TRIBOLOGY SERIES, 26

ENGINE TRIBOLOGY

C. M. TAYLOR
(Editor)

ELSEVIER

ENGINE TRIBOLOGY

TRIBOLOGY SERIES

Advisory Board

Vol. 1 Tribology – A Systems Approach to the Science and Technology of Friction, Lubrication and Wear (Czichos)
Vol. 2 Impact Wear of Materials (Engel)
Vol. 3 Tribology of Natural and Artificial Joints (Dumbleton)
Vol. 4 Tribology of Thin Layers (Iliuc)
Vol. 5 Surface Effects in Adhesion, Friction, Wear, and Lubrication (Buckley)
Vol. 6 Friction and Wear of Polymers (Bartenev and Lavrentev)
Vol. 7 Microscopic Aspects of Adhesion and Lubrication (Georges, Editor)
Vol. 8 Industrial Tribology – The Practical Aspects of Friction, Lubrication and Wear (Jones and Scott, Editors)
Vol. 9 Mechanics and Chemistry in Lubrication (Dorinson and Ludema)
Vol. 10 Microstructure and Wear of Materials (Zum Gahr)
Vol. 11 Fluid Film Lubrication – Osborne Reynolds Centenary (Dowson et al., Editors)
Vol. 12 Interface Dynamics (Dowson et al., Editors)
Vol. 13 Tribology of Miniature Systems (Rymuza)
Vol. 14 Tribological Design of Machine Elements (Dowson et al., Editors)
Vol. 15 Encyclopedia of Tribology (Kajdas et al.)
Vol. 16 Tribology of Plastic Materials (Yamaguchi)
Vol. 17 Mechanics of Coatings (Dowson et al., Editors)
Vol. 18 Vehicle Tribology (Dowson et al., Editors)
Vol. 19 Rheology and Elastohydrodynamic Lubrication (Jacobson)
Vol. 20 Materials for Tribology (Glaeser)
Vol. 21 Wear Particles: From the Cradle to the Grave (Dowson et al., Editors)
Vol. 22 Hydrostatic Lubrication (Bassani and Piccigallo)
Vol. 23 Lubricants and Special Fluids (Stepina and Vesely)
Vol. 24 Engineering Tribology (Stachowiak and Batchelor)
Vol. 25 Thin Films in Tribology (Dowson et al., Editors)
Vol. 26 Engine Tribology (Taylor, Editor)

TRIBOLOGY SERIES, 26

ENGINE TRIBOLOGY

Edited by

C.M. Taylor
Institute of Tribology
Department of Mechanical Engineering
The University of Leeds
Leeds, U.K.

ELSEVIER
Amsterdam • London • New York • Tokyo 1993

ELSEVIER SCIENCE PUBLISHERS B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

Library of Congress Cataloging-in-Publication Data

Engine tribology / edited by C.M. Taylor.
p. cm. -- (Tribology series ; 26)
ISBN 0-444-89755-0 (alk. paper)
1. Tribology. 2. Internal combustion engines--Lubrication.
I. Taylor, C. M. (Christopher Malcolm) II. Series.
TJ1075.E55 1993
621.43--dc20 93-25342
CIP

ISBN 0 444 89755 0

This book is printed on acid-free paper

Printed in The Netherlands

PREFACE

Studies of the lubrication of machine elements have been undertaken at the University of Leeds for over a century. Coincidentally the automobile has been part of our lives for a similar period. The ability to analyse in some depth the important frictional components of the internal combustion engine was significantly enhanced by the advent of the digital computer which became generally available around 1960, a few years before the word Tribology was coined.

In the last thirty years lubrication analysts and experimentalists have turned their attention successively to examination of engine bearings, piston assemblies and valve trains. The energy crisis of the mid-1970's stimulated an increased awareness of the scarcity of our natural resources for power generation and the need for their efficient use. Concurrently, there has been a growing recognition of the fragility of our environment and associated pressures through the green lobby to reduce emissions of noxious oxides, particulates and greenhouse gases. Through fuel consumption and exhaust considerations the automobile is a significant factor on a global scale in both these areas.

Customer expectations and international competition are obliging car and commercial vehicle manufacturers to produce more efficient and cleaner products in shorter product cycle times. The consideration of **Engine Tribology** has a leading role to play in helping to achieve these goals. Specific areas of interest include: design influences on fuel economy and emissions; new materials (ceramics, steels, coatings, lubricants, additives); low viscosity lubricants; and low heat rejection (adiabatic)engines.

Research and design in the area of **Engine Tribology** have formed a major thrust of activities in the Institute of Tribology within the Department of Mechanical Engineering at Leeds for over a quarter of a century. This has included the association of two of the series of annual Leeds-Lyon Symposia on Tribology with the general area. In 1982 the topic addressed was 'Tribology of Reciprocating Engines' whilst in 1990 the subject was 'Vehicle Tribology', both meetings being at Leeds. In response to a perceived need the Institute of Tribology mounted a three day continuing professional education course dealing with the subject of **Engine Tribology**. The programme was directed at young professional engineers in the automobile, automotive products, petroleum and

associated industries. The course also proved attractive to those in the academic and research and development environments and a lively meeting was held in September 1991 with delegates from all over Europe. The chapters of this book, updated, comprise the written material presented to attendees.

The structure of the book is such that firstly some basic features of tribology particularly associated with internal combustion engines are dealt with. That is, lubrication analysis relevant to plain bearings, Hertzian contact theory and elastohydrodynamic lubrication associated with cams and followers and friction and wear in a general context. After a general overview of lubrication mechanisms in an automobile engine, successive pairs of chapters examine engine bearings, valve trains, (cam and followers) and piston assemblies. For each machine element a background introduction is followed by design interpretations and a consideration of future developments. Final the important topic of materials, solids and lubricants, is focused upon in the concluding chapters.

The staff of the Institute of Tribology are grateful to those external contributors who have facilitated the publication of this book, and we all join in hoping that it proves to be an interesting and valuable resource in the important study of **Engine Tribology**.

C M Taylor
Institute of Tribology
Department of Mechanical Engineering
University of Leeds
United Kingdom

May 1993

CONTENTS

Engine Tribology / C.M. Taylor (Editor)

Short Journal Bearing LubricationTheory

D A Jones

Department of Mechanical Engineering, The University of Leeds, Leeds LS2 9JT, England

1. INTRODUCTION

This lecture builds from the most basic concepts of viscous flow to develop the governing equation of fluid film lubrication, the 'Reynolds Equation', and continues from that point to generate a model for the operating parameters of the journal bearing specifically the 'Narrow Journal Bearing Theory'.

1.1 Notation

F	Force generally or journal bearing friction force
A	Area
G	Shear Modulus
U, U_1, U_2	Surface velocities in x direction
V_1, V_2	Surface velocities in y direction
C, D	Constants of integration
W	Journal bearing load capacity
$W_{\theta=0}, W_{\theta=\frac{\pi}{2}}$	Journal bearing load capacity components
E	Energy dissipated in the journal bearing (E=Fwr)
C	Specific heat of lubricant
x, y, z	Coordinates
u, v, w	Velocity components (in x, y, z directions)
p	Pressure
h	Film thickness (in z direction)
$\dot{m}_x, \dot{m}_y$	Mass flowrate (in x and y directions)
b.	Bearing width
r	Journal bearing radius
d	Journal bearing diameter
c	Journal bearing radial clearance
e	Journal bearing eccentricity

w	Journal bearing specific loading $\left(w = \frac{W}{bd}\right)$
q	Lubricant volume flowrate in journal bearing
ϕ	Shear strain
τ	Shear stress
$\dot{\phi}$	Rate of strain
θ	Angular coodinate
ε	Journal bearing eccentricity ratio $\left(\varepsilon = \frac{e}{c}\right)$
$\Delta\theta$	Temperature rise in oil expelled from journal bearing
η	Dynamic viscosity
ρ	Fluid density
ω	Angular velocity
ψ	Journal bearing attitude angle
μ	Journal bearing friction coefficient $\left(\mu = \frac{F}{W}\right)$

2 STRAIN AND FLOW

2.1 Elastic Solids

An ideally elastic material obeys Hooke's Law ie. Under stress an equilibrium strain develops, the two being proportional to one another. So for the elastic block of Figure 1 subject to shear:

$$\frac{F}{A} = G\frac{\delta x}{\delta z} = G\phi \text{ (for small } \phi)$$

or $\tau = G\phi$ (1)

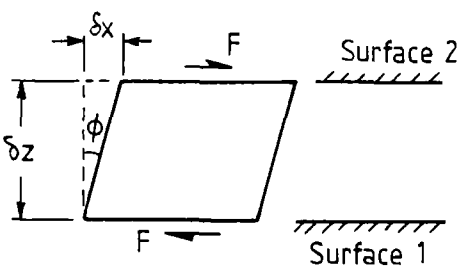


FIGURE 1

The constant G is the 'Shear Modulus' or 'Modulus of Rigidity'. Ideally, too, the elastic strain is fully recovered on removal of the stress. Many real materials behave in a reconisably similar manner.

2.2 Viscous Fluids

A complication arises if the gap between the two surfaces of Figure 1 is filled with a fluid; the strain ϕ increases without limit and the *rate* of straining $\dot{\phi}$ becomes dependant on the imposed stress τ. Again, in the simplest case, there will be proportionality

$$\text{ie. } \tau = \eta\dot{\phi} = \eta\frac{d\dot{x}}{dy} = \eta\frac{du}{dy} \qquad (2)$$

The constant η is the 'Dynamic Viscosity' of the fluid and a fluid obeying the law is a 'Newtonian Fluid'. Many real fluids behave similarly though more complex relationships are common.

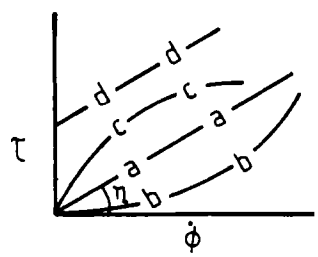


FIGURE 2

The line aa of Figure 2 represents a Newtonian Fluid, bb and cc Non-Newtonian Fluids and dd an ideal plastic material.

3. EQUATION OF MOTION

Figure 3 identifies the forces acting on a fluid element of area $\delta x \delta z$ moving in a direction x with a velocity u due to a pressure p=p(x) and a shear stress $\tau=\tau(z)$. For the moment pressure and shear stresses will be taken to be invariant in the y direction.

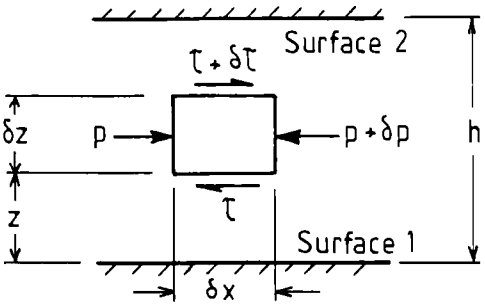


FIGURE 3

Note that pressure variations in the z direction will be negligible for small values of h since the medium is a fluid and that shear stress gradients in the x direction will be negligible compared with those in the z direction provided that the taper $\frac{dh}{dx}$ is small.

For steady motion in the x direction, per unit width (in the y direction)

$$p\delta z - (p + \delta p)\,\delta z - \tau\delta x + (\tau + \delta t)\,\delta x = 0$$

Or, in the limit $\frac{dp}{dx} = \frac{d\tau}{dz}$ (3)

4. VISCOUS FLOW SPECIES

Substituting from Equation (2) into Equation (3) and ignoring potential viscosity variations in the z direction

$$\frac{dp}{dx} = \tau\frac{d^2u}{dz^2} \tag{4}$$

Since $\frac{dp}{dx}$ is taken to be invariant in z, Equation (4) may be integrated when, introducing suitable boundary conditions, ($u=U_1$, when $z=0$ and $u=U_2$ when $z=h$) yields:

$$u = U_1 + (U_2 - U_1)\frac{z}{h} - \frac{h^2}{2\eta}\left(\frac{z}{h}\right)\left(1 - \frac{z}{h}\right)\frac{dp}{dx} \tag{5}$$

There are two components of velocity in Equation 5, a linear, velocity induced or 'Couette' component and a parabolic, pressure induced 'Poiseiuille' component as shown in Figure 4.

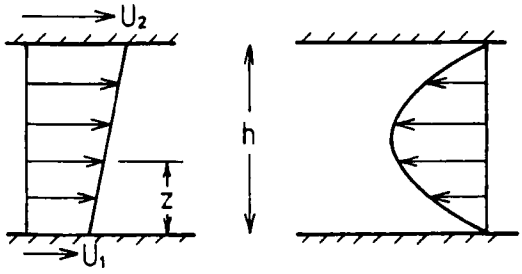


FIGURE 4

5. THE SLOT FLOW EQUATION

Equation (5) may be used to calculate the mass flowrate in the x direction through a narrow slot of width b (in the y direction).

$$\dot{m}_x = \int_o^h b\rho u dz \qquad (6)$$

Or, substituting in Equation (6) from Equation (5), ignoring density variations in y and integrating:

$$\dot{m}_x = \rho b \left\{ \left(\frac{U_1 + U_2}{2} \right) h - \frac{h^3}{12\eta} \frac{dp}{dx} \right\} \qquad (7)$$

Equation (7) is the well known 'Slot Flow Equation', valuable in itself, but here developed mainly as a step towards the famous 'Reynolds Equation'.

6. CONTINUITY AND THE REYNOLDS' EQUATION

6.1 One-Dimensional Flow

Equation (5) defines the components of the velocity u taking place in the x direction. It permits u to vary in y but ignores additional possible velocity components v and w, ie. it represents 'one dimensional' flow as does the derivative Equation (7). In the absence of other flow constituents the continuity condition requires that $\dot{m}_x$ = Const.

$$\text{Or} \quad \frac{d\dot{m}_x}{dx} = 0 \tag{8}$$

$$\text{ie.} \quad \frac{d}{dx}\left\{\frac{\rho h^3}{12\eta}\frac{dp}{dx}\right\} = \frac{d}{dx}\left\{\rho\frac{(U_1 + U_2)h}{2}\right\} \tag{9}$$

Equation (9) then is a one dimensional version of the Reynolds' Equation for viscous flow in a thin film.

6.2 Two-Dimensional Flow

Continuity will be achieved in two dimensional flow if:

$$\frac{\partial\dot{m}_x}{\partial x} + \frac{\partial\dot{m}_y}{\partial x} = 0 \tag{10}$$

In this $\dot{m}_x$ will be given by Equation (11) an obvious derivative of Equation (7), and $\dot{m}_y$ by the similar Equation (12).

$$\frac{\dot{m}_x}{b} = \rho\left\{\frac{(U_1 + U_2)h}{2} - \frac{h^3}{12\eta}\frac{\partial p}{\partial x}\right\} \tag{11}$$

$$\frac{\dot{m}_y}{b} = \rho\left\{\frac{(V_1 + V_2)h}{2} - \frac{h^3}{12\eta}\frac{\partial p}{\partial y}\right\} \tag{12}$$

From Equations (10) to (12) then, Reynolds' Equation in two dimensional flow will be given by:

$$\frac{\partial}{\partial x}\left\{\frac{\rho h^3}{12\eta}\frac{\partial p}{\partial x}\right\} + \frac{\partial}{\partial y}\left\{\frac{\rho h^3}{12\eta}\frac{\partial p}{\partial y}\right\} = \frac{\partial}{\partial x}\left\{\rho\frac{(U_1 + U_2)h}{2}\right\} \tag{13}$$

ie. Equation (13) represents a more complete form of the Reynolds Equation. It will be limited only by the condition that V_1=V_2=0 which will be seen to be realistic given the journal bearing configuration of Figure 5.

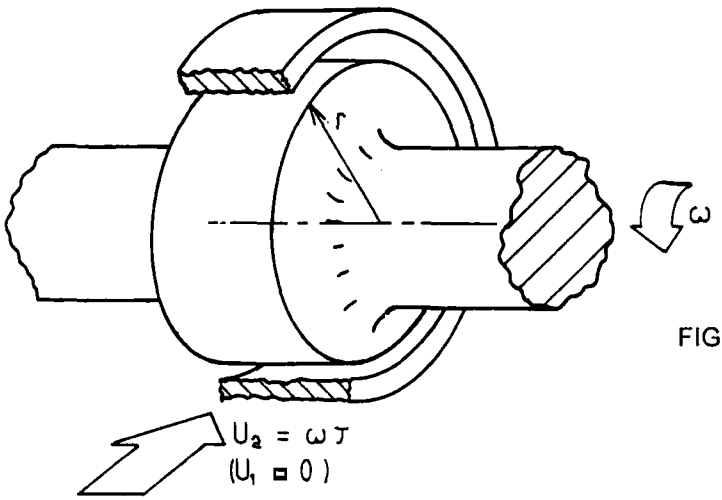


FIGURE 5

6.3 Terminology

In the foregoing we have discussed 'one' and 'two' dimensional *flows* and our terminology is that commonly employed in the general field of fluid dynamics. Authors in the more limited field of fluid film lubrication frequently refer to these as, respectively, 'two' and 'three' dimensional *films*.

6.4 Analytical Solutions of Reynolds' Equations

The solution of Equation (13) involves integration which may proceed in principle for the case where h=h(x) either where the second term is negligible by comparison with the first or vice-versa.

The first approach implies negligible gradients in the y direciton or a rather broad film. this is known as a 'wide bearing' solution and was first attempted by Osborne Reynolds (1886) and later solved more completely by Sommerfeld (1904).

The alternative assumes that the first term is relatively small so that pressure induced flow in the x direction is small, as would be the case for a 'narrow bearing', an approach favoured by Michell (1905) and by Dubois and Ocvirk (1955).

Typically the plain journal bearings associated with reciprocating internal combustion engines tend to be narrow in relation to their diameters and we shall concentrate on the narrow journal bearing analysis.

7. THE NARROW JOURNAL BEARING

7.1 The Pressures Generated

Working from Equation (13) and ignoring pressure induced flow in the x direction we see that:

$$\frac{\partial}{\partial y}\left\{\frac{\rho h^3}{12\eta}\frac{\partial p}{\partial y}\right\} = \frac{\partial}{\partial x}\left\{\frac{(U_1 + U_2 h)}{2}\right\}$$

Or, with $U_1=0$ and $U_2 = U$, for constant density and viscosity:

$$\frac{d}{dy}\left\{h^3 \frac{dp}{dy}\right\} = 6\eta U \frac{dh}{dx} \qquad (14)$$

We may integrate equation (14) for a correctly aligned bearing (for which h=h(x) only) to give

$$h^3 \frac{dp}{dy} = 6\eta U y \frac{dh}{dx} + C$$

The constant of integration C will be zero since $\frac{dp}{dy} = 0$ where $y = 0$ by symmetry. (ie. On the centreline of the bearing).

$$\text{Thus } p = \frac{3\eta U y^2}{h^3}\frac{dh}{dx} + D$$

Since p=0, or ambient, at the edges of the bearing where $y = \pm \frac{b}{2}$

$$\therefore p = -\frac{3\eta U}{h^3}\left\{\frac{b^2}{4} - y^2\right\}\frac{dh}{dx} \qquad (15)$$

Equation (15) demonstrates that positive, or super-ambient pressures will be developed where a film is contracting.

7.2 The Film Shape

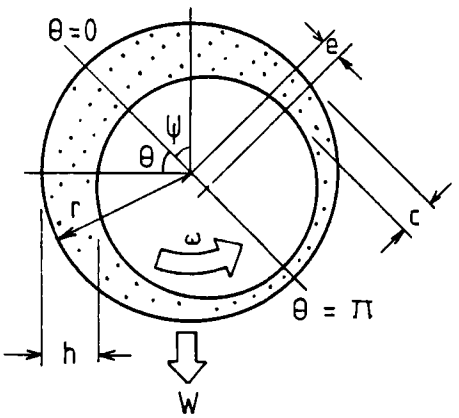


FIGURE 6

Figure 6 grossly exaggerates the dimensions of the fluid film to make it clear that the film thickness is given by:

$h = c + e\cos\theta$

Or, with $\varepsilon = \frac{e}{c}$ (known as the 'eccentricity ratio')

$$h = c(1 + \varepsilon\cos\theta) \tag{16}$$

To put the diagram into perspective; the shaft is to assume an eccentric position relative to the bush defined by the 'eccentricity' e and the 'attitude angle' ψ when loaded with W as shown.

It will be seen to be convenient to measure the x of Equation (15) from the point of maximum film thickness, where θ (of Equation (16)) is zero.

ie. $x = r\theta$ (17)

Hence $$\frac{dh}{dx} = \frac{dh}{d\theta} \cdot \frac{d\theta}{dx} = -\frac{c\varepsilon}{r}\sin\theta \tag{18}$$

and, substituting from Equations (16) and (18) in (15);

$$p = \frac{3\eta U}{rc^2} \cdot \frac{\varepsilon\sin\theta}{(1 + \varepsilon\cos\theta)^3} \cdot \left\{\frac{b^2}{4} - y^2\right\} \tag{19}$$

7.3 The Load Capacity and Attitude Angle

The bearing's load capacity may then be estimated from an integration routine. Dubois and Ocvirk (1953), observing that a liquid film is unable to sustain significant sub-ambient pressures because of cavitation, considered only the convergent part of the film. Two separate components of load may be evaluated:

$$\underset{\theta=0}{W} = \int_0^{\pi}\int_{-\frac{b}{2}}^{+\frac{b}{2}} p.r\cos\theta.dyd\theta$$

and $$\underset{\theta=\frac{\pi}{2}}{W} = \int_0^{\pi}\int_{-\frac{b}{2}}^{+\frac{b}{2}} p.r\sin\theta.dyd\theta$$

In these forms $\underset{\theta=0}{W}$ is the component of load in the direction $\theta=0$ with $\underset{\theta=\frac{\pi}{2}}{W}$ as the perpendicular component. Substitution for pressure from Equation (19) yields three integrals:

$$\int_{-\frac{b}{2}}^{+\frac{b}{2}} \left\{\frac{b^2}{4} - y^2\right\} dy = \frac{b^3}{6}$$

$$\int_0^{\pi} \frac{\sin\theta\cos\theta\, d\theta}{(1+\varepsilon\cos\theta)^3} = \frac{-2\varepsilon}{(1-\varepsilon^2)^2}$$

and $$\int_0^{\pi} \frac{\sin^2\theta\, d\theta}{(1+\varepsilon\cos\theta)^3} = \frac{\pi}{2}\cdot\frac{1}{(1-\varepsilon^2)^{\frac{3}{2}}}$$

Integrals of the type $\int \frac{\sin^l\theta.\cos^m\theta}{(1+\varepsilon\cos\theta)^n}.d\theta$ are best evaluated using the method of Booker (1965).

It follows that $\underset{\theta=0}{W} = -\frac{\eta U b^2}{c^2}\cdot\frac{\varepsilon^2}{(1-\varepsilon^2)^2}$

and $\underset{\theta=\frac{\pi}{2}}{W} = \frac{\pi}{4}\frac{\eta U b^3}{c^2}\cdot\frac{\varepsilon}{(1-\varepsilon^2)^{\frac{3}{2}}}$

Then with, $W^2 = \underset{\theta=0}{W^2} + \underset{\theta=\frac{\pi}{2}}{W^2}$

we see that the bearing load capacity is

$$W = \frac{\pi}{4} \cdot \frac{\eta U b^3}{c^2} \cdot \frac{\varepsilon}{\left(1-\varepsilon^2\right)^2} \left[1 + \left\{\left(\frac{4}{\pi}\right)^2 - 1\right\}\varepsilon^2\right]^{\frac{1}{2}} \qquad (20)$$

It is often convenient to consider a 'specific bearing load' w with $w = \frac{W}{bd}$.

$$\left(\frac{w}{\eta\omega}\right)\left(\frac{c}{r}\right)^2 \cdot \left(\frac{d}{b}\right)^2 = \frac{\pi}{2} \cdot \frac{\varepsilon}{\left(1-\varepsilon^2\right)^2} \left[1 + 0.621\,\varepsilon^2\right]^{\frac{1}{2}} \qquad (21)$$

Barwell (1979) discusses several alternative forms of Equation (21) for various purposes. Figure 7 shows how the dimensionless load capacity varies with the eccentricity ratio.

