickness and friction, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 230, 1503–1515, 2016a.

660   Wheeler, J., Vergne, P., Fillot, N., and Philippon, D.: On the relevance of analytical film thickness EHD equations for isothermal point contacts: Qualitative or quantitative predictions?, Friction, 4, 369–379, https://doi.org/10.1007/s40544-016-0133-7, 2016b.

Wymer, D. and Cameron, A.: Elastohydrodynamic lubrication of a line contact, Proceedings of the Institution of Mechanical Engineers, 188, 221–238, 1974.

Zhang, X. and Glovnea, R.: An experimental investigation of grease lubricated EHD contact subjected to normal sinusoidally variable loading, Tribology International, 147, 106 272, 2020.

Zhang, X., Glovnea, R., Morales-Espejel, G., and Félix-Quiñonez, A.: The Effect of Working Parameters upon Elastohydrodynamic Film Thickness Under Periodic Load Variation, Tribology Letters, 68, 1–10, 2020.

665