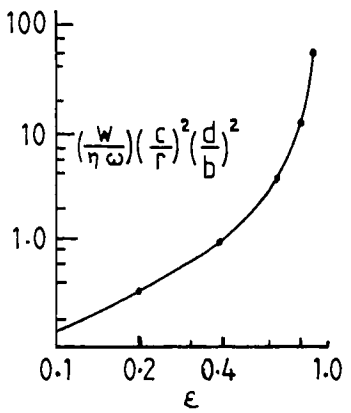


FIGURE 7

The bearing's 'attitude angle' Ψ is also available since

$$\tan\psi = \left\{\frac{W_{\theta=\frac{\pi}{2}}}{W_{\theta=0}}\right\}$$

whence $$\psi = \tan^{-1}\left\{\frac{-\pi}{4} \cdot \frac{\left(1-\varepsilon^2\right)^{\frac{1}{2}}}{\varepsilon}\right\} \qquad (22)$$

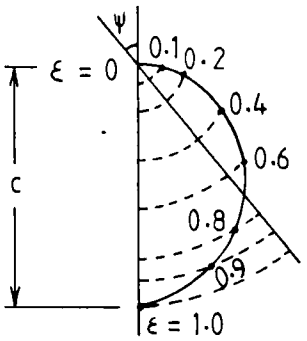


FIGURE 8

Figure 8 shows the trajectory of the shaft within the radial clearance for various eccentricity ratios. It implies that the shaft moves perpendicular to the applied load for the lightest loads and in the direction of the load for the highest values of load.

7.4 Bearing Friction and Friction Coefficient

Where there is no significant pressure induced flow in the direction of surface motion we see that

$$\tau = \eta\frac{U}{h} = \frac{\eta\omega r}{c(1+\varepsilon\cos\theta)}$$

If, following Dubois and Ocvirk (1953), then, we assume the clearance space to be filled with fluid, the friction force may be estimated by

$$F = \int_0^{2\pi}\int_{-\frac{b}{2}}^{+\frac{b}{2}}\tau r.\,dyd\theta$$

These yield $$F = \frac{\eta\omega br^2}{c}\cdot\frac{2\pi}{(1-\varepsilon^2)^{\frac{1}{2}}} \qquad (23)$$

The friction coefficient is

$$\mu = \frac{F}{W}$$

Whence $$\mu = 2\left(\frac{c}{r}\right)\left(\frac{d}{b}\right)^2\cdot\frac{1}{\varepsilon}\,\frac{(1-\varepsilon)^{\frac{3}{2}}}{(1+0.621\varepsilon^2)^{\frac{1}{2}}} \qquad (24)$$

7.5 Energy Dissipation and Lubricant Temperature

The energy expended within the bearing is given by $E=FU=F\omega r$

$$\text{ie. } E = \frac{\eta\omega^2 br^3}{c} \cdot \frac{2\pi}{\left(1-\varepsilon^2\right)^{\frac{1}{2}}} \tag{25}$$

Considering now continuity within the diverging part of the film we see that, with q as the volume flowrate required to make good the difference between the flow crossing the minimum and maximum film thickness sections:

$$q = \frac{bU\left(h_{max} - h_{min}\right)}{2} = b\omega rc\varepsilon \tag{26}$$

Clearly q also represents the flow expelled from the sides of the bearing in the convergent film.

Now with $\Delta\theta$ as the mean temperature rise in the expelled fluid (and ignoring any heat conducted away from the bearing);

$$E = \rho qC\Delta\theta = \rho b\omega rc\varepsilon C\Delta\theta \tag{27}$$

Or, equating (26) and (27),

$$\Delta\theta = 2\pi \cdot \frac{\eta\omega}{\rho C}\left(\frac{r}{c}\right)^2 \cdot \frac{1}{\varepsilon\left(1-\varepsilon^2\right)^{\frac{1}{2}}} \tag{28}$$

The mean temperature rise in the fluid expelled from a bearing has been found to provide a reliable means of estimating the effective temperature of the fluid within the film, hence of its viscosity and ultimately of viscous friction in the bearing.

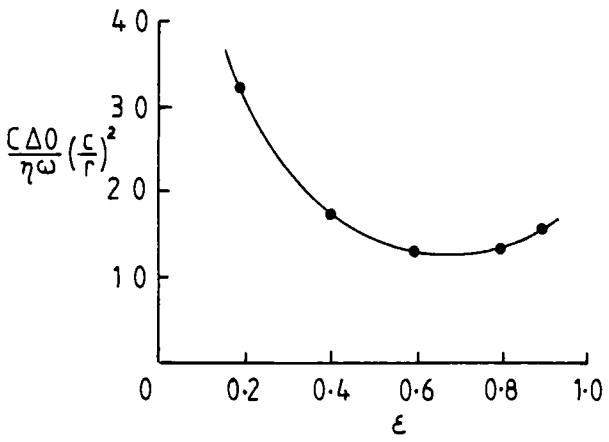


FIGURE 9

Figure 9 shows the variation of dimensionless temperature rise with eccentricity ratio. The minimum $\left(\text{at } \varepsilon = \frac{1}{\sqrt{2}}\right)$ is of greater theoretical than practical interest.

REFERENCES

Barwell F.T. (1979) 'Bearing Systems. Principles and Practice.' Oxford U. Press.

Booker J.F. (1965) 'A table of the journal-bearing integral.' Trans. A.S.M.E. J. Basic Engg. 87, 533-5.

Dubois G.B. and Ocvirk F.W. (1953) 'Analytical derivation and experimental evaluation of short-bearing approximation for full journal bearings.' N.A.C.A. Rep. No. 1157.

Michell A.G.N. (1905). 'Die Schmiering ebener Flachen.' Z. fur Math. Phys. 52 123-38.

Reynolds O. (1886). 'On the theory of lubrication and its application to Mr Beauchamp Tower's experiments, including an experimental determination of the viscosity of olive oil.' Phil. Trans. R. Soc. 177 (1), 157-234.

Sommerfield A. (1904). 'Zur hydrodynamischen theorie der Schmiermittelreibung.' Z. fur Math. Phys. 50 97-155.

Engine Tribology / C.M. Taylor (Editor)

Elastohydrodynamic Lubrication Theory

D A Jones

Department of Mechanical Engineering, The University of Leeds, Leeds LS2 9JT, England

1. GENERAL OBSERVATIONS

The load capacity of the conventional, liquid-lubricated, journal bearing benefits from the close conformity between the two surfaces. The thin film generated can be of considerable extent in that configuration so that the mean pressures generated within the film can be modest. Common forms of thrust bearing have films of similar proportions. The same cannot be said of gear tooth engagements, cams and followers, nor of the contact between rolling elements and raceways in rolling contact bearings. In these the contact areas are small and high interfacial pressures must be generated with or without a liquid lubricant. The evidence is that such arrangements frequently operate without significant metallic contact despite pressures of the order of 10^4 atmospheres. These are said to be 'elastohydrodynamically' lubricated in that the shape of the high modulus elastic boundaries to the conjunction become modified by virtue of the hydrodynamic pressures generated.

At first it may seem surprising that a fluid film can sustain such high pressures. However any film generated will be very thin, there will be significant elastic deformation of the surroundings and the viscosity of a mineral oil lubricant is strongly dependant on pressure, (there is an approximately exponential relationship).

Two aspects of the elastohydrodynamic theory require attention; the *extent* of the fluid film and its *thickness*. The latter will be particularly important in that it will define the surface finish conditions required to avoid metallic contact. The former may be defined approximately from a consideration of only the elastic nature of the conjunction, the 'Hertzian Contact Theory'. In principle the latter is a more difficult problem since it will require a simultaneous solution of elastic and hydrodynamic equations, the 'E H L Theory'. In practice useful results can be obtained in E H L by approximate methods, the 'Grubin Theory'.

It would be impossible to do justice to the research effort expended in resolving these issues in the short time available, and we shall only outline the background. Our main objective will be to introduce the results of the work in an utilitarian form to meet the needs of designers.

The author thanks Professor Chris Taylor who created the large Appendix to this presentation.

1.1 Notation - Hertzian Contact of Spheres

- a Radius of Hertzian contact.
- A,B Contacting spheres
- $e_{A,B}$ Surface elastic displacements of A,B.
- $E_{A,B}$ Young's modulus of elasticity for A,B.
- E' Equivalent elastic modulus $\left\{\frac{1}{E'}=\frac{(1-\nu_A)^2}{E_A}+\frac{\left(1-\nu_B{}^2\right)}{E_B}\right\}$
- f A geometrical feature of the Hertzian contact.
- F Total load on the contact
- $g_{A,B}$ Geometrical features of A,B
- p Hertzian pressure (Sometimes with "max" and "mean" subsripts)
- r A radial co-ordinate
- $r_{A,B}$ Radii of A,B
- R 'Equivalent' radius $\frac{1}{R}=\frac{1}{r_A}+\frac{1}{r_B}$
- z A linear coordinate
- δ The approach of remote points in the spheres (sometimes with subsripts A,B)
- $n_{A,B}$ Poisson's ratio for A,B

1.2 Notation - Elastohydrodynamic Lubrication of Cylinders

Dimensional Qualities

- A,B Contacting cylinders
- e Surface elastic displacement of equivalent cylinder
- $e_{A,B}$ Surface elastic displacements of A.B
- $E_{A,B}$ Young's modulus of elasticity for A,B
- E' Equivalent elastic modulus $\frac{1}{E'}=\frac{1}{2}\left\{\frac{\left(1-\nu_A{}^2\right)}{E_A}+\frac{\left(1-\nu_B{}^2\right)}{E_B}\right\}$
- F' Load per unit contact length
- g Geometrical feature of equivalent cylinder
- $g_{A,B}$ Geometrical features of A,B
- h Film thickness
- h_s Specific film thickness (defined in context)
- h_m Film thickness where $\frac{dp}{dx}=0$
- h_{cen} Centreline film thickness
- H_{cen} Dimensionless centreline film thickness

H_{min} Dimensionless minimum film thickness
p Pressure

q Reduced pressure $\left[q=\frac{1}{\alpha}\left(1-e^{-\alpha p}\right)\right]$

$r_{A,B}$ Radii of cylinders A,B

R Equivalent radius of cylinder $\left\{\frac{1}{R}=\frac{1}{r_A}+\frac{1}{r_B}\right\}$

s Linear coordinate

$U_{A,B}$ Surface velocity of A,B

U Entraining velocity $\left[U=\frac{1}{2}\left(U_A+U_B\right)\right]$

x Linear coordinate

α Pressure coefficient of viscosity

η Dynamic viscosity at operating conditions

η_o Dynamic viscosity at atmospheric pressure and operating temperature

ν Poission's ratio

ρ Fluid density

Dimensionless Quantities

G Materials parameter ($G=\alpha E'$)

H_{cen} Centreline film thickness $\left(H_{cen}=\frac{h_{cen}}{R}\right)$

H_{min} Minimum film thickness $\left(H_{min}=\frac{h_{min}}{R}\right)$

U_e Speed parameter $\left(U_e=\frac{\eta_o U}{E'R}\right)$

W'_e Load parameter $\left(W'_e=\frac{F'}{E'R}\right)$

2. HERTZIAN CONTACT THEORY

2.1 Introduction

Hertz (1886) developed a theory describing the properties of the small area of intimate contact generated when two large elastic bodies are pressed together. In particular, he dealt with the contact between relatively massive ellipsoidal bodies. By way of example we shall consider the special case of two spheres in contact. The results from Hertz' general analysis are summarised in the Appendix.

2.2 Spheres in Contact

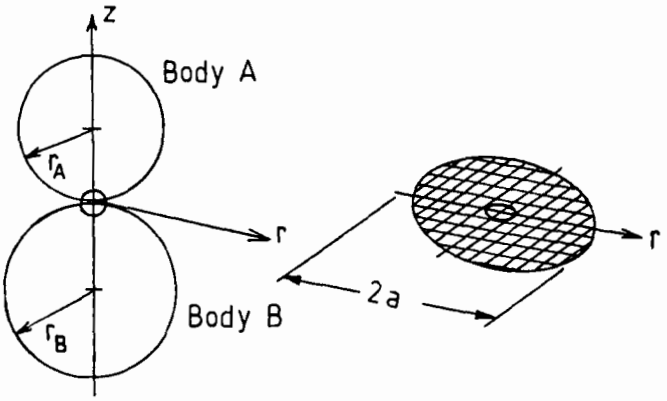


FIGURE 1

Figure 1 shows the model situation; two elastic spheres of radii r_A and r_B are in contact at 0 where a shallow and slightly dished circular interface of radius a develops due to elastic deformation under the influence of the applied load F.

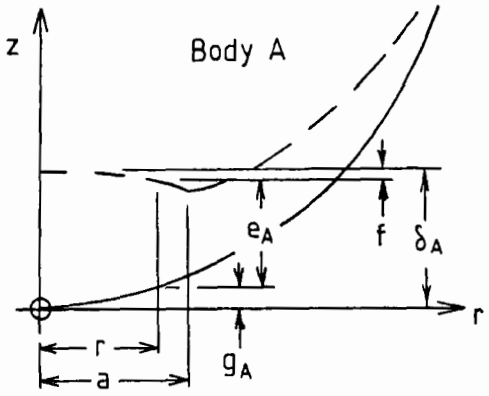


FIGURE 2

In Figure 2 the undeformed shape of the body A is shown in full line, the deformed shape in dashed line. The centreline approach of A to B is δ_A (and this is equal to the approach of a point far removed from the contact). With g_A and f as geometrical features of the conjunction at the radius r and e_A as the elastic deformation of the surface at r;

$$\delta_A = g_A + e_A + f$$

Clearly $\delta_B = g_B + e_B - f$

and $$\delta = \delta_A + \delta_B = g_A + g_B + e_A + e_B \qquad (1)$$

It is easy to see that

$$g_A + g_B = \frac{r^2}{2}\left\{\frac{1}{r_A}+\frac{1}{r_B}\right\} = \frac{r^2}{2R} \tag{2}$$

ie. we have defined an 'equivalent radius' R for the conjunction such that;

$$\frac{1}{R} = \frac{1}{r_A}+\frac{1}{r_B} \tag{3}$$

(Note that, for the more general problem of elliptical surfaces in contact, a slightly different definition applies. This is so for the Appendix). In principle the displacements e_A and e_B are calculable from elasticity theory given a pressure distribution p=p(r).

Hertz' assumed distribution was:

$$p = p_{max}\left\{1-\left(\frac{r}{a}\right)^2\right\}^{\frac{1}{2}} \tag{4}$$

Following which the displacements are:

$$e_{AB} = \frac{\pi}{4}\frac{\left(1-\nu^2_{AB}\right)}{E_{AB}}\frac{p_{max}}{a}\left(2a^2-r^2\right) \tag{5}$$

(See, for example, Johnson (1985).

or, with

$$E'_{AB} = \frac{E_{AB}}{\left(1-\nu^2_{AB}\right)}$$

$$e_A + e_B = \frac{\pi}{4}\frac{p_{max}}{a}\left(2a^2-r^2\right)\left\{\frac{1}{E'_A}+\frac{1}{E'_B}\right\}$$

(Note that in the Appendix, for historical reasons $\frac{1}{E'} = \frac{1}{2}\left\{\frac{1}{E'_A}+\frac{1}{E'_B}\right\}$

$$e_A + e_B = \frac{\pi}{4}\frac{p_{max}}{E'a}\left(2a^2-r^2\right) \tag{6}$$

$$\therefore \delta = \frac{r^2}{2R}+\frac{\pi}{4}\frac{p_{max}}{E'a}\left(2a^2-r^2\right) \tag{7}$$

Clearly then:

$$\delta = \frac{\pi}{2}\frac{p_{max}a}{E'} \tag{8}$$

and $$a = \frac{\pi}{2}\frac{p_{max}R}{E'} \tag{9}$$

If the load applied to the conjunction is F, it will be seen that

$$F = \int_0^a p(r).2\pi r.dr = \frac{2}{3}p_{max}\pi a^2 \tag{10}$$

Thus $$p_{max} = \frac{3}{2}.\frac{F}{\pi a^2} = \frac{3}{2}p_{mean} \tag{11}$$

$$a = \left(\frac{3}{4}\frac{FR}{E'}\right)^{\frac{1}{3}} \tag{12}$$

$$\delta = \left(\frac{9}{16}\frac{F^2}{(E')^2 R}\right)^{\frac{1}{3}} \tag{13}$$

The Equations (4) and (11) - (13) then describe the significant characteristics of the Hertzian contact between two spheres. They were developed by way of example only and are simplified versions of equations given in the Appendix.

3. SOLUTIONS TO ELASTOHYDRODYNAMIC PROBLEMS

3.1 Introduction

Grubin (1949) developed an approximate analytical solution to the problem of the elastohydrodynamic lubrication of rolling cylinders (the 'line contact' problem) long before the first numerical solutions became available. His equation for centreline film thickness in the conjunction is:

$$H_{cen} = 1.95\frac{(GU_e)^{\frac{8}{11}}}{(W'_e)^{\frac{1}{11}}} \tag{14}$$

This equation is very similar to the latest computer generated form (Dowson and Higginson (1959)) quoted in the Appendix:

$$H_{cen} = 3.06\frac{G^{0.56}.U_e^{0.69}}{(W'_e)^{0.10}} \tag{15}$$

Grubin made the assumption that the shape of the elastic surfaces at inlet to the fluid film was the same whether the conjunction was lubricated or not.

Equation (14) demonstrated an important feature of the EHL film; the film thickness is very insensitive to the load applied. This early prediction was remarkably accurate. For example, taking suitable values ($W'_e = 3 \times 10^{-5}$, $G = 5000$ and $U_e = 10^{-11}$) we find centreline film thickness of 2.46×10^{-5} (Grubin) and of 2.63×10^{-5} (Dowson and Higginson), a remarkably close agreement. Although brilliant conceptually the Grubin approach was incapable of revealing what proved to be very important detail.

Early numerical solutions were developed by Petrusevich (1951) and by Dowson and Higginson (1959) for the line contact problem.

Archard and Kirk (1961) produced an analytical solution to the 'point contact' problem (representing, for example, a ball or a plane) using a method similar to Grubin. Their equation for the centreline film thickness was:

$$H_{con} = \frac{1.4(GU_e)^{0.74}}{W_e^{0.073}} \qquad (16)$$

The Equation (16) should be compared with the latest (Hamrock and Dowson (1977)) result:

$$H_{con} = 3.05 \frac{G^{0.49} . U_e^{0.68}}{W_e^{0.073}} \qquad (17)$$

Equation (17) is actually a simplified version of Hamrock and Dowson's more general 'elliptical contact' expression given in the Appendix and may readily be derived therefrom.

The Equations (16) and (17) again yield remarkably similar estimates of the centreline film thickness, though not quite so good as for line contact and again the analytical approach to the problem fails on interesting detail.

The problem of the elliptical contact is particularly important in that it is representative of the conditions in a conventional 'deep groove' ball bearing. This was attacked numerically by Chang (1970) and by Hamrock and Dowson (1977). The Appendix also gives the most recent version of the Hamrock and Dowson minimum film thickness equation.

3.2 Numerical Solutions of the Line Contact Problem

By way of example we shall outline the processes involved in resolving a relatively simple E H L problem. Specifically the problem concerns the fluid film separating two parallel axis elastic cylinders rolling together. The problem would relate to a cylindrical roller in a rolling contact bearing or to a gear tooth.

In the simplest approach, thermal effects in the film, possible transitions to plastic flow in the solids and the effects of finite length in the cylinders is neglected as is any dynamic loading of the conjunction. It will be reasonable to ignore the effects of compressibility in the lubricant but impossible to avoid modelling the large variations of fluid viscosity with pressure.

With these limitations the pressures generated will depend on film shape and the lubricant viscosity whilst the film shape and viscosity will, in turn, depend on the pressures developed. A complication will arise due to the effects of cavitation in the fluid film, but viscosity variations may be represented within a modified version of the equation representing flow in the fluid film, the Reynolds' equation. In practise then it will be necessary to solve, simultaneously, an elasticity equation defining film shape, a modified version of the Reynolds' equation, and an algorithm defining the onset of cavitation. We shall look at each element in turn.

3.2.1 The Theoretical Model

Figure 3 represents a pair of elastic cylindrical rollers of radii r_A and r_B, with an 'equivalent' single cylinder of radius R near a plane.

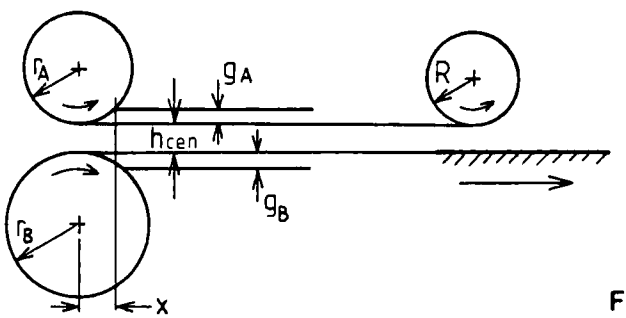


FIGURE 3

It will be seen that, as in paragraph 2.2, the film shape will be defined, for the pair of rollers, by:

$$h = h_{cen} + g_A + g_B + e_A + e_B \qquad (18)$$

And, for the single equivalent roller by:

$$h = h_{cen} + g + e \qquad (19)$$

In this form, again as in 2.2, the convenient simplification to a single cylinder near a plane will be achieved if:

$$g = \frac{x^2}{2R} \qquad (20)$$

$$\frac{1}{R}=\frac{1}{r_A}+\frac{1}{r_B} \tag{21}$$

and $$\frac{1}{E'}=\frac{1}{2}\left(\frac{1}{E'_A}+\frac{1}{E'_B}\right) \tag{22}$$

3.2.2 Surface Displacements

A concentrated line load applied to the boundary of a semi-infinite elastic body in elasticity theory yields surface displacements which go to infinity under the load line according to the expression:

$$e=\frac{2F'}{\pi E'}\ell n(s)+\text{Const}$$

This is illustrated in Figure 4.

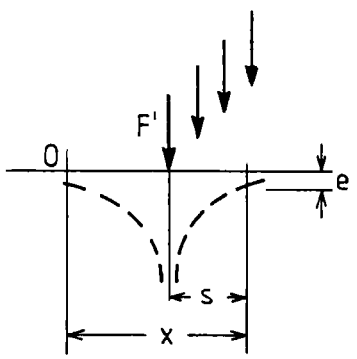


FIGURE 4

Clearly direct numerical integration for the displacements due to an arbitrary pressure profile p=p(x) is impractical. However, provided that the pressure profile can be expressed in a form permitting formal integration, the displacements remain finite everywhere. Dowson and Higginson (1959) made use of the latter feature, they approximated their pressure profiles by a limited number of terms in a polynomial. (The reference gives details of the process).

3.2.3 A Modified Reynolds' Equation

The Reynolds' equation for one-dimensional flow is:

$$\frac{d}{dx}\left\{\frac{\rho h^3}{12\eta}\frac{dp}{dx}\right\}=\frac{d}{dx}\left\{\rho\frac{(U_A+U_B)h}{2}\right\}$$

Whence, for constant density, and writing $U=\frac{(U_A+U_B)}{2}$

$$\frac{d}{dx}\left\{\frac{h^3}{\eta}\frac{dp}{dx}\right\} = 12U\frac{dh}{dx} \qquad (23)$$

Integrating:

$$\frac{dp}{dx} = 12\eta U\left\{\frac{h - Const}{h^3}\right\}$$

The constant may be evaluated if $h=h_m$ where $\frac{dp}{dx} = 0$ so that

$$\frac{dp}{dx} = 12\eta U\left\{\frac{h - h_m}{h^3}\right\} \qquad (24)$$

Equation (23) is an integrated form of the Reynolds' equation and may, in principle, be solved where the relationships $h=h(x,p)$ and $\eta=\eta(p)$ are known.

There are a number of empirical equations representing the variation $\eta=\eta(p)$. One such is

$$\eta = \eta_O e^{\alpha p} \qquad (25)$$

In this α is a pressure exponent of viscosity and η_o is the viscosity at ambient pressure and operating temperature.

Substituting from (24) in (23):

$$e^{-\alpha p}\frac{dp}{dx} = 12\eta_0 U\left\{\frac{h - h_m}{h^3}\right\}$$

Or writing $\alpha q = 1 - e^{-\alpha p}$ (26)

We see that $\frac{dq}{dx} = e^{-\alpha p}\frac{dp}{dx}$

So that q may be regarded as a 'reduced pressure' being the pressure which would be developed in a fluid of constant viscosity η_o. Then the integrated form of Reynolds' equation in terms of reduced pressure becomes:

$$\frac{dq}{dx} = 12\eta_0 U\left\{\frac{h - h_m}{h^3}\right\} \qquad (27)$$

Equation (27) is the modified form of Reynolds' equation referred to earlier and is to be solved simultaneously with the equation (19). (The elastic

displacement e in equation (19) is to be derived by the process described in paragraph 3.2.2, the quantity g by equation (20)).

In practise, the direct approach to the resolution of the problem, by which an assumed film shape would be used to calculate a pressure profile, which would then be used to develop a revised film shape iteratively, does not always converge so that equation (27) proves to have limited usefulness. (It is restricted mainly to the inlet film region). The method of solution preferred by Dowson and Higginson used an assumed pressure profile to generate two alternative film shapes, one resulting from the elasticity and one from an "inverse" hydrodynamic calculation, a method which they found to be rapidly convergent.

3.2.4 The Inverse Hydrodynamic Calculation

Equation (23) may be rewritten in terms of a reduced pressure as:

$$\frac{d}{dx}\left(h^3\frac{dq}{dx}\right) = 12\eta_0 U\frac{dh}{dx} \qquad (28)$$

Which, differentiating, gives:

$$h^3\frac{d^2q}{dx^2} - \frac{dh}{dx}\left\{12\eta_0 U - 3h^2\frac{dq}{dx}\right\} = 0 \qquad (29)$$

There will be two points of inflexion in the reduced pressure curve, for which

$$\frac{d^2q}{dx^2} = 0$$

At one of these points $\frac{dh}{dx} = 0$ and the film thickness is at a minimum. At the other $\frac{dq}{dx} = \frac{4\eta_0 U}{h_a^2}$ and the special value h_a can be deduced as $h_a = \left(\frac{4\eta_0 U}{\frac{dq}{dx}}\right)^{\frac{1}{2}}$

Substituting for $\frac{dq}{dx}$ in Equation (27) yields the value of h_m as

$$h_m = \frac{2}{3}h_a \qquad (30)$$

Rearranging Equation (27) yields a cubic equation in $\frac{h}{h_m}$ reading:

$$\left(\frac{h_m^2}{12\eta_0 U}\frac{dq}{dx}\right)\left(\frac{h}{h_m}\right)^3 - \left(\frac{h}{h_m}\right) + 1 = 0 \qquad (31)$$

The Equation (31) can then be solved to define the required values of h(x).

3.2.5 Boundary Conditions in Terms of Pressure

Obviously the fluid film pressure must start to build in the inlet zone, where the film is convergent. In terms of the minute scale of the complete elastohydrodynamic conjunction, then, the pressure must be ambient (taken as zero) well away from the high pressure region. This is implicit and a suitable inlet boundary location may be readily established by trial and error. (In "starved" E H L contacts the lubricant supply is assumed to be such as to limit the location of the inlet boundary of pressure rise). Similarly there will be a point in the outlet region where the film pressure will return to zero. We shall name that the 'outlet boundary location'. Beyond that point there will be no further pressure drop since conventional lubricants are incapable of sustaining significant sub-ambient pressures because of the expulsion of dissolved air from the liquid, a process known as cavitation.

Consider then a point at the outlet boundary, not only will pressure be zero but the pressure gradient $\frac{dp}{dx}$ will also be zero. This is known as a 'cavitation boundary condition' and an interative approach is required to locate the outlet boundary accordingly. Any realistic mathematical modelling of the E H L conjunction will take account of cavitation at film outlet.

3.2.6 Results

A single solution of the E H L problem will thus generate a string of numbers representing p=p(x) and h=h(x) which satisfy the elasticity and Reynolds' equations and are appropriate in terms of the inlet and oulet boundary locations. The load capacity per unit length of cylinder may then be calculated by a numerical integration routine.

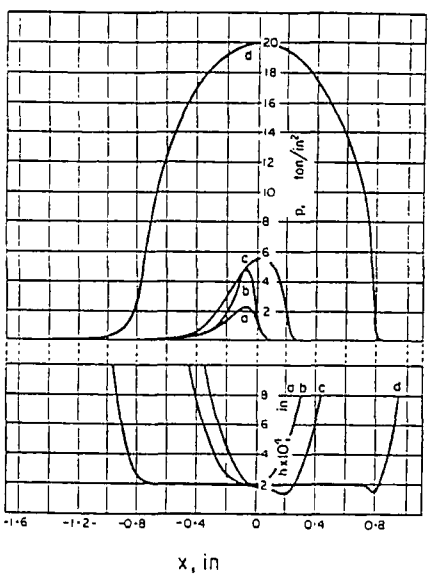
$$F' = \int_{xin}^{xout} p\,dx$$

A large number of such result sets may then be used to generate approximate expressions sumarising the work in terms of the dimensionless centreline film thickness of Equation (15) or, more important, the minimum film thickness:

$$H_{min} = 2.65 \frac{U_e^{0.7} G^{0.54}}{(W'_e)^{0.13}} \qquad (32)$$

Such expressions relate, in dimensionless form, film thickness to speed, materials properties and load. Equation (32) demonstrates the same insensitivity to load as did Grubin's Equation (14) for the central film thickness. We shall now examine some of Dowson and Higginson's results to demonstrate the interesting properties of the E H L conjunction.

Figure 5 shows that neither the pressure dependence of viscosity not the elastic nature of the conjunction is, of itself, particularly important in influencing the load capacity of the EHL conjunction, it is the two together which produce the startling effect. The load capacity is proportional to the area of the pressure profile.

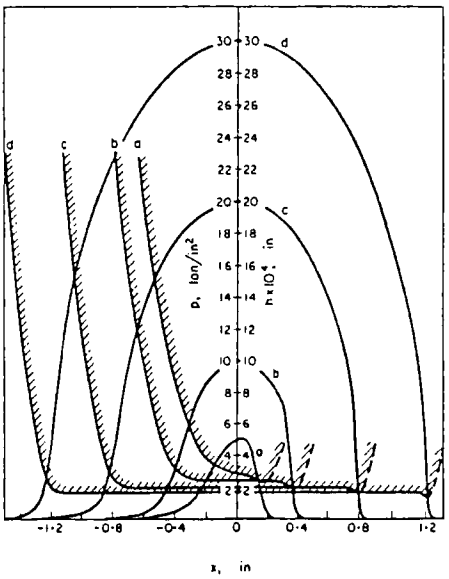


Pressure distributions and film shapes for the same centre-line film thickness. (a) Constant viscosity, rigid cylinders. (b) Pressure-dependent viscosity, rigid cylinders. (c) Constant viscosity, elastic cylinders. (d) Pressure-dependent viscosity, elastic cylinders.

[Reproduced from the *Journal of Mechanical Engineering Science*, Vol. 1, No. 1, pp. 6–15 (1959) by kind permission of the Institution of Mechanical Engineers.]

FIGURE 5

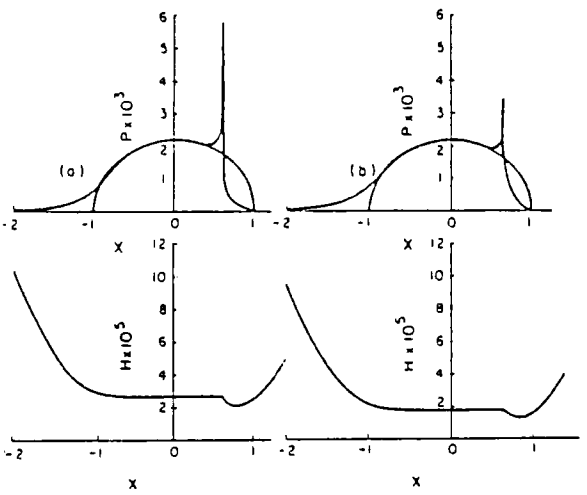
Figure 6 shows the effect of increasing load on an E H L conjunction and clearly demonstrates its insensitivity in terms of both central and minimum film thicknesses. These calculations referred to low values of the dimensionless speed and materials parameters (U_e and G) and thus do not show the interesting pressure spikes associated with Figures 7 and 8 which refer to more realistic conditions. The film thicknesses predicted here are sufficient to avoid contact between rollers having high quality surface finishes.



Pressure distributions and film shapes with pressure-dependent viscosity and elastic cylinders.

[Reproduced from the *Journal of Mechanical Engineering Science*, Vol. 1, No. 1, pp. 6–15 (1959) by kind permission of the Institution of Mechanical Engineers.]

FIGURE 6



(a) G = 5000 (b) G = 2500

FIGURE 7

Pressure distributions and film shapes. $W = 3 \times 10^{-5}$. $U = 10^{-11}$. (a) $G = 5000$. (b) $G = 2500$.

[Reproduced from the *Journal of Mechanical Engineering Science*, Vol. 2, No. 3, pp. 188–194 (1960) by kind permission of the Institution of Mechanical Engineers.]

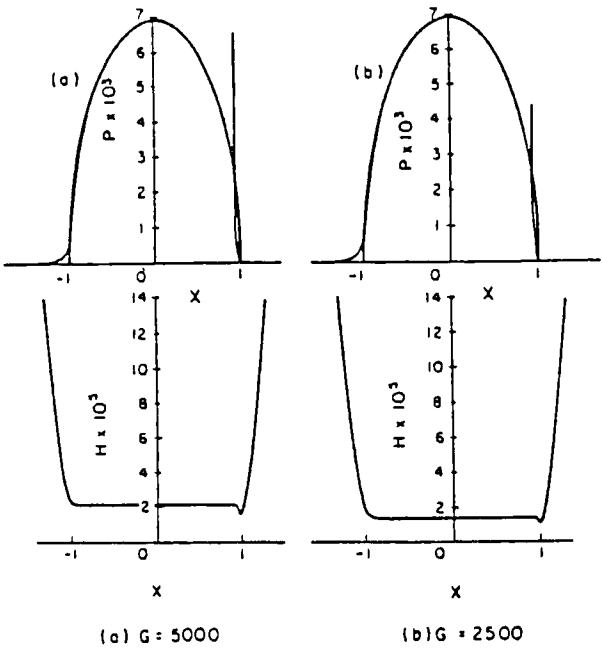
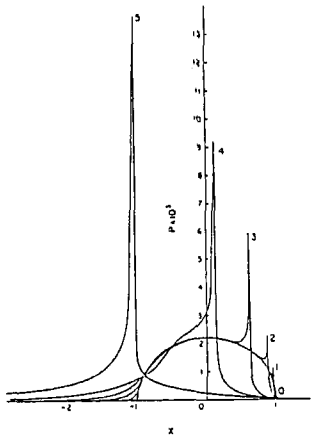


FIGURE 8

Pressure distributions and film shapes. $W = 3 \times 10^{-4}$. $U = 10^{-11}$. (a) $G = 5000$. (b) $G = 2500$.

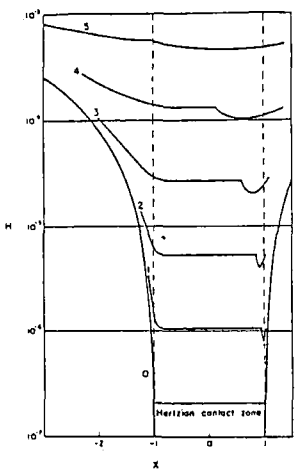
[Reproduced from the *Journal of Mechanical Engineering Science*, Vol. 2, No. 3, pp. 188–194 (1960) by kind permission of the Institution of Mechanical Engineers.]

Figure 9 shows the extraordinary effect of extreme speed on the E H L conjunction:



Pressure distributions for an incompressible lubricant. $W = 3 \times 10^{-5}$, $G = 5000$. $U = (0)$ 0 (dry contact), (1) 10^{-12}, (2) 10^{-12}, (3) 10^{-11}, (4) 10^{-10}, (5) 10^{-9}.

[Reproduced from the *Journal of Mechanical Engineering Science*, Vol. 4, No. 2, pp. 121–6 (1962) by kind permission of the Institution of Mechanical Engineers.]



Film shapes for an incompressible lubricant. $W = 3 \times 10^{-5}$, $G = 5000$. $U = (0)$ 0 (dry contact), (1) 10^{-12}, (2) 10^{-12}, (3) 10^{-11}, (4) 10^{-10}, (5) 10^{-9}.

[Reproduced from the *Journal of Mechanical Engineering Science*, Vol. 4, No. 2, pp. 121–6 (1962) by kind permission of the Institution of Mechanical Engineers.]

FIGURE 9

The numerical processes for the solution of E H L point and elliptical contacts is similar to that described though much more complex in practise.

REFERENCES

Archard J.F. and Kirk M.T. (1961) 'Lubrication at point contacts'. Proc. Roy. Soc. Ser.A 261, 532-550

Cheng H.S. (1970) 'A numerical solution to the elastohydrodynamic film thickness in an elliptical contact'. J. Lubr. Tech. 92, (1), 155-162

Dowson D. and Higginson G.R. (1959) 'A numerical solution to the elastohydrodynamic problem'. J.Mech. Eng. Sci. 1 (1), 7-15

Grubin A.N. (1949) 'Fundamentals of the hydrodynamic theory of lubrication of heavily loaded cylindrical surfaces'. In Russian. Available as D.S.I.R. Translation CTS-235

Hamrock B.J. and Dowson D. (1977) 'Isothermal elastohydrodynamic lubrication at point contacts. Part III - Fully flooded results'. J. Lubr. Tech. 99, (2), 264-276

Hertz H. (1886) 'The contact of elastic solids'. J. Reine Angew. Math. 92, 156-171

Johnson K.L. (1985) 'Contact Mechanics'. Cambridge University Press

Petrusevich A.L. (1951) 'Fundamental conclusion from the contact-hydrodynamic thoery of lubrication'. Izv. Akad. Nauk. SSSR (OTN), 2, 209.

APPENDIX ELASTOHYDRODYNAMIC LUBRICATION THEORY

Hertzian Contact Analysis and the Calculation of Elastohydrodynamic Film Thickness

The purpose of the design presented is to enable the determination of the characteristics of dry Hertzian contacts and the elastohydrodynamic film thickness in lubricated concentrated contacts.

(i) The case of line contact and point (or elliptical) contact are considered separately (see Figures 1A, 2A and 3A)
(ii) For Hertzian point contact the principal planes of the solids are taken to be coincident. If this is not the case, ESDU Design Item 78035 may be consulted.
(iii) The minimum lubricant film thickness formula and that for the film thickness at the centre of the Hertzian contact are presented for analysis of elastohydrodynamic contacts. The former has more significance but the latter may be useful in estimates of traction, (Figure 4A).
(iv) For elastohydrodynamic point contacts the direction of lubricant entrainment may be along either principal axis of the contact ellipse. If this is not the case, CHITTENDEN et al, Proc. Roy. Soc., A 397, 1985, p 271-294, may be consulted in order to modify the data quite simply.
(v) No consideration is given to the determination of contact stresses or relating these to failure criteria. The ESDU Design Items on Contact Stresses provide such data.
(vi) An example calculation for a rolling element bearing is given.
(viii) Illustrations

Figure 1A Illustration of line and point or elliptical contacts
Figure 2A Geometry of line contacts
Figure 3A Geometry of point of elliptical contacts
Figure 4A Illustration of elastohydrodynamic line contact
Figure 5A Flow chart for example calculations
Figure 6A Equivalent geometries for inner and outer ring contacts of a ball bearing

HERTZIAN LINE CONTACT

Notation

A,B	contacting cylinders
b	semi-axis of contact (Figure 2A)
E	Young's modulus of elasticity
E'	equivalent elastic modulus $\left(\frac{1}{E'} = \frac{1}{2}\left(\frac{(1-\upsilon_A^2)}{E_A} + \frac{(1-\upsilon_B^2)}{E_B}\right)\right)$

F	total load on contact (F = F' L)
F '	load per unit contact length
L	contact length
p	normal pressure
p_{max}	maximum Hertzian pressure
p_{mean}	mean normal pressure
r_{Ay}, r_{By}	radii of curvature of contacting cylinders in principal plan (Figure 2A) ($r_{Ax} = r_{Bx} = \infty$)
R_y	equivalent radius or curvature (Figure 2A) ($R_x = \infty$) (gives same separation of solids with undistorted condition)
R	given by curvature sum $\left(\frac{1}{R} = \frac{1}{R_x} + \frac{1}{R_y}\right)$ (hence R = R_y)
y,z	cartesian coordinates (contact conditions invariant with x) (Figure 2A)
δ	approach of points on cylinders remote from the contact
ν	Poisson's ratio

HERTZIAN LINE CONTACT

Formulae

$$\frac{1}{R_y} = \frac{1}{r_{Ay}} + \frac{1}{r_{By}} \qquad (1)^*$$

$$(R_x = r_{Ax} = r_{Bx} = k = \infty)$$
$$R = R_y \qquad (2)$$

$$p = p_{max}\left(1 - \left(\frac{y}{b}\right)^2\right)^{\frac{1}{2}} \qquad (3)$$

$$p_{max} = \frac{2F'}{\pi b} \qquad (4)$$

$$p_{mean} = \frac{\pi}{4} p_{max} \qquad (5)$$

$$b = \left(\frac{8}{\pi}\frac{F'R}{E'}\right)^{\frac{1}{2}} \qquad (6)$$

$$\delta = \frac{2}{\pi}\frac{F'}{E'}\left(\frac{2}{3} + \ell n\left(\frac{2\pi E'}{F'}(r_{Ay} + r_{By})\right)\right) \qquad (7)^{**}$$

* The radii (r_{Ay}, r_{By}) as used in the formula adopt positive or negative values depending on whether the solids are convex or concave in the vicinity of the undeformed contact. If the centres of curvature of the cylinders are on opposite sides of the contact point both radii will be positive, if not one will be negative.

** For identical materials. The determination of the approach of remote points on the solids is more complex for line contacts than point contacts. See,

K L JOHNSON, Proc. I.Mech.E. 1982, Vol. 196, No. 39, p 363-378

HERTZIAN LINE CONTACTS

Some useful non-dimensional relationships

$$W'\left(\frac{F'}{E'R}\right)$$

$\frac{b}{R}$	$\left(\frac{8}{\pi}W'\right)^{\frac{1}{2}}$
$\frac{\delta}{R}$	$\frac{2}{\pi}W'\left(\frac{2}{3}+\ell n\left(\frac{2}{W'}\cdot\frac{(r_{Ay}+r_{By})^2}{r_{Ay}.r_{By}}\right)\right)$
$\frac{p_{max}}{E'}$	$\left(\frac{W'}{2\pi}\right)^{\frac{1}{2}}$
$\frac{p_{max}}{p_{mean}}$	$\frac{4}{\pi}$
$\frac{b}{R}.\frac{E'}{p_{max}}$	4

HERTZIAN POINT OR ELLIPTICAL CONTACTS

Notation

A,B	Contacting bodies
E	Young's modulus of elasticity
E^1	equivalent elastic modulus $\left(\frac{1}{E^1}=\frac{1}{2}\left(\frac{1-\upsilon_A^2}{E_A}+\frac{1-\upsilon_B^2}{E_B}\right)\right)$
p	normal pressure
p_{max}	maximum Hertzian pressure
p_{mean}	mean normal pressure
R	given by curvature sum $\left(\frac{1}{R}=\frac{1}{R_x}+\frac{1}{R_y}\right)$
x,y,z	cartesian coordinates (Figure 3A)
δ	approach of points on ellipsoids remote from the contact
υ	Poisson's ratio

HERTZIAN POINT OR ELLIPTICAL CONTACTS

Formulae

$$\frac{1}{R_x}=\frac{1}{r_{Ax}}+\frac{1}{r_{Bx}},\frac{1}{R_y}=\frac{1}{r_{Ay}}+\frac{1}{r_{By}} \quad (1)^*$$

$$\frac{1}{R}=\frac{1}{R_x}+\frac{1}{R_y} \quad (2)$$

$$p=p_{max}\left(1-\left(\frac{x}{a}\right)^2-\left(\frac{y}{b}\right)^2\right)^{\frac{1}{2}} \quad (3)$$

$$p_{max}=\frac{3F}{2\pi ab} \quad (4)$$

$$p_{mean}=\frac{2}{3}p_{max} \quad (5)$$

$$a=\left(\frac{6k^2{}_1FR}{\pi E}\right)^{\frac{1}{3}} \quad (6)$$

$$b = \left(\frac{6 l_1 F R}{\pi k E'}\right)^{\frac{1}{3}} \tag{7}$$

$$k = \frac{a}{b} = \left(\frac{R_x}{R_y}\right)^{\frac{2}{3}} \tag{8}$$

$$\delta = l_2 \left(\left(\frac{9}{2 l_1 R} \right) \left(\frac{F}{\pi k E'} \right)^2 \right)^{\frac{1}{3}} \tag{9}$$

$$l_1 = 1 + \left(\frac{\pi}{2} - 1\right)\left(\frac{R_y}{R_x}\right), \quad l_2 = \frac{\pi}{2} + \left(\frac{\pi}{2} - 1\right) \ell_n \left(\frac{R_x}{R_y}\right) \tag{10}$$

- The radii (r_{Ax} ,r_{Bx} ,r_{Ay} ,r_{By}) as used in the formulae adopt positive or negative values depending on whether the solids are convex or concave in the vicinity of the undeformed contacts. If the centres of curvature of the ellipsoids in one plane are on opposite sides then both radii in that plane will be positive, if not one will be negative.

NOTE the principal radii of curvature of each contacting ellipsoids are here taken to be in the same two planes (for cases where the principal planes of one solid do not coincide with those of the other see ESDU Design Item 78035).

HERTZIAN POINT OR ELLIPTICAL CONTACTS

Some useful non-dimensional relationships for the contact between two spheres or a sphere and a plane

$$W = \left(\frac{F}{E'R^2} \right)$$

$\frac{b}{R}$	$(3W)^{\frac{1}{3}}$
$\frac{\delta}{R}$	$\frac{1}{2}(9W^2)^{\frac{1}{3}}$
$\frac{P_{max}}{E'}$	$\frac{1}{2\pi}(3W)^{\frac{1}{3}}$
$\frac{P_{max}}{P_{mean}}$	$\frac{3}{2}$
$\frac{b}{R}.\frac{E'}{p_{mean}}$	2π

ELASTOHYDRODYNAMIC FILM THICKNESS - LINE CONTACTS

Notation

A,B	contacting cylinders
e	direction of lubricant entrainment (subscript)
E	Young's modulus of elasticity
E'	equivalent elastic modulus $\left(\frac{1}{E'}=\frac{1}{2}\left(\frac{(1-\upsilon_A^2)}{E_A}+\frac{(1-\upsilon_B^2)}{E_B}\right)\right)$
F	total load on contact (F = F L)
F'	load per unit contact length
h_{cen}	film thickness at centre of Hertzian contact (Figure 4A)
h_{min}	minimum film thickness (Figure 4A)
L	contact length
R_e	equivalent radius of curvature in direction of lubricant entrainment
u	entraining velocity $\left(u=\frac{1}{2}(u_A+u_B)\right)$
α	viscosity-pressure coefficient
η_o	dynamic viscosity at atmospheric pressure and operating temperature
υ	Poisson's ratio

Non-dimensional groupings

G	dimensionless materials parameter ($\alpha E'$)
H_{cen}	dimensionless central film thickness $\left(\frac{h_{cen}}{R_e}\right)$
H_{min}	dimensionless minimum film thickness $\left(\frac{h_{min}}{R_e}\right)$

U_e dimensionless speed parameter $\left(\frac{\eta_o u}{E'R_x}\right)$

W'_e dimensionless load parameter $\left(\frac{F'}{E'R_x}\right)$

ELASTOHYDRODYNAMIC FILM THICKNESS - LINE CONTACTS

Formulae

$$h_{min} = 2.65R_x\left(\frac{\eta_o u}{E'}R_x\right)^{0.7}(\alpha E')^{0.54}\left(\frac{F'}{E'R_x}\right)^{-0.10}$$

$$h_{cen} = 3.06R_x\left(\frac{\eta_o u}{E'R_x}\right)^{0.69}(\alpha E')^{0.56}\left(\frac{F'}{E'R_x}\right)^{-0.10}$$

or in dimensionless form

$$H_{min} = 2.65U_e^{0.7}G^{0.54}W_e^{-0.13}$$

$$H_{cen} = 3.06U_e^{0.69}G^{0.56}W_e^{-0.10}$$

NOTE: The entraining direction of the lubricant (e) may quite arbitrarily coincide with either the x or y coordinate direction. If x is chosen to represent the length of the contact in which conditions are invariant, then $R_e = R_y$.

ELASTOHYDRODYNAMIC FILM THICKNESS - POINT OR ELLIPTICAL CONTACTS

Notation

A,B contacting ellipsoids

e direction of lubricant entrainment (subscript) - alone either the major or minor axis of the contact ellipse.

E Young's modulus of elasticity

E' Equivalent elastic modulus $\left(\frac{1}{E'}=\frac{1}{2}\left(\frac{(1-\upsilon_A^2)}{E_A}+\frac{(1-\upsilon_B^2)}{E_B}\right)\right)$

F total load on contact

h_{cen} film thickness at centre of Hertzian contact

h_{min} minimum film thickness

R_e equivalent radius of curvature in direction of entrainment (see e)

R_s equivalent radius of curvature in direction normal to lubricant entrainment (see s)

s direction normal to lubricant entrainment (subscript) - along either the major or minor axis of the contact ellipse

u entraining velocity $\left(u=\frac{1}{2}(u_A+u_B)\right)$ - along either the major or minor axis of the contact ellipse (in the 'e' direction)

α viscosity-pressure coefficient

η_o dynamic viscosity at atmospheric pressure and operating temperature

ν Poisson's ratio

Non-dimensional groupings

G dimensionless materials parameter ($\alpha E'$)

H_{cen} dimensionless central film thickness (h_{cen}/R_e)

H_{min} dimensionless minimum film thickness (h_{min}/R_e)

U_e dimensionless speed parameter ($\eta_o u/E'R_e$)

W_e dimensionless load parameter ($F/E'R_e^2$)

ELASTOHYDRODYNAMIC FILM THICKNESS - POINT OR ELLIPTICAL CONTACTS

Formulae

Introduce for brevity the notation

$$H^* = \frac{(h/R_e)}{\left(\frac{\eta_o u}{E'R_e}\right)^{0.68}(\alpha E')^{0.49}\left(\frac{F}{E'R_e^2}\right)^{-0.073}} = \frac{H}{U_e^{0.68}G^{0.49}W_e^{0.073}}$$

$$H^*_{min} = 3.68\left(1-e^{-0.67(R_s/R_e)2/3}\right)$$

then

$$H^*_{cen} = 4.31\left(1-e^{-1.23(R_s/R_e)2/3}\right)$$

NOTE:

(i) the entraining direction of the lubricant (e) may be along either of the principal axes of the contact ellipse, with the side leakage direction (s) then being associated with the remaining principal axis. Thus if the entraining direction (normally the minor axis) is designated the y axis in the cartesian system, then $R_s=R_x$ and $R_e=R_y$, and vice-versa if the x direction is chosen to be that of lubricant entrainment.

(ii) the notation adopted permits a simple extension to the case when lubricant entrainment is at an arbitrary angle to the principal axes (Reference CHITTENDEN et al, Proc. Roy. Soc. Lond, A 397, 1985, p 271-294).

EXAMPLE

Calculation of the Hertzian contact conditions and elastohydrodynamic film thicknesses at the inner and outer ring contacts of the most heavily loaded ball of a single row, deep groove, radial ball bearing (contact angle =0°)

Problem specification

Inner ring diameter (at contact)	($2 R_i$)	52.291 mm
Outer ring diameter (at contact)	($2 R_o$)	77.706 mm
Ball diameter	($2 r$)	12.7 mm
Number of balls	(z)	9
Inner ring groove radius	(r_i)	6.604 mm
Outer ring groove radius	(r_o)	6.604 mm
R.M.S. surface finish of balls	(σ_b)	0.045 μm
R.M.S. surface finish of rings	(σ_r)	0.125 μm

Operating conditions

Radial load	(F_R)	8900 N
Inner ring angular velocity	(Ω_i)	400 rads/s
Outer ring angular velocity	(Ω_o)	0 rads/s
Lubricant viscosity at atmospheric pressure and effective operating temperature.	(η_o)	0.01 Ns/m^2
Viscosity-pressure coefficient	(α)	2.3×10^{-8} m^2/N
Young's modulus for both rings and balls	(E)	2×10^{11} N/m^2
Poisson's ratio for both rings and balls	(ν)	0.3

For flow chart of calculations see Figure 5A.

(1) LOAD ON CONTACTS

For a bearing which is geometrically perfect (zero radial internal clearance) the load on the most heavily loaded ball is given by

$$F = \frac{4.37F_R}{Z} = \frac{4.37 \times 8900}{9} = 4321\text{N}$$

To allow for the actual radial internal clearance an iterative procedure is necessary. This results in an actual load on the most heavily loaded ball (see reference) of

$$\underline{F = 4513\ \text{N}}$$

Neglecting centrifugal effects this acts on the inner and outer ring contacts.

(2) EQUIVALENT CONTACT GEOMETRY

The calculation of the equivalent radii of curvature (Figure 3A) rendering an equivalent ellipsoid and plane separated by the same gap (with no distortion) as for the contacting solids is detailed in Figure 6A. Using subscripts i and o for inner and outer rings and noting the cartesian coordinates adopted,

$$\frac{1}{(R_x)_i} = \frac{1}{r} - \frac{1}{r_i} = \frac{1}{6.35} - \frac{1}{6.604}, (R_x)_i = 165.1\text{mm}$$

$$\frac{1}{(R_y)_i} = \frac{1}{r} + \frac{1}{R_i} = \frac{1}{6.35} + \frac{1}{0.5(52.291)}, (R_y)_i = 5.11\text{mm}$$

$$\frac{1}{(R_x)_o} = \frac{1}{r} - \frac{1}{r_o} = \frac{1}{6.35} - \frac{1}{6.604}, (R_x)_o = 165.1\text{mm}$$

$$\frac{1}{(R_y)_o} = \frac{1}{r} - \frac{1}{R_o} = \frac{1}{6.35} - \frac{1}{0.5(77.706)}, (R_y)_o = 7.59\text{mm}$$

The curvature sums are therefore

$$\frac{1}{(R)_i} = \frac{1}{(R_x)_i} + \frac{1}{(R_y)_i} = \frac{1}{165.1} + \frac{1}{5.11}, (R)_i = 4.957\text{mm}$$

$$\frac{1}{(R)_o} = \frac{1}{(R_x)_o} + \frac{1}{(R_x)_o} = \frac{1}{165.1} + \frac{1}{7.59}, (R)_o = 7.256\text{mm}$$

(3) HERTZIAN CONTACT CONDITIONS

$$E' = \frac{E}{(1-\upsilon^2)} = \frac{2x10^{11}}{0.91} = 22x10^{11}\text{N/m}^2$$

	INNER RING	OUTER RING
k	$\left(\frac{165}{5.11}\right)^{\frac{2}{3}} = 10.1$	7.79
I_1	$1+\left(\frac{\pi}{2}-1\right)\left(\frac{5.11}{165.1}\right) = 1.0177$	1.026
I_2	$\frac{\pi}{2}+\left(\frac{\pi}{2}-1\right)\ell n\left(\frac{165.1}{5.11}\right) = 3.55$	3.33
a	$\left(\frac{6x10.1x1.0177x4513}{\pi x2.2x10^{11}} x \frac{4.957}{10^3}\right)^{\frac{1}{3}} = 2.72\text{mm}$	2.61mm
b	$\frac{2.72}{10.1} = 0.27\text{mm}$	0.33mm
p_{max}	$\frac{3 \times 4513}{2x\pi x2.72x0.27x10^{-6}} = 2.93\frac{\text{GN}}{\text{m}^2}$	$2.50\frac{\text{GN}}{\text{m}^2}$
δ	$3.55\left(\left(\frac{9000}{2x1.0177x4.957}\right)\left(\frac{4513}{\pi x10.1x2.2x10^{11}}\right)^2\right)^{\frac{1}{3}} = 25.6\mu\text{m}$	25.0μm

(4) KINEMATIC CONDITIONS

The entraining velocity (u) may readily be determined as 6.252 m/s (see reference)

(5) ELASTOHYDRODYNAMIC FILM THICKNESS

$R_s = R_x$, $R_e = R_y$

	INNER RING	OUTER RING
$U_e = \frac{\eta_e u}{E' R_y}$	$\frac{0.01 \times 6.252}{2.2 \times 10^{11} \times 5.11 \times 10^{-3}} = 5.56 \times 10^{11}$	3.74×10^{-11}
$G = \alpha E'$	$2.3 \times 10^{-8} \times 2.2 \times 10^{11} = 5060$	5060
$W_e = \frac{F}{E' R_y^2}$	$\frac{4513}{2.2 \times 10^{11} \times (5.11 \times 10^{-3})^2} = 7.86 \times 10^{-4}$	3.56×10^{-4}
h_{min}	$3.68 \times (5.11 \times 10^{-3}) \times (5.56 \times 10^{-11})^{0.68} \times (5060)^{0.49} \times (7.86 \times 10^{-4})^{-0.073} \times (1 - e^{-0.67 \times 10.1}) = 0.22 \mu m$	$0.26 \mu m$
Film Thickness Parameter λ	$= \frac{h_{min}}{(\sigma_b^2 + \alpha_r^2)^{\frac{1}{2}}} = \frac{0.22 \times 10^{-6}}{(.045^2 + .125^2)^{\frac{1}{2}} \times 10^{-6}} = 1.66$	1.96

With high values of ellipticity ratio (k), as here, application of the line contact elastohydrodynamic film thickness formulae with an equivalent load per unit length based on (say) equal maximum Hertzian contact pressure would give good film thickness prediction.

REFERENCE: B J HAMROCK and D DOWSON "Ball Bearing Lubrication", 1981, Wiley-Interscience (Example p 330, modified)

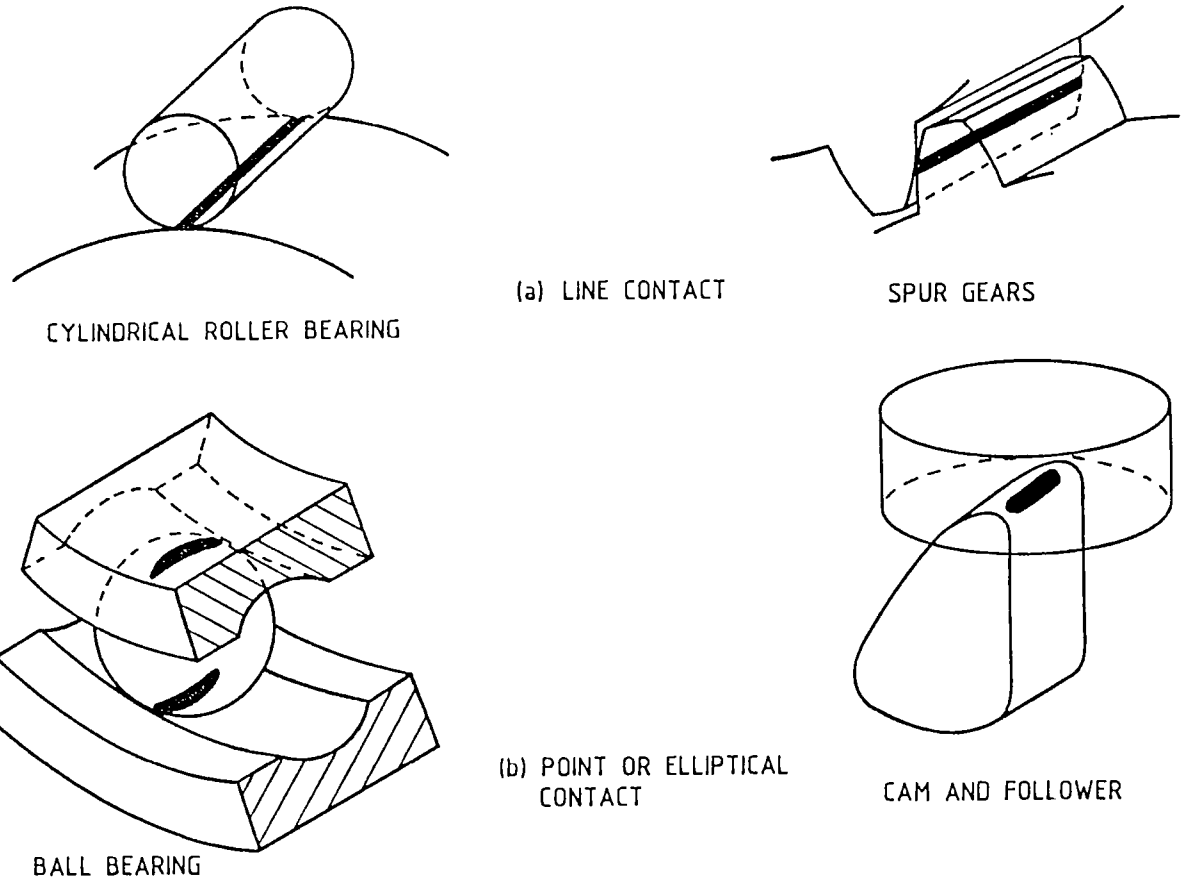


Figure 1A Illustration of Hertzian contact patch

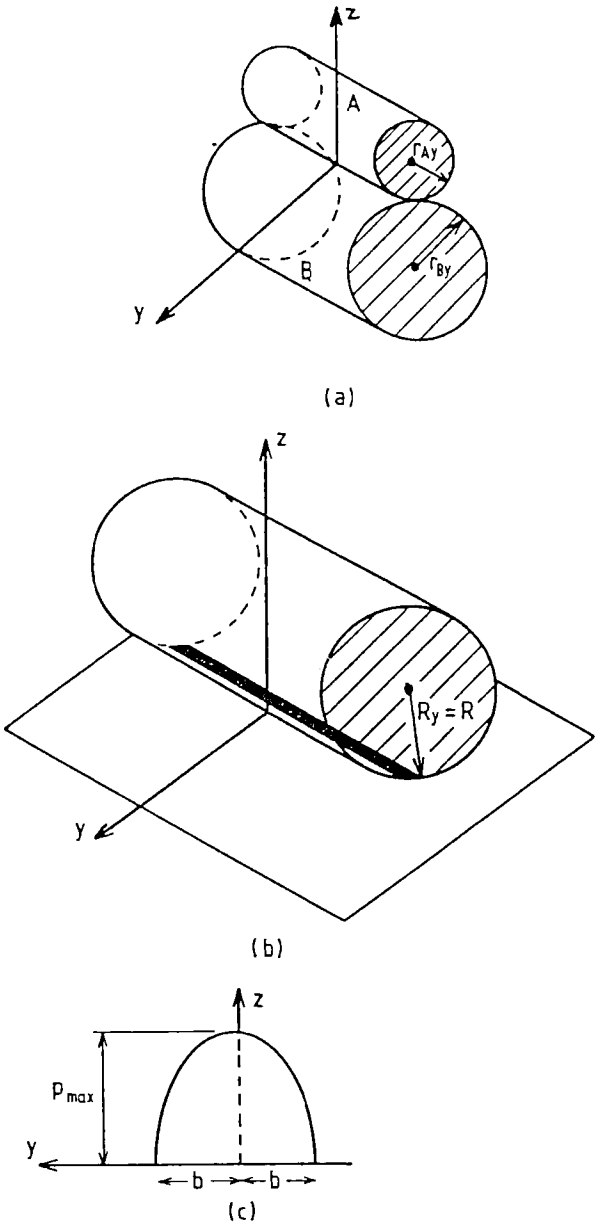


Figure 2A Geometry of line contacts
(a) Contact of two cylinders
(b) Equivalent geometry with contact patch
(c) Semi-elliptical normal pressure distribution on contact patch

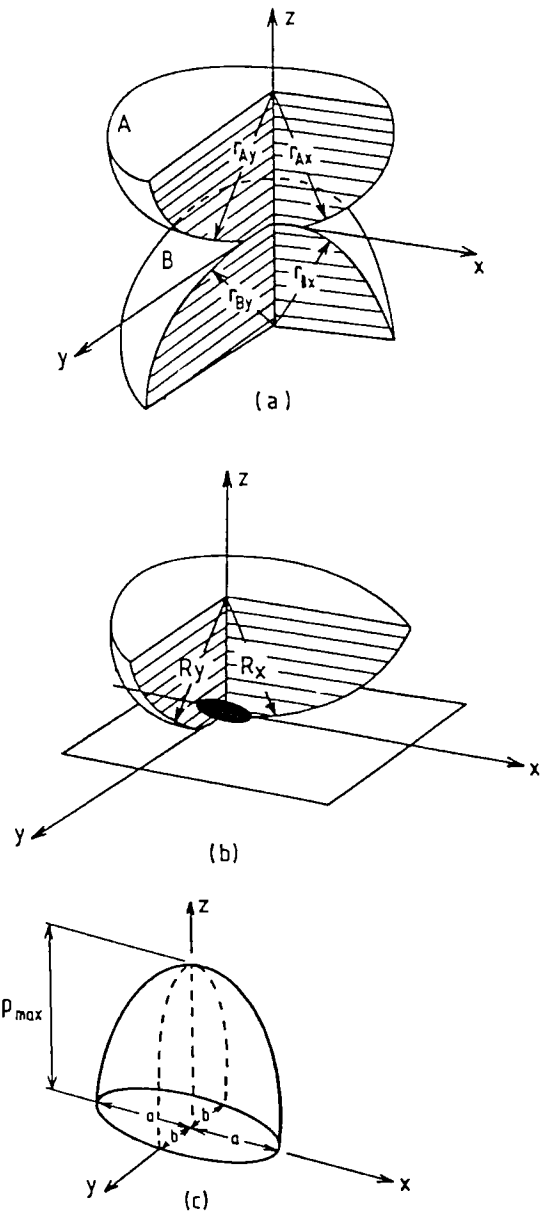


Figure 3A Geometry of point or elliptical contacts
(a) Contact of two ellipsoidal solids
(b) Equivalent geometry with Hertzian contact patch
(c) Semi-elliptical normal pressure distribution on contact patch

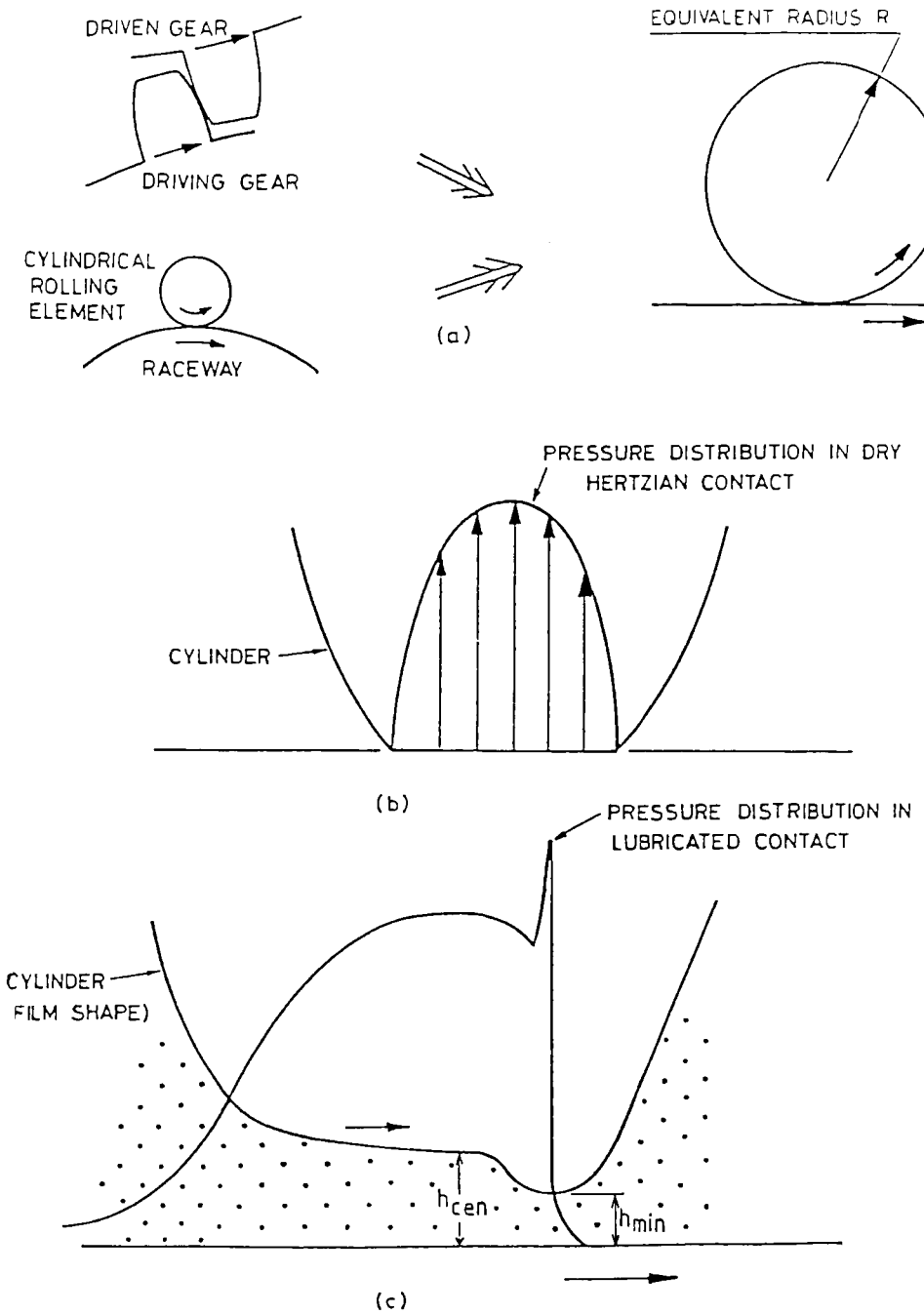


Figure 4A Illustration of elastohydrodynamic line contacts

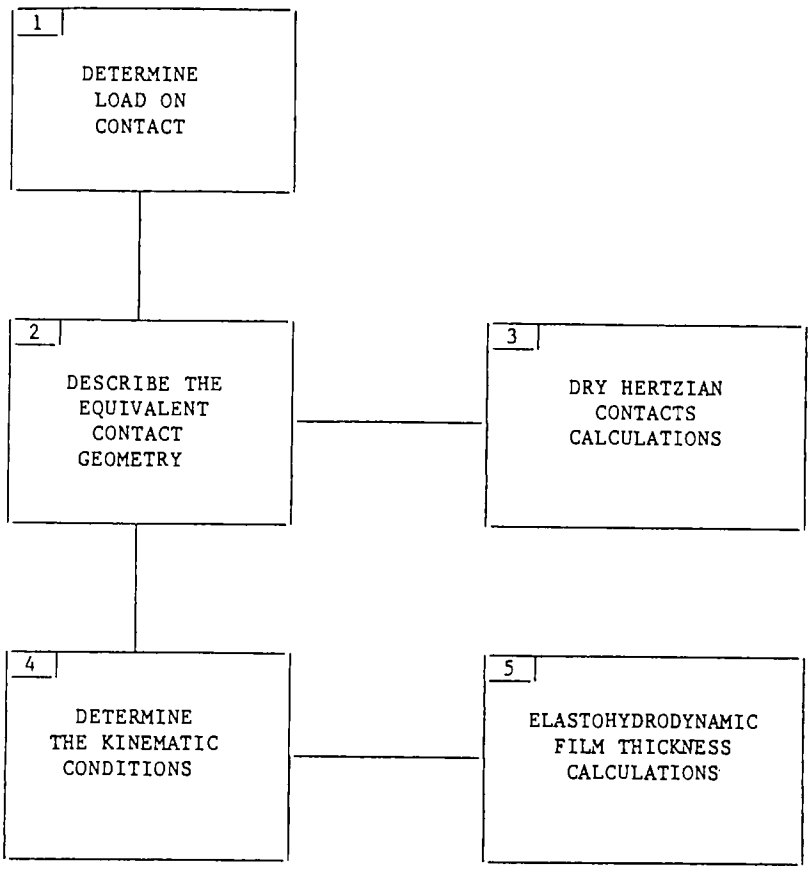


Figure 5A Flow chart for example calculations

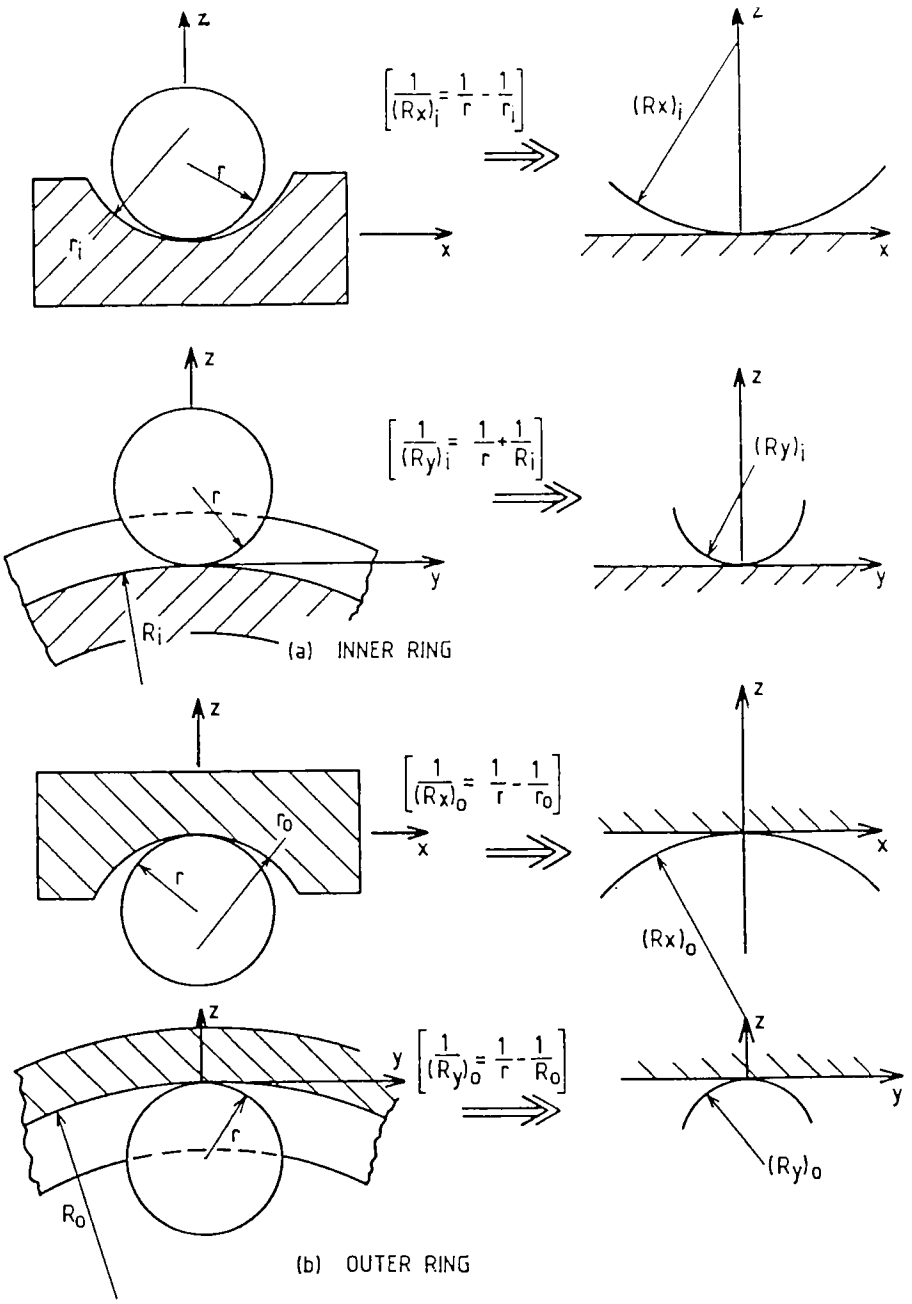


Figure 6A Equivalent geometries (same gap between undistorted solids) for inner and outer ring contacts of a ball bearing

Engine Tribology / C.M. Taylor (Editor)

Dry and boundary lubricated sliding friction and wear for engine component materials

T. H. C. Childs

Department of Mechanical Engineering, University of Leeds, Leeds LS2 9JT

Abstract

This chapter reviews by the development of friction mechanism maps the responses of sliding surfaces to friction and wear in the absence of hydrodynamic lubrication. It initially considers the classical views of adhesive and abrasive friction before discussing in more detail the variety of plastic contact flows that can occur, depending on surface roughness slope and interfacial adhesion. It then considers the conditions that result in elastic contact. Mechanisms of friction and wear in the different regimes of surface stressing are described.

1. INTRODUCTION

The purpose of this chapter is to review and provide an understanding of what happens to sliding surfaces when fluid film lubrication breaks down. This occurs when the lubricant film thickness becomes less than the heights of the surface roughness, perhaps because of slow or zero motion between the surfaces, too heavy a load or because of loss of fluid viscosity at high temperatures. Friction and wear rates invariably rise, sometimes with immediate catastrophic consequences.

The mechanisms and magnitudes of friction and wear in the absence of fluid lubrication depend on the surfaces' mechanical properties: their elastic moduli and hardness (and for ceramics also their brittle properties) and the shear strength of any remaining surface films; on the surface roughness: the important quantity is the surface asperity slope; and on their chemical reactivity with the surrounding atmosphere: commonly wear occurs not by the removal of material with the composition of the bulk but by removal of surface reaction films. The lecture will consider the influence of these on friction and wear both to form a view as to what might be minimum magnitudes of friction and wear in the absence of fluid lubrication, and to establish design requirements (such as surface finish) for avoiding early failures.

The chapter is divided into four parts. Bowden's and Tabor's original views (1,2) of adhesive and ploughing friction and wear at plastic micro-contact areas between sliding solids are first briefly reviewed for their physical insight and as a base from which later developments have occurred. A closer consideration is then given to the ways in which plastic contacts can flow, in order to distinguish between benign, surface smoothing or wave, flows and conditions of prow or chip formation that lead to rapid failure. It is now recognised that run-in surface contacts are not plastic but elastic: the third section will establish the conditions for elastic contact and represent these in a friction mechanism

map. Finally the range of wear mechanisms that accompany the different friction mechanisms is explored.

2. INITIAL CONCEPTS OF FRICTION AND WEAR

2.1. Adhesive and deformation (or ploughing or abrasive) friction

Bearing surfaces are usually loaded so lightly that even in the absence of fluid films they touch over only a small fraction of their apparent contact area. The interface at the real contacts can be inclined to the sliding direction. Figure 1 illustrates an artificially simple example in which the roughness of the upper surface is imagined to be a collection of wedges of constant slope θ and the interface normal stresses p and tangential stresses s are the same on each contact and only act on the leading face of each. Both s and p contribute to the reaction against the friction force F and the normal load W. Resolution of forces gives the friction coefficient μ to be

$$\mu = F/W = (p \sin\theta + s \cos\theta)/(p \cos\theta - s \sin\theta) \quad (1)$$

In general s, p and θ may vary from contact to contact and (1) may be modified to include this. Two special cases are particularly simple to consider. If $\theta=0$,

$$\mu_{adh} = s/p \quad (2)$$

and if $s=0$

$$\mu_p = \tan\theta \quad (3)$$

Equations (2) and (3) are the well known expressions for the adhesive and ploughing components of friction (1), the first being a ratio of mechanical properties, depending most variably on the interface shear strength s, and the second depending on surface roughness. They both are in accord with Amonton's Laws that state that friction coefficient is independent of load and apparent area of contact.

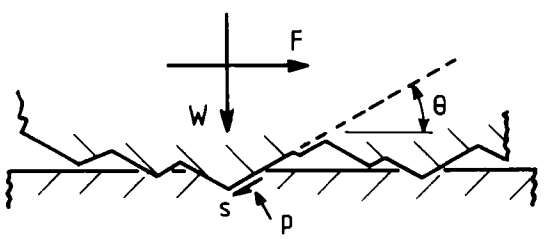


Figure 1. A sliding contact between two surfaces

2.1.1. Adhesive friction and junction growth.

At the time that equation (2) was developed it was assumed that real areas of contact between metals would be plastically stressed. The interaction between s and p was used to explain the high sliding friction coefficients, 1.0 or larger, that were observed between clean ductile metals. The key point was noted that the area of a plastically loaded junction would grow if a friction force were added to it. By analogy with the combined loading and shear of a block between two parallel platens, figure 2a, for which p and s can be related by the Tresca yield criterion equation 4a, p and s acting on an adhesive friction junction, figure 2b, may be assumed to be related by an equation of the form of 4b, where α and β are assumed to be constants, and k is the shear yield stress of the plastically loaded material.

$$p^2 + 4s^2 = 4k^2 \tag{4a}$$

$$p^2 + \alpha s^2 = \beta k^2 \tag{4b}$$

Equation (4b) may be used to relate the area A of the junction when acted on by F and W together to its area A_o when acted on by W alone:

$$(A/A_o)^2 = 1 + \alpha(F/W)^2 \tag{5}$$

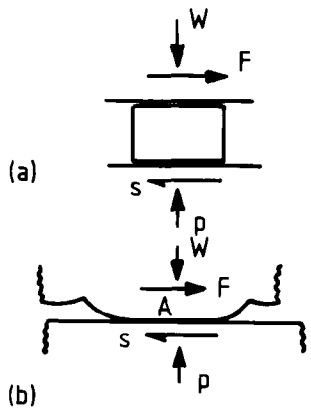
The contact pressure when the area is A can be substituted in (2), and after some rearrangement an expression for the friction coefficient allowing for the junction growth is

$$\mu = \frac{s}{k}\left\{\beta - \alpha\left(\frac{s}{k}\right)^2\right\}^{-\frac{1}{2}} \tag{6}$$

If α and β are constants, then equation (4b) suggests that they will be equal, because when p=0 and for an extremely clean interface, s may be expected to equal k. If $\alpha = \beta$, and when s=k, the friction coefficient according to equation (6) becomes infinite. There is a further expectation from equation (4b) that β=25, for in the absence of sliding the contact has the geometry of a hardness indent and p may be identified with the hardness H=5k [2]. In that case, equation (6) indicates that if either α/β or s/k is reduced from 1.0 to 0.95, the predicted friction coefficient falls from infinity to less than 1.0: the prediction of infinity is critically dependent on the assumptions of the model (equation 4b) as well as on the cleanness of the interface.

2.1.2. Ploughing friction and geometry.

In the absence of adhesion, the prediction of equation (3) for ploughing friction varies slightly with the details of the assumed shape of the contact. Figure 3 shows a cone scratching a metal flat. If the contact pressure is assumed uniform over the interface and the interface is assumed to be just the area of intersection between the indenter and the flat (i.e. extra contact due to deformation of the flat ahead of the indenter is ignored), integration over the interface of the components of pdA resolved in the directions of F and



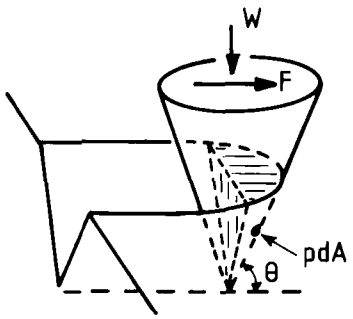


Figure 2. Combined loading and shear of (a) a block, (b) a friction junction
Figure 3. A cone scratching a flat surface

W leads to equation (7a) for the friction coefficient. A similar calculation for a sphere ploughing a flat leads to equation 7b, where tanθ is the slope of the sphere at the edge of the contact.

$$\mu = [2/\pi]\tan\theta \tag{7a}$$

$$\mu = [4/(3\pi)]\tan\theta \tag{7b}$$

These differ from equation (2) by a simple numerical factor, but apparently confirm the essential geometric nature of the ploughing friction component.

2.2. Adhesive and abrasive wear

The adhesive and abrasive models of friction in figures 2 and 3 may be used to discuss wear (2,3). Suppose that in the adhesive friction case of figure 2 the contact is circular with diameter 2a and that every time it slides through its own length it produces a hemispherical wear particle of volume $(2/3)\pi a^3$. If junction growth is neglected a may be related to the load on the contact,

$$W = \pi a^2 H \tag{8}$$

and an expression developed for the total wear volume V in sliding distance L:

$$V = (1/3)(W/H)L \tag{9}$$

Similarly, in the abrasive conditions of figure 3, if it is assumed that all the material displaced from the groove behind the indenter becomes wear debris and that the pressure p is also equal to H, it may be shown that

$$V = (2\tan\theta/\pi)(W/H)L \tag{10}$$

Both equations (9) and (10) have the form of the wear law now associated with Archard's name (4):

$$V = K(W/H)L \quad \text{or} \quad V = kWL \tag{11}$$

where K is dimensionless and is known as the wear coefficient and k has inverse stress units and is called the specific wear rate. Equations (9) and (10) suggest that K has values, depending on θ, of the order 0.1 to 1.0.

2.3. Limitations of the initial concepts

Equations (2) to (11) present the simplest description of friction and wear with any claim to reality. Yet as far as ploughing deformation is concerned they give no information about the flow pattern of material round the indenter, and a consideration of deformation combined with adhesive friction is avoided. In the case of adhesive friction, the theory takes s/k (which may be thought of as a surface contamination factor) as given. It has nothing to say about the nature of interfacial rheology that controls s. Further by no means all metal sliding contacts are plastic: hard polished surfaces and many run-in surfaces suffer elastic or elasto-plastic deformation. Finally, as summarised in figure 4, equations (9) and (10) grossly overestimate measured wear rates (5). Further review of these topics is developed in the remainder of this lecture.

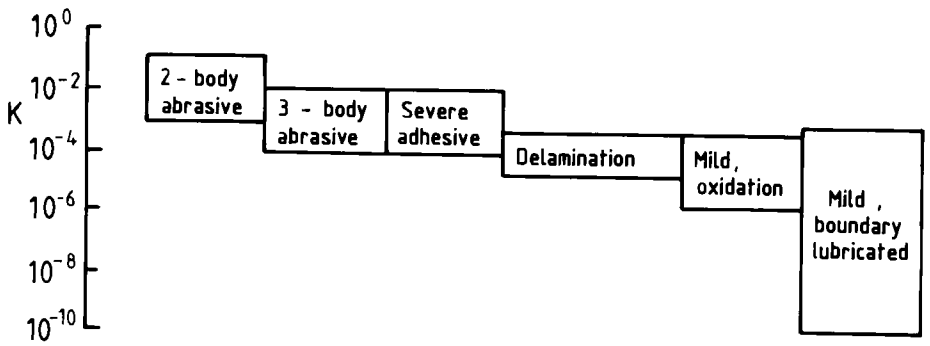


Figure 4. The range of wear coefficients for wear modes in section 5

3. PLASTIC FLOW OF METAL SLIDING FRICTION CONTACTS

When newly installed metal bearings are first run, the contacts usually deform plastically. The surfaces can either become smoother and run-in or become rougher and fail. Which occurs depends on the surface roughness slopes and the surface shear strength

ratio s/k. The purpose of this section is to establish what are the critical values of these variables and to give insight into the mechanisms of surface failure that can occur when they are exceeded, either because the as-formed surface is too rough or poorly lubricated or because the effective roughness is increased by the presence in the contract of trapped wear debris or hard particulate contamination.

Model experiments in which a metal is scratched by a hard cone or sphere (6,7) show that the metal can flow round the indentor as a wave, without any material removal, if the indentor is blunt and well lubricated. At the other extreme metal is removed as swarf or a chip. In between, initial sliding of the indentor results in displacement of a prow or wedge of metal that becomes trapped ahead of the indentor. If that entrapment is stable further sliding occurs by wave flow round the wedge. Alternatively the wedge may be expelled from the contact. The formation process is then repeated. The conditions of slope and adhesion that lead to these flows may be studied for plane strain flows of ploughing by a wedge shaped asperity, for rigid plastic, non-workhardening metals, by means of slip-line field theory. The results may then be compared with experiments on three-dimensional flows with real metals. Slip-line field theory may also be used to analyse junction growth more rigourously than in the discussion of the previous section.

3.1. Essentials of slip-line field plasticity analysis

The maximum shear stress in a plane strain flow of a non-hardening plastic material is everywhere k. Slip-line field theory provides a means of calculating the trajectories of the maximum shear stress (the slip-lines) and the variations of hydrostatic pressure throughout the plastic region.

The hydrostatic pressure variations are the starting point. Figure 5a shows slip lines α and β which at point 0 are inclined at ϕ to the cartesian axes x,y. Analysis of stress at 0 gives

$$\sigma_{xx} = -p-k \sin 2\phi$$

$$\sigma_{yy} = -p+k \cos 2\phi \qquad (12)$$

$$\tau_{xy} = k \cos 2\phi$$

Substitution of equations (12) into the stress equilibrium equations

$$\partial\sigma_{xx} / \partial x + \partial\tau_{xy} / \partial y = 0 \; ; \; \partial\sigma_{yy} / \partial y + \partial\tau_{xy} / \partial x = 0 \qquad (13)$$

and considering the case when $\phi = 0$, gives

$$d(p+2k\phi) = 0 \text{ in the direction of an } \alpha \text{ line}$$

$$d(p - 2k\phi) = 0 \text{ in the direction of a } \beta \text{ line.} \qquad (14)$$

Thus if p is known anywhere in the field, it may be calculated anywhere else, provided the trajectories of the slip lines are known. Equations (14) themselves provide rules for obtaining the trajectories. The variation of ϕ along an α or β line must be such that the calculated pressure change is zero in traversing, in figure 5a for example, from O to A, A to B, B to C and back to O. Two common boundary conditions to determine the extent of the field and to give a starting point for pressure calculations are the rough surface and free surface conditions. Figure 5b shows slip lines intersecting, at an angle ζ a rough surface on which the shear stress is s. Force equilibrium parallel to the surface gives

$$\zeta = 0.5\cos^{-1}(s/k) \tag{15}$$

At a free surface (figure 5c), the conditions that $s = 0$ requires the slip lines to intersect the surface at an angle of $\pi/4$, while zero normal stress requires $p = k$.

A more detailed account of the construction of slip-line fields, of the calculation of velocities in the fields and of force and velocity boundary conditions may be found in standard texts [8,9].

3.2. Plastic models of ploughing

Figure 6 shows two possible slip line field models of wave and chip forming flows when a rigid wedge slides over a plastic flat [10-12]. It is required by their geometry that

$$\theta \leq \zeta \text{ ,for the wave flow}$$
$$\theta \geq \pi/2-\zeta \text{ , for the chip flow} \tag{16}$$

where ζ is defined in equation (15). The consequent range of validity of these fields is mapped on to the s/k, θ plane in figure 7a and the friction coefficients calculated from them are shown in figure 7b. For the wave flow, for example, friction coefficient is calculated from equation (1): force equilibrium normal to the wedge surface gives

$$p_n/k = p_2/k + \sin 2\zeta \tag{17}$$

Equation (14) and the free surface boundary condition give

$$p_2/k = 1+2\phi_1 \tag{18}$$

The value of ϕ_1 is determined by the volume conservation requirement that the tip of the wedge lies at the level of the surface:

$$\sin\theta = \sqrt{2}\sin\xi\sin\left(\frac{\pi}{4}+\xi-\phi_1-\theta\right) \tag{19}$$

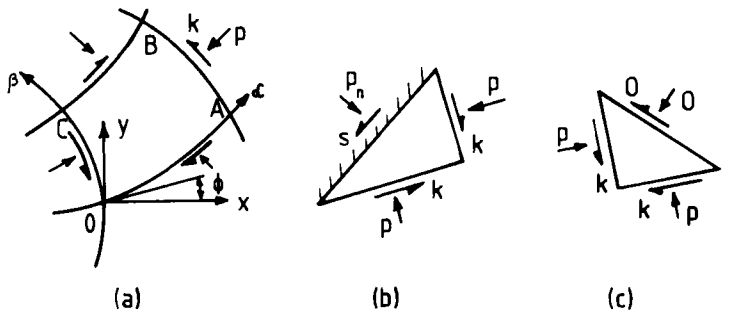


Figure 5. (a) α and β slip lines inclined at θ to cartesian axes x, y at 0 and diagrams concerning (b) rough and (c) free surface boundary conditions

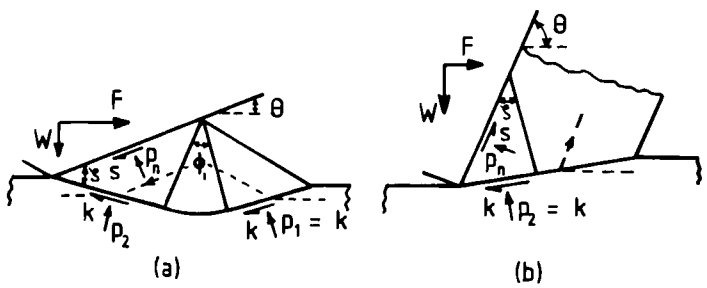


Figure 6. (a) Wave and (b) chip forming flows over a wedge tool

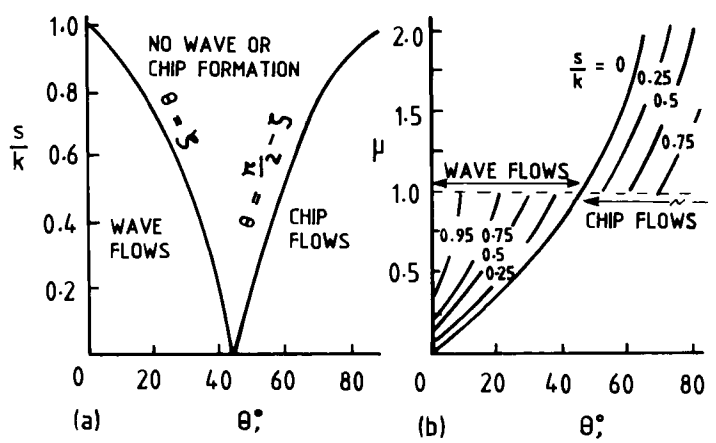


Figure 7. (a) flow mechanism map and (b) friction coefficients for the fields of figure 6.

For the wave flow the effect of adhesion, ($s/k > 0$), is to increase the friction coefficient relative to its value when $s/k = 0$. For the chip forming flow the sliding direction of the metal over the wedge is opposite to that of the wave flow. This leads to a reduction in the friction coefficient the greater is s/k.

Equations (16) and figure 7a show that for a range of θ about 45°, and of increasing width the greater is s/k there is no wave or chip forming flow. It may be surmised that these are the conditions that lead to prow or wedge formation. However the fields of figure 6 are not the only possible ones (13). Surface plastic flow is highly non-unique. Further fields are shown in figure 8: more complex wave flows (I), combined wave and wedge flows (II), wedge flows (III), combined wave and cutting flows (IV) and cutting flows with and without stagnant zones (V and VI). Their range of validities are shown in figure 9a. It is seen that for wave flows to be the only possibility (region I),

$$\theta < 0.5\,\zeta \qquad (20)$$

and that the region III where prow formation is the only possibility is much reduced relative to that in figure 7a. The non-unique flow fields result in a more complex dependence of friction on s/k and θ (figure 9b). For $\mu < 0.4$ or > 1.0 the situation is close to that seen in figure 7b but at intermediate values of μ friction is not uniquely determined by s/k and θ. Its value becomes less determinate the larger is s/k, as indicated by the areas enclosed by the dashed, dotted and dash-dotted lines for $s/k = 0.5$, 0.75 and 1.0 respectively. These conclusions of plane strain plasticity theory are qualitatively supported by experiments on real surfaces (14).

In summary, to avoid any possibility of prow or chip formation surface finish and lubrication should satisfy equation (20). As a rough rule, this will be the case when $\mu < 0.4$.

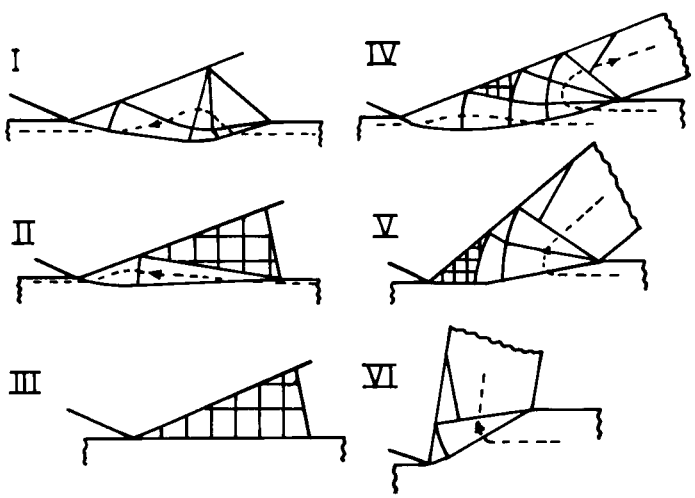


Figure 8. Further wave, wedge and chip forming flows

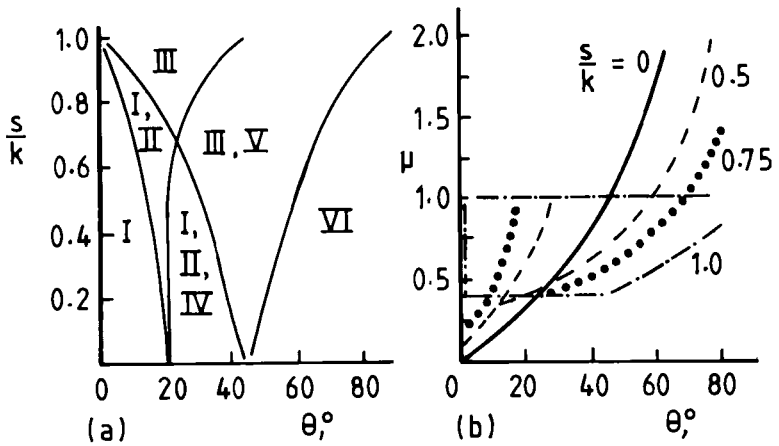


Figure 9. (a) flow mechanism map and (b) friction coefficients for the fields of figure 8

3.3. Junction growth

The considerations of the previous section visually accord with abrasive conditions. Slip line field models of sliding a plastic rough surface over a rigid flat have also been constructed (11, 14, 15). They show that equation (4b) only approximately describes the combined stress conditions in the junction. Although equation (5) reasonably describes the junction growth that occurs (with α increasing from about 12 to 20 as asperity slope θ decreases from 30° to 0°), the sizes of α and β conspire to maintain μ from equation 6 less than or equal to 1.0. As μ approaches 1.0 further steady deformation of the asperity becomes impossible. For two surfaces of equal hardness, a transition to scoring is anticipated as is observed experimentally (16), or fragments of the softer surface may transfer to the harder, as will be considered in section 5.

4. THE TRANSITION TO ELASTIC CONTACT CONDITIONS

The contact between metals is not always plastic. In this section the conditions for elastic contact of a model asperity - a sphere or cylinder - pressed statically on to a flat surface are first considered. This is an extension of the Hertzian contact analysis introduced in the context of ehl in an earlier lecture. This is then extended to a criterion for elastic or plastic contact of static rough surfaces, before finally the conditions for elastic contact of sliding surfaces are developed and added to the friction mechanism map of figure 9a. The surface material and geometry dependence of friction is then fully discussed.

4.1. Elastic hertzian contact and its breakdown

Figure 10a sketches a section through an elastically loaded sphere or cylinder, radius R, on an elastic flat, showing the Hertzian elliptical contact pressure over the contact width 2a. It is well known (18) that the mean contact pressure $\bar{p}$ is

$$\bar{p} = nE^*(a/R) \tag{21}$$

where n = 0.42 for a sphere or 0.39 for a cylinder and

$$1/E^* = (1 - \nu_1^2)/E_1 + (1 - \nu_2^2)/E_2 \tag{22}$$

where suffix 1 and 2 denotes properties of the two bodies. Equation (21) has the form of Hookes Law if a/R is identified as the mean strain in the contact. a/R is also geometrically the slope of the sphere or cylinder at the edge of the contact before its deformation. The relation between mean contact pressure and slope may be emphasised, and will be made use of later, by comparing equation (21) with the expression for the mean contact stress between an elastic wedge of slope θ and a flat

$$\bar{p} = 0.5E^*\tan\theta \tag{23}$$

It is also well-known that the maximum shear stress occurs not at the surface but at a depth 0.48a (for a sphere) or 0.785a (for a cylinder) beneath the centre of the contact. Its value is $0.46\bar{p}$ for a sphere or $0.38\bar{p}$ for a cylinder (for an elastic wedge a maximum shear stress of $0.64\bar{p}$ occurs at its apex). Yielding commences when the maximum shear stress reaches 0.5Y (according to the Tresca criterion). The values of $\bar{p}/Y$ and of the deformation severity parameter $(E^*/Y)(a/R)$ or $(E^*/Y)\tan\theta$ when this occurs, deduced from the numbers in this paragraph are listed in Table 1.

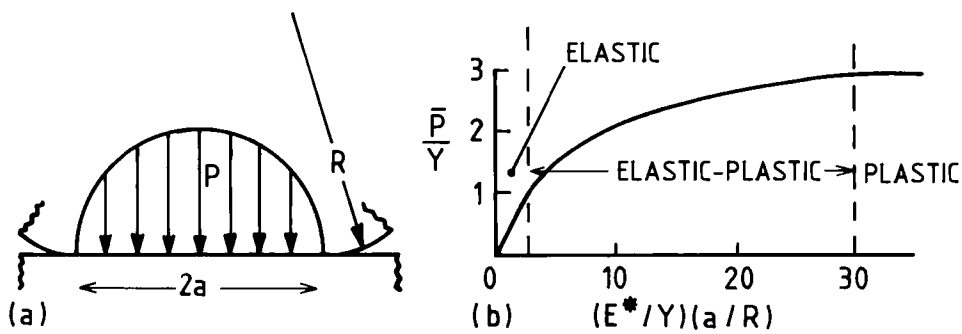


Figure 10. (a) a Hertzian contact, (b) the variation of $\bar{p}$ up to and beyond yield

Table 1. Critical values for the onset of yielding

Contact Shape	Sphere on flat	Cylinder on flat	Wedge on flat
$\bar{p}/Y$	1.1	1.3	0.8
$(E^*/Y)(a/R)$	2.6	3.3	-
or $(E^*/Y)\tan\theta$	-	-	1.6

Figure 10b follows the change in $\bar{p}/Y$ with $(E^*/Y)(a/R)$ for the case of a sphere on a flat. Up to first yield the relation is linear, as indicated by equation (21). As loading increases the plastic zone beneath the surface increases in size until at about $(E^*/Y)(a/R) = 10$ it breaks through to the free surface, when $\bar{p}/Y=2$. Finally as $(E^*/Y)(a/R)$ increases to 30 the indentation becomes fully plastic and $\bar{p}/Y$ stabilises at the value of 3, the hardness of a material equalling 3Y (18). The relevance of this to the contact stresses between static rough surfaces is the subject of the next section.

4.2 Elastic and plastic contact between static rough surfaces

Figure 11a shows the static contact between a flat and a rough surface modelled as an array of equal spheres, all with their centres at the same distance from the flat. For such a surface an increase of the load W results in an increase of the load per contact and of the width of each contact. In the elastic state this results in the increase of the contact pressure $\bar{p}$ (equation 21) and, if sliding were to be attempted, the adhesive friction coefficient s/p would fall. This contradicts Amonton's Law that friction is independent of load. If, however, as illustrated in figure 11b, the spheres are distributed at varying heights above the flat, an increase in load can result in an increase in the number in contact, a smaller increase in the average contact pressure and a lesser change in friction coefficient. This situation was first analysed by Greenwood and Williamson (19) and later developed by Archard and Whitehouse (20) and reviewed by Archard (21).

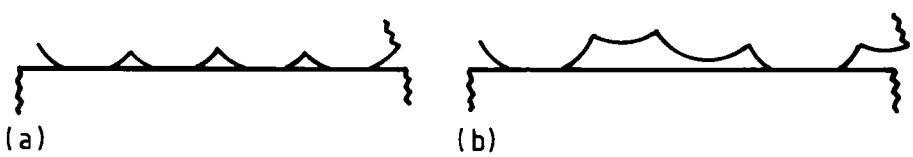


Figure 11. Models of (a) regular and (b) random height rough surfaces

Greenwood and Williamson approximated roughness to an array of spherical tips of constant radius R but with peak heights distributed with a standard distribution σ_s. They showed that mean contact pressure was almost totally independent of load for realistic values of σ_s and bearing loads. They found

$$\bar{p} = nE^* (\sigma_s/R)^{0.5} \qquad (24)$$

where n = 0.56 for an exponential distribution of peak heights or 0.3 to 0.4 for a Gaussian height distribution. A close similarity to equation 21 is seen if $(\sigma_s/R)^{0.5}$ is identified with a/R or surface roughness slope.

The similarity may be developed to consider whether the average contact will be elastic or plastic. $\bar{p}/Y$ will vary in the same way with $(E^*/Y)(\sigma_s/R)^{0.5}$ as p/Y for a single contact varies with (E*/Y)(a/R), shown in figure 10b. The quantity $(E^*/Y)(\sigma_s/R)^{0.5}$ is known as the plasticity index for the surface (although other definitions use the ratio E*/H or E*/k). Table 2 lists typical values of (E*/Y) for the contacts of different classes of materials, together with typical surface slopes $(\sigma_s/R)^{0.5}$ produced by different manufacturing processes, and records values of the resulting plasticity indices. It shows that many newly manufactured bearing surfaces will have plasticity indices between 2.5 and 30 and will thus have contacts in the elasto-plastic state.

Table 2. Plasticity index values for different materials and finishes

Material combination	**Typical (E*/Y) range**	**Typical $(\sigma_s/R)^{0.5}$ range**		
		smooth: 0.005 to 0.1	medium: 0.1 to 0.2	rough: 0.2 to 0.5
metal on metal	50 to 500	0.25 to 50	5 to 100	10 to 250
polymer on metal	15 to 60	0.08 to 6	1.5 to 12	3 to 30
ceramic on metal	50 to 100	0.25 to 10	5 to 20	10 to 50

Other quantities of interest may be calculated. The mean diameter 2a of a contact is shown (19) to be

$$2a = 2(\sigma_s R)^{0.5} \qquad (25)$$

If a surface runs-in to a plasticity index close to 2.5, the mean real contact pressure will be close to Y. The degree of contact (the ratio of the real to apparent contact) will then be p_{nom}/Y where p_{nom} is the nominal contact pressure. Table 3 lists σ_s, R and 2a for data measured by Greenwood and Williamson (19). Contacts of diameter about 10μm are

Table 3. Some examples of σ , R and 2a

Contacting surfaces	**σ_S , μm**	**R , μm**	**2a , μm**
Bead blasted aluminium	1.37	13	8
Ground mild steel	0.6	15	6
Ground, lightly polished steel	0.5	85	13
Well polished steel	0.1	500	14
Worn, run-in steel	0.065	250	8
Hard steel roller bearing	0.024	150	4

common. Table 4 lists the degree of contact for automotive bearing surfaces loaded to typical maximum values: values range from about 5 to 50%. A feature that has been glossed over in this discussion is that values of σ_s and R vary with measurement sampling interval: a sensible value must be chosen for this.

Table 4. Maximum degrees of contact

Engine component	Y, MPa	Maximum recommended nominal pressure, MPa	Degree of contact, %
Piston ring	600	20	3
Cam and tappet	2000	1000	50
Crankshaft bearing	100	20	20

4.3. The influence of sliding on the contact of rough surfaces

Figure 10b, when interpreted as an indication of the dependence of contact stress on plasticity index $(E^*/Y)(\sigma_s/R)^{0.5}$ shows that as plasticity index increases from 2.5 to 30 the contact changes from elastic, to elasto-plastic, to totally plastic. The question arises: how does sliding affect the transition from elastic to plastic deformation?

4.3.1. Surface shear strength characteristics.

An initial question is how best to describe the surface shear traction in elastic conditions? In plasticity analyses, as in section 3, the surface shear strength is usually taken to be independent of contact pressure, as in equation 26a, but in elastic analyses s is assumed proportional to pressure, equation 26b.

$$s/k = m = \text{constant} \quad (26a)$$

$$s/p = \mu = \text{constant} \quad (26b)$$

The appropriate form probably depends on the nature of the interfacial shear film. There is experimental evidence that interfacial shear strength of thin films is proportional to contact pressure for graphite and molybdenum disulphide (22), for polymer films (23), and for lubricant fluids at high shear rates (24), while for soft metal films, shear strength is independent of pressure, figure 12.

4.3.2. Modification of Hertzian contact stress.

The addition of the surface traction 26a or b to an elastically loaded Hertzian contact modifies the maximum shear stress distribution round the contact in two ways. The region of largest maximum shear stress, which in the absence of traction lies at a depth from 0.5a to 0.8a below the centre of the contact (depending on the contact shape), is displaced towards the surface. For small values of m or μ ($m \leq 0.5$, $\mu \leq 0.25$) the size of this maximum shear stress is hardly changed from $0.3p_o$ where p_o is the maximum Hertzian contact stress (the ratio of p_o to the mean pressure $\bar{p}$ considered in section 4.1 is 3/2 for a point contact and $4/\pi$ for a line contact). At the same time a second region of large maximum stress develops at the surface. This is shown in Figure 13, in units of μp_o , for a Hertzian line contact. The two surface shear distributions, 26a and b, have been chosen to have the same mean value, in figure 13a; the magnitudes of the corresponding surface maximum shear stresses, figure 13b, have been calculated from information in (25). The constant surface traction distribution gives rise to an infinite maximum shear τ_{max} at $x/a = \pm 1$; it will therefore cause yielding at any value of p_o. However this is

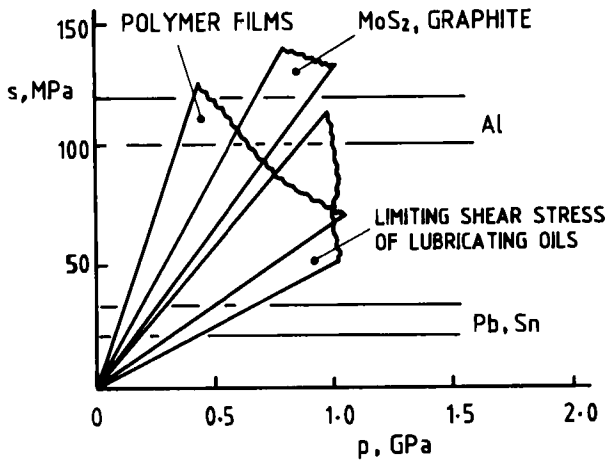


Figure 12. Contact pressure dependence of shear strength for a range of materials

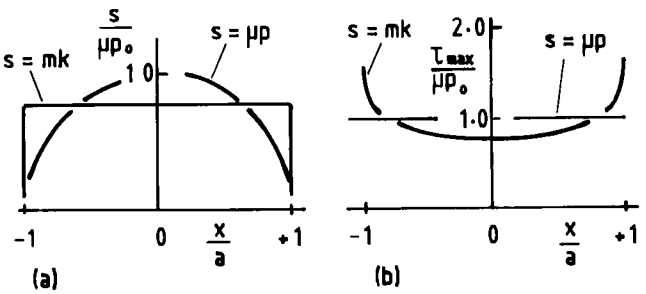


Figure 13. (a) surface tractions of equation 26 and (b) corresponding maximum surface shear stresses

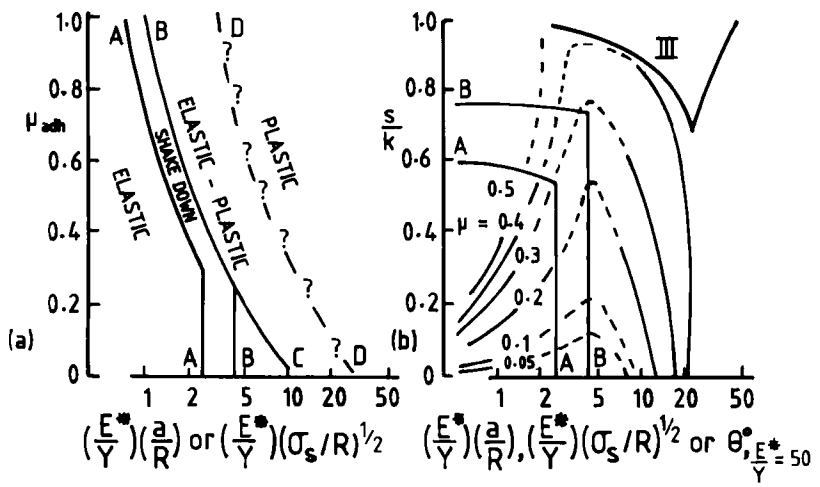


Figure 14. Flow mechanism maps, (b) with contours of constant friction coefficient

extremely localised. It may be important for the wear of contacts but will be ignored here in the context of friction. It will be assumed that an adequate approximation of both the τ_{max} distributions of figure 13b is

$$\tau_{max} = \mu p_o \qquad (27)$$

4.3.3. Sliding contact stress state mapping.

The maximum shear stress of equation 27 modifies the values of the plasticity index at which transitions from elastic to plastic stressing occur. Johnson (11) has provided a map showing how the stress state depends on μ and p_o/k in a single Hertzian contact (where p_o is the maximum Hertz stress as before). It may be shown that p_o/k is approximately equal to $(E^*/Y)(a/R)$, and we have already discussed the equivalence of $(E^*/Y)(a/R)$ and plasticity index, so his map may be directly used with a change of scale to produce figure 14a. The $(E^*/Y)(a/R)$ scale has been made logarithmic. To the left of line AA, deformation is always elastic. Repeated sliding in conditions just to the right of AA induces residual stresses that displace the elastic boundary to BB (shakedown limit). Between the lines BB and BC sub-surface plastic strain occurs but it is constrained to an elastic order of magnitude. To the right of BC plastic strain is still constrained elastically but it occurs at the surface. The position of the boundary DD above which flow is fully plastic is not known with any certainty, except that at $\mu=0$ it passes through $(E^*/Y)(a/R)=30$; it has been drawn arbitrarily parallel to BC.

4.3.4. Friction mechanism mapping.

Within the elastic region, the deformation or ploughing component of friction is zero, apart from any elastic hysteresis. A relationship may be developed between the effective value of s/k and μ, in order to transpose the elastic boundary of the map of figure 14a on to the scale of figures 7a or 9a, in order in turn to develop a unified map for combined adhesive and deformation friction, taking elastic deformation into consideration. Equation 28 is a series of manipulations: the first is a rewriting of s/k to introduce the average contact stress $\bar{p}$; the second recognises that $s/\bar{p}$ is the adhesive friction coefficient and that $\bar{p}/k$ is cp_o/k (where c is 2/3 for a point contact or $\pi/4$ for a line contact); finally p_o/k has been replaced by $(E^*/Y)(a/R)$ as previously introduced in section 4.3.3.

$$s/k \equiv (s/\bar{p})(\bar{p}/k) \equiv c\mu(p_o/k) \equiv c\mu(E^*/Y)(a/R) \qquad (28)$$

Equation 28 has been used to transform the boundaries AA and BB from figure 14a on to a map of s/k versus $(E^*/Y)(a/R)$. Figure 14b is the result. The lines AA and BB from figure 14a are drawn for a line contact, $c=\pi/4$. The region III marked on the map is the region of prow formation from figure 9a, for the particular case of $(E^*/Y)=50$. The reason for this choice of (E^*/Y) is introduced later. Thus figure 14b is a combined elastic and plastic friction mechanism map, with elastic adhesive friction in the region round the origin bounded by AA, prow or wedge formation in the region III, and elastic-plastic to plastic wave flows between AA/BB and III. At higher values of slope parameter, abrasive machining flows commence.

Contours of constant friction coefficient have also been added to figure 14b. Within the elastic region, μ is an adhesive friction coefficient and rearrangement of equation 28 gives

$$\mu = (s/k)(1/c)\ [(E^*/Y)(a/R)]^{-1} \qquad (29)$$

Contours are drawn for μ from 0.05 to 0.5. In the plastic region μ depends on combined adhesive and deformation components and as already considered in section 3.2 depends on s/k and mean surface slope, either a/R, $\tan\theta$ or $(\sigma_s/R)^{0.5}$ as appropriate. In order to map μ values from figure 9 on to the scale of figure 14b, a value must be selected for (E^*/Y). For the special case of sliding steel on steel (E^*/Y) may be close to 50. For that value, the scale of $(E^*/Y)(a/R)$ of figure 14b is then numerically approximately the same as that of surface slope in degrees (1 radian is 57°), for the range of slopes shown. Lines of $\mu = 0.2$, 0.3 and 0.4 in the plastic state have been added to figure 14b for this case. Finally the elastic and plastic friction coefficient contours have been extended into the elastic-plastic region on the basis that they should meet one another smoothly: the lines are broken to indicate that their detailed positions are speculative in that region.

4.3.5. Commentry on the friction mechanism map.

The map (figure 14b) may be discussed in terms of three friction coefficient magnitudes: $\mu < 0.3$, $0.3 < \mu < 0.4$ and $\mu > 0.4$. For $\mu < 0.3$, surface deformation will either be elastic, elastic-plastic with sub-surface constrained plastic deformation or a plastic wave flow. In the transition range of μ between 0.3 and 0.4 there is an increasing danger, as s/k and surface slope increase, that plastic flow may concentrate on the surface and verge on wedge or prow formation. If μ should become greater than 0.4, there is a possibility that this is because of extreme surface smoothness in an elastic contact state, but if the elastic limit should be exceeded there is a rapid transition to destructive plastic wedge or machining flows. It must therefore be prudent to regard $\mu = 0.3$ as an upper value for long-lived components, quite apart from considerations of energy losses in sliding.

The map may also be used to consider minimum attainable friction coefficients in the absence of fluid film lubrication. For example, to achieve $\mu < 0.05$ for steel sliding on steel certainly needs a combined surface roughness slope less than between 5° and 10°, which is not difficult to achieve (see Table 2, where slopes $(\sigma_s/R)^{0.5}$ are in radians), but it also requires s/k to be less than 0.1. In section 4.3.1 the nature of surface shear strength was briefly discussed. Figure 12 shows that most classes of material surface contaminants seem to have a shear strength that increases with normal stress, in this case the contact stress. A sliding surface which has run-in to the region AA-BB of figure 14b will have a mean real contact stress of about 2k. There is thus a lower limit to s/k which is twice the gradient of the s/p data in figure 12. That is 0.1 for the non-metal data in that figure. For the metal data only Pb or Sn films on the hardest substrates appear to offer low enough shear strengths for $\mu < 0.05$ (this example is offered as illustrative: it is not a recommendation for a practical solution to low friction coatings). It is a matter of continuing research to understand the rheology of thin surface layers, with the objective of

constructing new molecules with low shear strength yet load carrying and surface protective capacity.

5. MECHANISMS AND MAGNITUDES OF WEAR

The flow regime boundaries of figures 9a and 14 can also serve as foundations on which to build a wear mechanism map. Figure 15 is such a map. Two carefully made bearing surfaces might well have a plasticity index of around 5 to 20 (see Table 2) when first assembled; and be lubricated to have $s/k < 0.5$. This would place them in the 'initial wear' region of figure 15. Sliding causes surface changes. Entrapment or embedding of hard wear debris or contaminating grits could cause a surface roughening and a change of abrasive wear. Breakdown of lubrication could cause an increase of s/k, to cause a change to severe adhesive wear. Alternatively surface smoothing (running-in) could occur. Two distinctions have been made, between running-in to a high surface shear strength or friction coefficient level, when wear by metal surface fatigue or delamination has been observed; and to a low shear strength level, when conditions described as mild wear can be established. This last is the onlny acceptable wear state for long-lived engine components. The wear coefficients of all the wear states listed in figure 15 are recorded in figure 4. As a rough example, from equation 11, if a piston ring of hardness 2 GPA, operating at a nominal contact pressure of 20 MPa, with a stroke of 100 mm, were to operate entirely in a boundary lubrication regime with $K = 10^{-8}$ it would lose 10 μm in depth every million engine revolutions: thus 'low' wear rates are put in perspective!

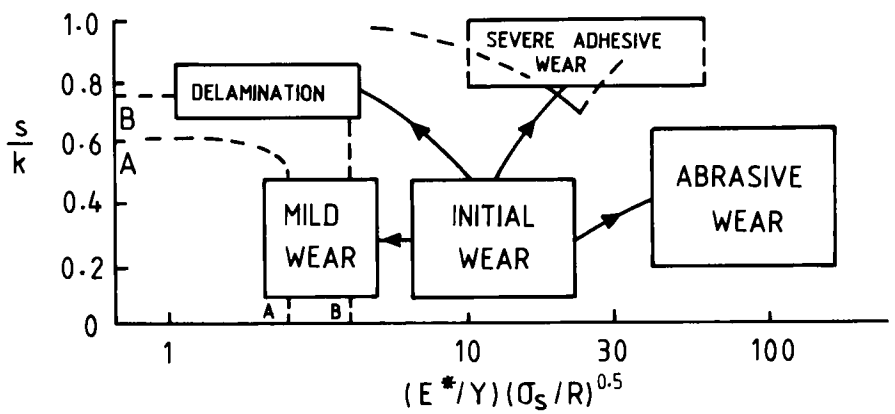


Figure 15. A wear mechanism map

The calculation above was made after dividing both sides of equation 11 by nominal contact area, thus converting wear volume V to wear depth s and load W to nominal contact pressure P, as recorded in equation 30a. A further refinement is to differentiate

with respect to time to create equation 30b in which U is sliding speed. Another manipulation of equation 11 is to multiply the right hand side by μ/μ, as has been done in equation 30c. μWL is the friction work: k/μ can then be interpreted as wear volume per unit friction work input. All these alternative forms have their value in giving insight into numerical values of k or K.

$$s = k P L \quad (30a)$$
$$ds/dt = k P U \quad (30b)$$
$$V = (k/\mu)\mu WL \quad (30c)$$

In the remainder of this section further points are made specific to each of the wear regimes identified in figure 15.

5.1. Abrasive wear

Abrasive wear is the most severe form that there is (as indicated in figure 4). The term 2-body abrasion signifies abrasives embedded in one of the sliding surfaces; 3-body abrasion refers to abrasives free to move relative to both sliding surfaces. The first thing to write about either is that a surface is protected from abrasion if it is more than 1.3 times as hard as the abrasive, and that some protection is given even when its hardness is only 0.8 that of the abrasive. Table 5 lists the hardness of some common abrasives (3). One of the most common engine contaminants is silica sand: table 5 suggests some protection would be given against abrasion by sand by hardening a surface to 6.5 GPa while full resistance would be given by hardening to 10.5 GPa. This is a range that can be achieved by steels and ceramics.

Table 5. The hardness of some abrasives

Material	HV, GPa	Material	HV, GPA	Material	HV, GPA
Diamond	80	Alumina	21	Zirconia	11
CBN	40	Tungsten carbide	19	Silica sand	8
Silicon carbide	25	Garnet	13	Glass	5

When abrasion does occur the concept behind equation 10, that wear volume should equal the scratch volume, forms a good starting point for considering wear rates. One may modify equation 10 to

$$V = f K' (W/H)L \quad (31)$$

where K' is the shape factor which for a conical abrasive is $2\tan\theta/\pi$ and f is the fraction of the scratch volume that appears as wear. f is found to depend at least on the abrasive shape and the work hardening rate and the microstructure of the abraded material. Figure 16a shows a collection of f values measured as a function of abrasive slope in single scratch tests. The initial rapid variations of f with slope are the transitions from wave to cutting flows but the constant values of f at higher slopes vary from 1 to 0.4, with different materials and abradors. A similar range of f is seen when scratching materials of

a range of work hardening characteristics by a single abrasive (figure 16b). Hard carbides in a metal microstructure can depress f but rarely to less than 0.2. Thus as far as engine component wear is concerned the only safe procedure is to avoid abrasion by either filtration, other forms of wear particle entrappment or by suitable surface hardening.

5.2. Severe adhesive wear

Severe adhesive wear, in the form of scuffing or scoring, is the failure mode that all metal bearing surfaces have to be guarded against. If s/k should rise to a high value, wedge formation, in the case of ploughing modes of plastic flow, or the exhaustion of juntion growth, leads to the transfer of metal asperities from one surface to the other. These particles act as keys for further transfer and layers build up that eventually break down to cause wear particles. Early studies in dry conditions were particularly thorough. For example for the sliding of brass on a hard tool steel (26,27) transfer fragments were

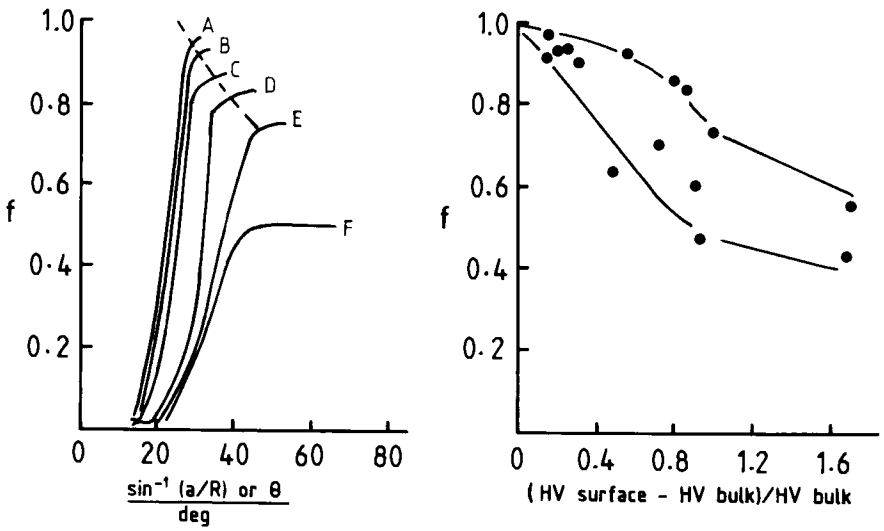


Figure 16. Variation of f with (a) slope and (b) work hardening, from (14)

typically 200 μm in diameter and 5 μm thick, while the wear debris could be 1 mm in diameter by 15 μm thick. Specific wear rates were 10^{-5} to 10^{-6} mm ^{2}/N.

It is essential that s/k does not reach high values in practical bearings. In the absence of special surface treatments, the natural lubricant that any metal (other than gold) has is its oxide skin. Wherever possible sliding components should be chosen to be harder than their oxides. For example severe wear is common at room temperature with steels of Vickers Hardness less than 5 GPa, but can be suppressed by hardening them to greater than this: the hardness of iron oxides is about 5 GPa (28). As a second line of defence sliding metals should be chosen to be insoluble in one another, to reduce the strength of any intermetallic junctions that should be formed. This is possibly why many of the plane bearing overlay metals are insoluble in iron(3).

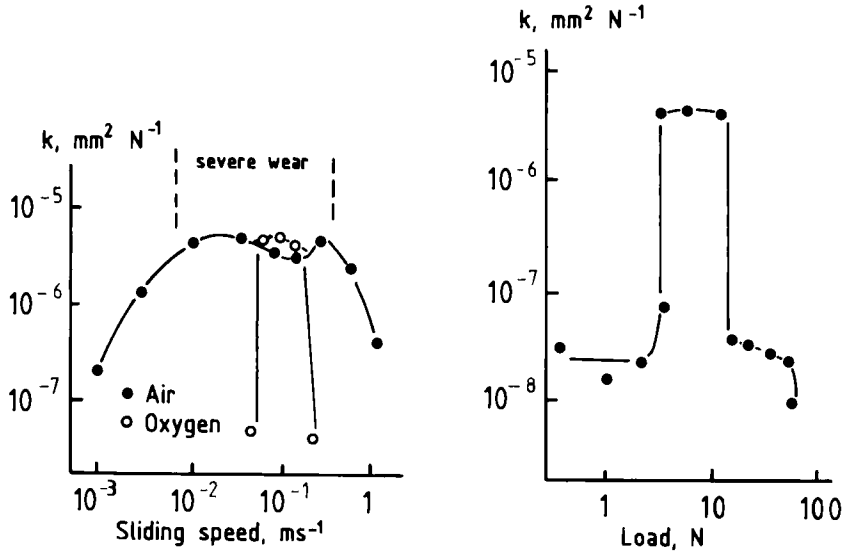


Figure 17. Mild to severe wear transitions (a) brass on steel, (b) steel on steel

Even if a protective surface film can exist on a metal, whether it does prevent severe wear depends on a competition between its removal in the wear process and its rate of reformation in the environment in which it finds itself. Figure 17 shows variations of specific wear rates with changes of load, sliding speed and environment for brass on steel and steel on steel in the dry sliding conditions already mentioned in the last paragraph. At low loads and speeds protective oxides form more quickly than they are removed. At high loads and speeds friction heating, causing accelerated oxidation (and for steel on steel surface hardening too), again results in oxide reformation countering its removal. Only in intermediate conditions does severe wear occur. The transitions to severe wear can occur very suddenly with change of operating conditions. Suppression of severe conditions by film formation from lubricants is a major development activity of oil companies which will be returned to in a later lecture.

5.3. Delamination wear

Both abrasive and severe wear definitely involve fully plastic flow. It is possible to run-in surfaces to a near elastic state in high friction conditions ($0.5 < \mu < 1.0$). Repeated loading of the contact results in fatigue crack growth and eventually results in wear. This mechanism was first observed by Suh and his co-workers (29). Wear debris was in the form of plates 1 to 10 μm thick, specific wear rates for steels were between $4x10^{-7}$ to $4x10^{-8}$ mm^2/N. These wear rates are still too high to be acceptable in sliding contacts. The study of crack growth in these circumstances is of more importance to the failure of sliding/rolling contacts such as occur in power transmission gearing and driving or braking rolling contacts. The study of the elastic-plastic strain increment per loading cycle is a matter of current research (30).

5.4. Mild wear

If surfaces can be induced to run-in to a near elastic state in low s/k conditions, then usefully low wear rates can be achieved. Although there have been mechanical analyses of elastic-plastic crack growth in these conditions too (31) it is not clear whether they are relevant to the study of mild wear as it is usually observed that surface removal does not involve the bulk material but the breakdown of the surface films that are responsible for the low s/k. Conditions of dry wear, when wear occurs by removal of oxide films, may be distinguished from boundary lubricated wear when surface films formed by reaction with oil additives are involved. In dry wear there is no single mechanism of oxidative wear. In the high load mild wear conditions of figure 17b a continual cycling between metal transfer, oxidation and removal of oxide was observed. In the high speed conditions of figure 17a a mixed oxide was continually formed between the brass and steel and wear occurred by the steady removal of that oxide: the steel wore at the same rate as the brass. In other circumstances thick oxide scales grow and flake off (32) and in other circumstances again oxide debris can be compacted to form a glaze on the metal surface, the wear finally coming from that layer (33). This wide range of mechanisms gives rise to a wide range of specific wear rates, from 10^{-7} to 10^{-9} mm^2/N. An even wider range of specific wear rates can be obtained by reaction film removal in boundary lubricated conditions, down to 10^{-13} mm^2/N (5,34).

6. CONCLUSION

The chapter has surveyed the different mechanisms of the sliding friction and wear mainly of metals. It has emphasised the importance of surface finish (surface slope) and interfacial shear strength (s/k) in controlling the severity of friction and wear. A key quantity is the plasticity index, the prodct of (E*/Y) and surface slope. Provided $s/k < 0.5$, a plasticity index < 2.5 ensures a predominantly elastic response of surfaces in contact. In the absence of fluid film lubrication, minimum sliding friction coefficients seem to be about 0.05 to 0.1: there is good reason to believe from rheological studies of surface films that there is a lower bound to the friction coefficient between sliders in solid contact and that it may be about 0.05. Minimum values of specific wear rate that have been observed are about 10^{-13} mm^2/N. Although these numbers have been produced in the context of metals sliding on metals, there is no reason to expect lower values for the sliding of metals on ceramics or of ceramics on ceramics (an extension of the lecture's content to the sliding of ceramics must await another time). The engine tribologist should for the time being assume these values as the minimum attainable and design components accordingly.

REFERENCES

1. Bowden F.P. and Tabor D. (1950) The Friction and Lubrication of Solids. Vol.I. Ch.5. Clarendon Press, Oxford
2. Bowden F.P. and Tabor D. (1964) The Friction and Lubrication of Solids. Vol.II. Chs.16,17. Clarendon Press, Oxford

3. Rabinowicz E. (1965) Friction and Wear of Materials. Chs.6,7. John Wiley, New York
4. Archard J.F. and Hirst W. (1956) "The Wear of Metals under Unlubricated Conditions", Proc.Roy.Soc. A236, 397-410.
5. Childs T.H.C. (1980) "The sliding wear mechanisms of metals, mainly steels". Tribology International, 13, 285-293.
6. Childs T.H.C. (1970) "The sliding of rigid cones over metals in high adhesion conditions", Int.J.Mech.Sci., 12, 393-403
7. Kayaba et al (1986) "Analysis of the abrasive wear mechanism by successive observations of wear processes in a scanning electron microscope", Wear 110, 419-430
8. Hill R. (1950) The Mathematical Theory of Plasticity Chs.6-8, Clarendon Press, Oxford
9. Johnson W. and Mellor P.B. (1983) Engineering Plasticity Ch.12, Ellis Harwood, Chichester
10. Challen J.M. and Oxley P.L.B. (1979) "An explanation of the different regimes of friction and wear using asperity deformation models", Wear 53, 229-243
11. Johnson K.L. (1985) Contact Mechanics Ch.7, Cambridge University Press, Cambridge
12. Lee E.W. and Shaffer B.W. (1951) "Theory of plasticity applied to a problem of machining", Trans ASME, J.Appl.Mech. 18, 405-413
13. Petryk H. (1987) "Slipline fields of sliding contact", Proc.Conf.Friction, Lubrication and Wear - fifty years on, pp 987-994, Mechanical Engineering Publications, I.Mech.E. London
14. Childs T.H.C. (1988) "The mapping of metallic sliding wear". Proc.I.Mech.E. 202Ptc, 379-395.
15. Johnson K.L. (1968) "Deformation of a plastic wedge by a rigid flat die under the action of a tangential force". J.Mech.Phys. Solids 16, 395-402
16. Collins I.F. (1980) "Geometrically self similar deformations of a plastic wedge under combined shear and compression by a rough flat die". Int.J.Mech.Sci. 22, 735-742
17. Cocks M. (1966) "Shearing of junctions between metal surfaces". Wear 9, 320-328
18. Johson K.L. (1985) Contact Mechanics Chs.5-7, Cambridge University Press, Cambridge
19. Greenwood J.A. and Williamson J.B.P. (1966) "Contact of nominally flat surfaces", Proc.Roy.Soc. A295, 300-319
20. Whitehouse D.J. and Archard J.F. (1970) "The properties of random surfaces of significance in their contact". Proc.Roy.Soc. A316, 97-121
21. Archard J.F. (1974) "Surface topography and tribology", Tribology Int. 7, 213-220
22. Briscoe B.J. and Smith A.C. (1982) "The interfacial shear strength of molybdenum disulphide and graphite films", Trans. ASLE 25, 349-354
23. Briscoe B.J. in Friedrich K. (ed) (1986) Friction and Wear of Polymer Based Composites. Ch.2. Elsevier, Amsterdam
24. Evans C.R. and Johnson K.L. (1986) "The rheological properties of elastohydrodynamic lubricants". Proc.I.Mech.E. 200 Ptc. 303-312
25. as reference 11, but chapters 2 and 7.

26. Kerridge M. and Lancaster J.K. (1956) 'The stages in the process of severe metallic wear'. Proc.Roy.Soc. A236, 250-264
27. Hirst W. and Lancaster J.K. (1960) 'The influence of speed on metallic wear'. Proc.Roy.Soc. A259, 228-241
28. Welsh N.C. (1965) 'The dry wear of steels Parts I and II'. Phil.Trans.Roy.Soc. A257, 31-70
29. Suh N.P. et al. (1974) 'Further investigation of the delamination theory of wear'. Trans.ASME F96, 631-637
30. Bower A.F. and Johnson K.L. (1991) 'Plastic flow and shakedown of the rail surface in repeated wheel-rail contact'. Wear 144, 1-18
31. Hearle A.D and Johnson K.L. (1987) 'Cumulative plastic flow in rolling and sliding line contact'. Trans.ASME Jnl. Appl. Mech. 54, 1-5
32. Quinn T.F.J. et al. (1980) 'Developments in the oxidational theory of mild wear'. Tribology Int. 13, 153-158
33. Stott F.H. and Wood G.C. (1978) 'The influence of oxides in the friction and wear of alloys'. Tribology Int. 11, 211-218
34. Childs T.H.C. and Sabbagh F. (1989) 'Boundary lubricated wear of cast irons to simulate automotive piston ring wear rates'. Wear 134, 81-88.

Engine Tribology / C.M. Taylor (Editor)

LUBRICATION REGIMES AND THE INTERNAL COMBUSTION ENGINE

C M TAYLOR

Department of Mechanical Engineering, University of Leeds, Leeds LS2 9JT

1. INTRODUCTION

The continuing trend for automobile engines to have a higher output and yet at the same time to be more compact places considerable pressure upon the component designers. There is a requirement for more sensitive design predictions in order to achieve an improved specification. This is certainly true of the main frictional components of the engine, the bearings, the valve train and the piston assembly. It is also important that any improvement in the ability to operate in a more severe environment of load, speed or temperature, should not compromise the durability of any machine element.

In order to achieve a better design of the principal tribological devices of the internal combustion engine, it is vital to have a clear understanding of the major physical and chemical mechanisms responsible for the load carrying ability and failure mechanisms. The contribution will concentrate upon the lubrication regimes (or different modes of lubrication) which lead to the ability of lubricated, sliding/rolling contacts to sustain an applied load. These regimes will be put in a generalised, rationalised context and then related to the automobile engine.

It will be shown that different modes of lubrication are significant in the alternative tribological components of the engine and that a component may enjoy a range of lubrication regimes during a cycle. Since the importance of the physical and chemical properties of the lubricant and its additive package are different for alternative modes of lubrication, it will be recognised that critical compromises have to be made. A single oil is used to lubricate bearings, valve train and piston assembly and yet different properties will be desirable to optimise the performance of each.

This emphasises the significance of the lubricant as an engineering design material.

2. MODES OF LUBRICATION

2.1 Fluid Film Lubrication

If the lubricant film is sufficiently thick to prevent the opposing solids from coming into contact the condition is described as 'fluid film lubrication' and

the frictional characteristics of the combination can be explained within the framework of conventional fluid mechanics.

This condition is often referred to as the ideal form of lubrication since it provides low friction and a high resistance to wear. The behaviour of the contact is governed by the bulk physical properties of the lubricant, notably viscosity, and the frictional characteristics arise purely from the shearing of the viscous lubricant.

2.2. Boundary Lubrication

If the surfaces of the solids are not separated by the lubricant and contact takes place over an area comparable to that which develops in dry contact, the condition is called 'boundary lubrication', since the frictional characteristics are determined by the properties of the solids and the lubricant at their common interfaces.

In this case it is the physical and chemical properties of thin surface films of molecular proportions which govern the contact characteristics. The properties of the bulk lubricant are of minor importance and the coefficient of friction is essentially independent of viscosity. The surface action which determines the behaviour of boundary lubricants can be described in order of 'film strength' in the following terms:

a) Physically adsorbed layers of gaseous, liquid or solid lubricants.

b) Chemically adsorbed layers.

c) Films formed by chemical reaction.

2.2.1. Physically Adsorbed Layers

These are formed by short range inter-molecular forces and the process is reversible. A layer of lubricant one or more molecules thick becomes attached to the surfaces of the solids and this provides a modest protection against wear. The limit of 'film strength' is denoted by a transition temperature which is close to the 'melting point' of the lubricant.

An interesting feature of physically adsorbed layers is that they are essential to 'fluid film' lubrication. In fluid mechanics we assume that the fluid adjacent to a solid moves with the velocity of the solid - the condition of 'no-slip'.

Chemisorbed films offer moderate protection or film strength whilst films formed by chemical reaction exhibit the highest degree of surface protection.

The most common lubricant, mineral oil, consists of a large number of different hydrocarbon molecules. Most of these molecules are saturated hydrocarbons having chain or ring structures which are in equilibrium.

The molecules offer little protection to the solid bearing surfaces in the absence of fluid film lubrication. They consist of paraffinic, naphthenic and aromatic groups linked together to form complex molecules.

Mineral oils also contain small but important quantities of more reactive substances like sulphur, nitrogen and oxygen. These hetero-atoms readily react with newly exposed metal surfaces and valuable boundary lubricating properties can result.

It is usual to add small quantities (up to about 1-2%) of substances to base lubricants like mineral oils if chemisorbed or reaction films are required to provide extra protection for the bearing surfaces. These substances are called 'additives'.

2.2.2. Chemisorbed Films

These are generally produced by adding animal and vegetable fats and oils to the base oils. These additives contain long chain fatty acid molecules which exhibit great affinity for metals at their active ends. The usual representation of these polar molecules is like that of a carpet pile with the molecules standing perpendicular to the surface. These fatty acid molecules form metal soaps which are low shear strength materials giving rise to coefficients of friction in the range 0.1 to 0.15. The soap film has the thickness of one molecule (about 2.5nm) but it is very dense because of the preferred orientation of the molecules. For example, stearic acid will from a mono-molecular layer of the soap iron stearate containing about 10^{14} molecules per square centimetre of surface. The effectiveness of these films is limited by the melting point of the soap ($180^{o}C$ for iron stearate). It is essential to choose an additive which will react with the bearing metals. For example, inert metals like gold and platinum are not effectively lubricated by fatty acids.

Examples of fatty acid additives are stearic, oleic and lauric acid. The soap films formed by these acids might reduce the coefficient of friction to 50% of that obtained by a straight mineral oil. They provide satisfactory boundary lubrication at moderate loads, temperatures and speeds and are often successful in situations showing evidence of mild surface distress.

2.2.3. Films Formed By Chemical Reaction

These provide the greatest film-strength and are used in the most severe operating conditions. If the loads and sliding speeds are high, significant contact temperatures will be developed. It has already been noted that films formed by physical and chemical adsorption cease to be effective above certain transitional temperatures, but some additives start to react and form new high melting point inorganic solids at high temperatures. For example sulphur will start to react at about $100^{o}C$ to form sulphides with melting points in excess of $1000^{o}C$. Lubricants containing additives like sulphur, chlorine, phosphorus and zinc are often referred to as extreme pressure (EP) lubricants since they are effective in the most arduous conditions.

It is worth noting that oxygen from the atmosphere or free oxygen in the bulk lubricant is a valuable EP lubricant. It forms oxide layers which generally provide a low shear strength surface film capable of reducing friction and wear between bearing materials.

Increasing use is being made of pre-treated surfaces in bearings. Thin films of low shear strength material are deposited on hard backings for plain bearings. Polymers like PTFE, are used as dry bearings in solid or impregnated form.

Soaps and fats are often applied to the surfaces or metals before forming to improve die life (eg wire drawing). The pre-treatment of metal in a chemical bath to produce a surface layer of inorganic salt is used for both corrosion resistance and high film strength (eg phosphate coatings and sulphidising).

The coefficient of friction between chemically clean and dry metals is extremely high (at least unity). Oxidation of the surfaces normally brings the coefficient of dry friction down to about 0.5. The application of a straight mineral oil to the surfaces effects a further reduction to about 0.15 whilst a good boundary lubricant might bring the value down to about 0.08.

2.3. Mixed Lubrication

In the past it has been adequate for engineers to distinguish between the well known 'fluid film' and 'boundary' lubrication regimes, but it is clear that there is an important 'mixed' lubrication regime between the two. In full fluid lubrication separation of the opposing solids is complete. In boundary lubrication the laws of dry friction apply since the coefficient of friction is independent of load, speed and apparent area of contact. The actual value of the friction coefficient does, of course, depend upon the nature of boundary lubricant.

The transition from boundary to fluid film lubrication does not take place instantaneously as the severity of loading is decreased, but an increasing proportion of the load is carried by pressures within the fluid which fills most of the space between opposing solids. Indeed, it is often very difficult to eliminate fluid film lubrication effects to enable true boundary lubrication experiments to be performed, and there is evidence to suggest that micro-fluid film lubrication associated with a number of miniature bearings formed by surface irregularities is an important effect in some processes.

In the mixed lubrication regime the contact characteristics are determined by varying combinations of fluid film and boundary lubrication effects. Physical properties of the bulk lubricant and chemical properties of the boundary lubricant are all important. This is an important lubrication condition since many machine elements move intermittently and hence operate in the 'boundary', 'mixed' and 'fluid-film' regimes at different times (eg plain bearings during start up and shut down conditions).

2.4. Elastohydrodynamic Lubrication

It has now been demonstrated convincingly that effective hydrodynamic films often persist well into the region originally described as boundary lubrication.

This action occurs as a result of local elastic deformation of the bearing materials and it is thus most readily generated in highly stressed nominal point or line contacts like gears, rolling bearings and cams and tappets. In these situations the viscosity of the lubricant rises very considerably due to the high contact pressures and this further assists the formation of effective fluid films. The viscosity in a typical gear or rolling bearing contact may rise to several thousand times the atmospheric viscosity of the lubricant.

In a sense the recent development of understanding of elastohydrodynamic lubrication has tidied up our view of lubrication regimes. It has enabled a more precise definition to be given to boundary lubrication, but it has not removed the need to define a mixed lubrication regime. It is now clear that elastohydrodynamics plays an important part in the behaviour of conventional plain bearings under high loads and the subject of micro-elastohydrodynamic lubrication, which is concerned with fluid film lubrication in the vicinity of elastically deformed asperities, has done much to explain the apparent delay in the onset of true boundary lubrication under some operating conditions. It is important to note that elastohydrodynamic lubrication is a form of fluid film lubrication.

2.5. Surface Quality

The form of lubrication encountered in a machine is determined by the lubricant properties, load, speed, temperature overall geometry or bearing surfaces and surface quality. The load, speed and ambient temperature are often fixed parameters whereas the designer can often exercise his discretion over the other quantities. Once the lubricant has been selected and the overall geometry fixed the final factor which determines the mode of lubrication is surface quality.

Most bearing surfaces are turned or ground and sometimes polished and it is useful to erect a scale of typical dimensions as shown in Figure 1. Very great care has to be exercised if surfaces are to be produced with R_a values of a fraction of a micrometre and it is clearly undesirable to spend a great deal of time and effort making a surface better than it needs to be for successful operation. Indeed the running in process in which asperities are reduced in height without damage to the underlying material is a very good final treatment for most bearing surfaces. The process is often aided by pre-treatment of the surfaces or the addition of EP or other boundary lubricants during the initial stages of running.

Rolling bearings usually have rolling elements with a surface roughness (R_a) in the range 0.05 to 0.13μm, whilst the roughness of the tracks varies from 0.2 to 0.3μm. Plain bearing surfaces are usually rougher, in the range 0.25 to 1.3μm, whilst gears have similar surface qualities.

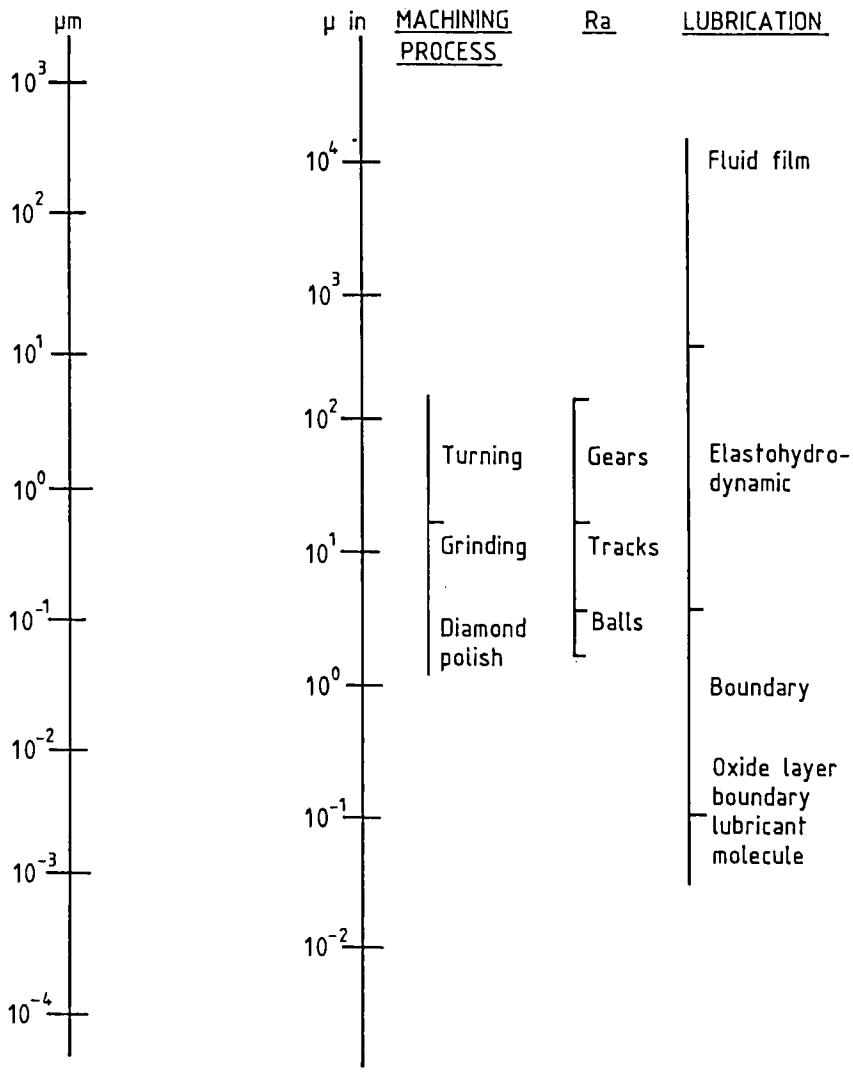

µm
10^3
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
µ in
10^4
10^3
10^2
10^1
10^0
10^{-1}
10^{-2}
MACHINING PROCESS
Turning
Grinding
Diamond polish
Ra
Gears
Tracks
Balls
LUBRICATION
Fluid film
Elastohydro-dynamic
Boundary
Oxide layer boundary lubricant molecule
TYPICAL DIMENSIONS

Figure 1

2.6. Transition From Fluid Film To Boundary (or mixed) Lubrication

It is convenient to relate transition from one lubrication regime to another to a dimensionless film thickness ratio defined as -

$$\text{Film thickness ratio} = \frac{\text{effective film thickness}}{\text{surface roughnesses } (R_a)}$$

2.6.1. Fluid Film Lubrication

The lubricant films are normally many times thicker than the surface roughness as measured on the R_a scale. The physical properties of the lubricant dictate contact behaviour and surface effects are entirely negligible.

The film thickness normally exceeds 1.0μm and the film thickness ratio is in excess of 10 and may even rise to 100. Films of this thickness are clearly insensitive to chemical action in surface layers of molecular proportions.

2.6.2. Elastohydrodynamic Lubrication

Local elastic distortion of the solids provides coherent fluid films and asperity interaction is prevented. For normal engineering contacts the film thickness is between 0.1μm and 1.0μm and the film thickness ratio is greater than 4 or 5.

2.6.3. Mixed Lubrication

The behaviour of the contact is governed by a mixture of 'boundary' and 'fluid film' effects. Some asperity contact may occur and interaction takes place between mono or multi-molecular layers of boundary lubricating films whilst a partial fluid-film lubrication action develops in the bulk of the space between the solids.

If the effective film thickness is calculated on the basis of normal fluid film or elastohydrodynamic action between smooth surface it is found to be in the range 0.05μm to 1μm with film thickness ratios between 1 and 5.

2.6.4. Boundary Lubrication

Surface interaction between mono or multi-molecular layers of boundary lubricants and the solids dominates the operation of the contact. Fluid film effects are negligible and there is considerable asperity contact. The frictional behaviour is similar to that encountered in dry friction between solids.

The surface films vary in thickness from 5nm to 10nm and the film thickness ratio is less than unity.

The lubrication regimes discussed above are illustrated in Figure 2 and the relationship between coefficient of friction and film thickness ratio is shown in Figure 3.

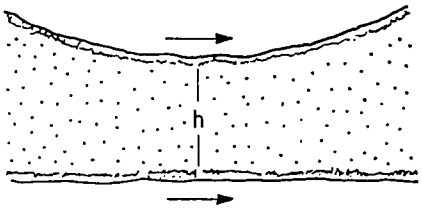
3. THE FRICTIONAL ELEMENTS OF THE INTERNAL COMBUSTION ENGINE

The internal combustion engine has served to power the motor car for a century. In energy efficiency terms automobile engineers would not claim to have been overly successful. Figure 4 shows the typical distribution of fuel energy. In the broadest terms, some 60% of the energy is uselessly dissipated in the form of heat to the environment, either from the surfaces of the engine and it auxiliaries, or down the exhaust pipe. A further 15% or so may be lost in mechanical actions, leaving about 25% of the original energy in terms of brake power available to the transmission and hence, after further significant losses, to supply traction at the wheels. The distribution of energy may show substantial differences from that indicated depending upon the type of engine, the load condition and specific operational situation. For example, whilst a vehicle is stationary at the traffic lights the whole of the fuel energy is essentially wasted.

An ideal engine operating on the Carnot cycle would, according to the Laws of Thermodynamics, achieve a maximum efficiency of about 85% with typical temperatures for heat generation and rejection. The low heat rejection and high efficiency cannot be realised in practice for a number of reasons. The substantial thermal losses associated with the automobile engine have quite naturally dictated that attempts at improvement since the energy crisis of nearly two decades ago have focused in this area. These developments have been wide ranging in nature, encompassing on the one hand combustion considerations and on the other low heat rejection engines. Maximum brake efficiencies of over 45% have been mooted, an increase of less than 10% over the efficient automotive diesel engine currently available. It was widely recognised around 1980 that despite the large energy losses associated with thermal aspects, the return on research investment associated with attempts at improvement were limited. Worthwhile gains were potentially to be had more effectively by consideration of more modest areas of loss. Attention was directed towards mechanical losses, a typical distribution of these being shown in Figure 5.

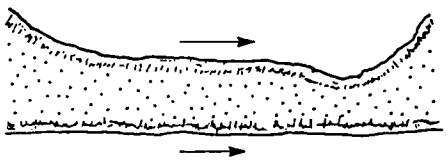
Some 80% of mechanical losses are generally accepted to be associated with the main frictional components: the engine bearings, the valve train and the piston assemblies, with the latter often credited with about half the total. The auxiliary losses, including pumping of fluids, are not small at about 20%. Indeed it is not unknown for an oil pump at limiting engine speeds to be less than 10% efficient and consume 2 to 3 kW. Assessment of the losses of the frictional pairs of an engine involves fluid film lubrication analysis, together with a range of other tribological aspects. The reliability and durability of the components must not be degraded in efforts to reduce losses.

Indeed, the need to make analyses more sensitive in order to improve durability whilst increasing rating is an important area of development. Lubrication regimes will be reviewed against this particular backcloth.



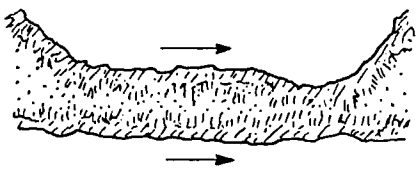
a Hydrodynamic

$h > 1\mu m$

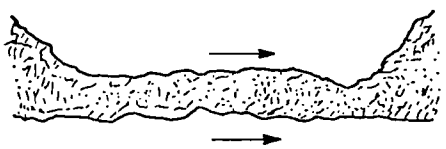


b Elastohydrodynamic

$h \sim 0.1 - 1\mu m$

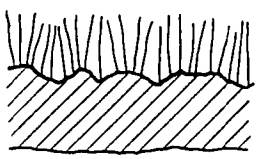


c Mixed



d Boundary

Typical dimensions of surface films

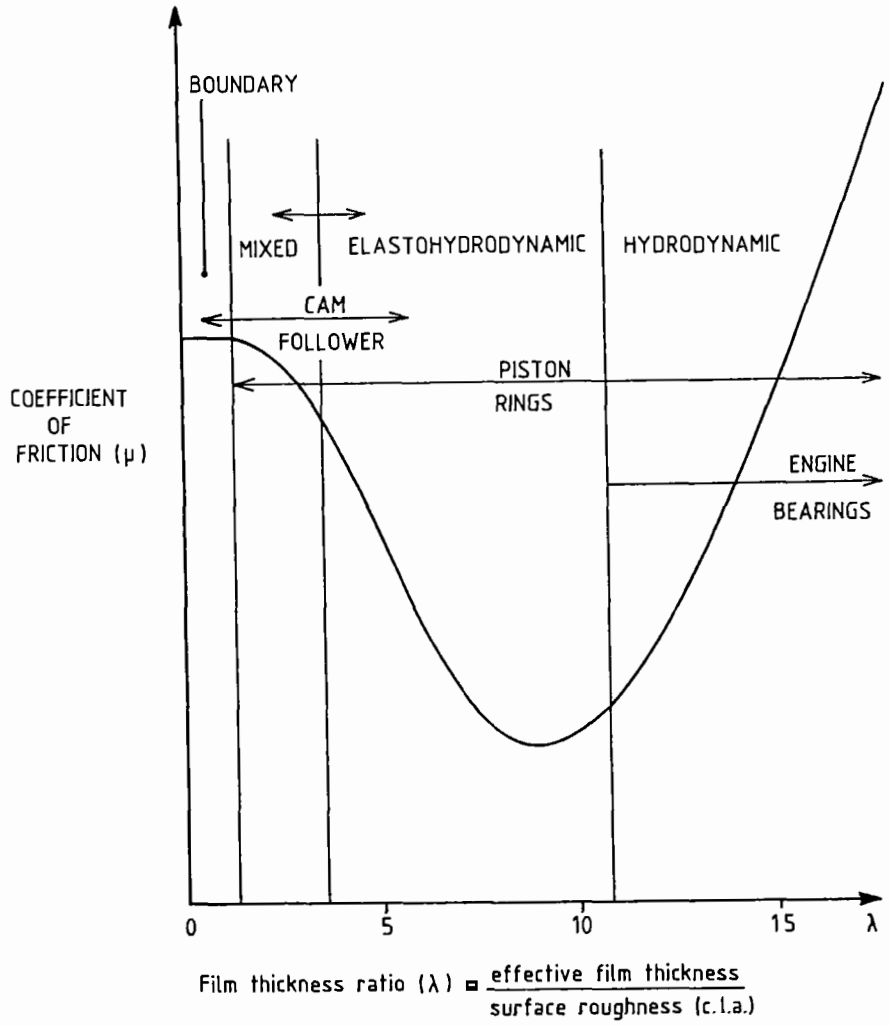


30 Å boundary lubricant molecule (3nm)

10 - 50 Å oxide layer (1 - 5nm)

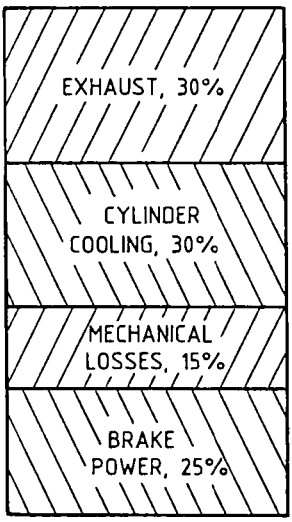
LUBRICATION REGIMES

Figure 2



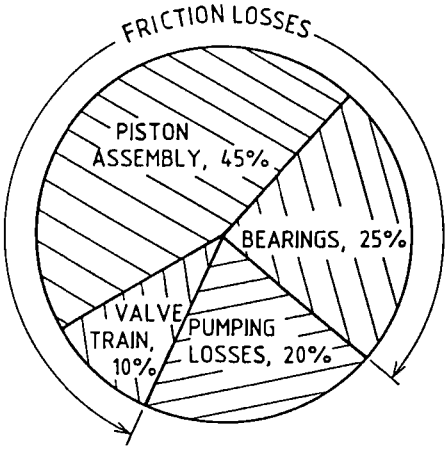
VARIATION OF COEFFICIENT OF FRICTION WITH FILM THICKNESS RATIO

Figure 3



TYPICAL DISSIPATION OF THE FUEL ENERGY IN AN INTERNAL COMBUSTION ENGINE

Figure 4



TYPICAL DISTRIBUTION OF MECHANICAL LOSSES IN AN INTERNAL COMBUSTION ENGINE

Figure 5

Figure 3, sometimes known as the Stribeck curve, has been introduced as a means of rationalising the alternative modes of lubrication which have been identified in terms of a map or diagram. It is possible to relate the lubrication characteristics of engine bearings, valve train (cam/follower) and piston rings to the various regions of the diagram. Since the components are subjected to dynamic circumstances, the mode of lubrication may vary during a single cycle.

3.1. Engine Bearings

In the design context the bearings are taken to operate with hydrodynamic lubrication. It is recognised that at start-up and shut down some element of surface interaction may occur and indeed that high speed rub situations may develop. Operational experience does, however, indicate that for the large majority of situations the journal and bearing surfaces are effectively separated. The engine bearings are seen to enjoy the most benign lubrication circumstances.

It is important that the bearing designer is not lulled into a false sense of security. In reality it is clear that effects of elastic distortion can be significant and that at least for some parts of the cycle the mode of lubrication would be better termed elastohydrodynamic. Indeed current predictions of minimum lubricant film thickness in automobile bearings are of the order of 1μm and increasing thought is being given to the influence of surface interaction phenomena. Design predictions of tribological behaviour for all the lubricated components of the engine are built upon a range of assumptions and can only be viewed as a benchmark in nature.

3.2. Cam And Follower

The cam and follower contact has proved to be the most difficult component of the automobile engine in terms of lubrication. Conventional wisdom would have it that it operates with boundary lubrication. That is, the choice of the solid materials and lubricant additives are seen as of paramount importance in ensuring that satisfactory reaction films are being continuously formed on the interacting surfaces.

Nevertheless it is now clear that hydrodynamics has a role to play in cam and follower lubrication and that mixed lubrication (and even elastohydrodynamic action at certain points during the cycle) is relevant. Whereas an engine bearing is a conformal contact (ie the bearing or sleeve wraps closely around the journal), all automobile cam and follower geometries are counterformal. Thus typical mean pressures are much higher for the cam and follower and film thicknesses are lower. In this context the mode of lubrication is clearly seen as being more severe.

3.3. Piston Rings

In the case of the lubrication of piston rings it is immediately evident that the relative sliding speed between rings and liner is substantial at mid-stroke and zero at the dead centres. It is this relative velocity that entrains lubricant into the contact and reflects in part the prospect of satisfactory film lubrication. Thus at mid-stroke a most satisfactory

thickness of lubricant film may exist between the compression rings and cylinder liner with hydrodynamic lubrication conditions. At the dead centres, with zero entraining velocity of the lubricant into the contact, satisfactory behaviour will depend upon squeeze film action and elastohydrodynamic/mixed/boundary actions will have significance in this vicinity. It is interesting to observe that conditions of zero lubricant entrainment velocity also occur in the case of engine bearings and cam and follower operation, although this is not as obvious as with the piston ring situation.

4. CONCLUSION

The physical and chemical actions associated with the modes of lubrication known as hydrodynamic, elastohydrodynamic, mixed and boundary have been identified and discussed. These regimes of lubrication have been related to the major tribological components of the automobile engine. These components rely upon different mechanisms for satisfactory lubrication and indeed each may enjoy a range of lubrication modes during a single cycle. This reflects the challenge and compromises facing the designer in improving the operational characteristics of these machine elements, in response to the demands for improved performance in the context of energy efficiency and emissions requirements.

This page intentionally left blank

Engine Tribology / C.M. Taylor (Editor)

ENGINE BEARINGS : BACKGROUND AND LUBRICATION ANALYSIS

C M TAYLOR

Department of Mechanical Engineering, University of Leeds, Leeds LS2 9JT

1. INTRODUCTION

Modern day design techniques for engine bearings date back some twenty five years. Many plain journal bearing applications may be treated as steadily loaded for design purposes, that is the applied load can be taken as essentially fixed in magnitude and direction of application. However, with engine bearings, such as the main or crankshaft bearings, the big-end and gudgeon pin bearings associated with the connecting-rod and even the camshaft bearings this is not the case. The load is dynamic for these situations, the magnitude and direction being a function of time. Consequently the journal or shaft does not remain at a fixed location within the clearance space. Rather it executes a cyclic orbit and one of the main purposes of the design procedures developed is to determine the locus of the shaft movement and hence the minimum film thickness occurring during the cycle. This is seen as an important design parameter.

During the 1960's Booker in the United States of America and Blok and his associates in The Netherlands were independently developing graphical design techniques to predict the locus of the journal centre in a dynamically loaded journal bearing. The background has been well described by Campbell et al (1968). Booker's approach, which has become known as the **Mobility Method,** has prevailed and is still the mainstay of engine bearing performance prediction. In fact the **Impulse Method** associated with Blok was somewhat more effective in terms of the maps which were developed for graphical orbit prediction. However, the concurrent development of the digital computer soon rendered graphical approaches obsolete and the effective design data presentation of Booker proved to be attractive.

The **Mobility Method** has developed into a reliable and robust technique which is capable of use by a competent engineer, as opposed to a fluid film bearing specialist. No matter which particular bearing is being considered the problem reduces to the same format. The cyclic loading for a given rotational speed is determined from a knowledge of the cylinder pressure variation, the inertial contribution of the reciprocating and rotating components and, of course, the geometry. This load is assumed to be known a priori. The kinematics of the crank and slider mechanism are straightforwardly calculated and the motion of the shaft can be determined by taking the hydrodynamically generated load capacity

due to the pressures in the fluid film to balance the applied load. In this approach a consideration of dynamic instabilities is therefore precluded.

It is not proposed to develop here the procedures for calculating the loading and kinematics of specific bearings. The general approach for the crank and slider mechanism is widely available in the literature and the design report of Martin (1974) is particularly useful. A polar load diagram for the big-end connecting rod bearing of a 6 cylinder petrol engine passenger car is shown in Figure 1 whilst that for the main bearing number 1 is given in Figure 2.

Attention will primarily be given to the development of the understanding of the lubrication equations appropriate to the situation and how these can be evolved into the **Mobility Method** design approach. The paper will conclude with an assessment of the limitations of the technique thus highlighting future developments which may be anticipated.

1.1 Notation

A,B	constants of integration (Appendix)
b	bearing width
c	radial clearance (=(R-r)=half diametral clearance)
d	Shaft diameter
e	eccentricity ($\dot{e}$ = de/dt)
$f(\xi,\eta)$	polynomial function (Booker (1969))
h	lubricant film thickness
M	mobility number/vector (with components in ε, ψ, ξ, η directions)
p	pressure
p_{max}	maximum instantaneous pressure
$\bar{p}_{max}$	maximum pressure ratio (= $p_{max}/(P/bd)$)
P	total instantaneous load (with components P_a, P_n)
r	shaft radius
R	bush radius
t	time
V_s	squeeze' velocity
x,y,z	Cartesian coordinates
β	angle between fixed axis and line of shaft(S) and bush (B) centres ($\dot{\beta}$ = dβ/dt)
ε	eccentricity ratio (= e/c) ($\dot{\varepsilon}$ =dε/dt)
η	lubricant dynamic viscosity
ξ,η	rectangular coordinates in load frame (Figure 4).
ϕ	angular coordinate (measured from maximum film thickness)
ϕ_1,ϕ_2	limits of full film region
ψ	attitude angle ($\dot{\psi}$ = dψ/dt)

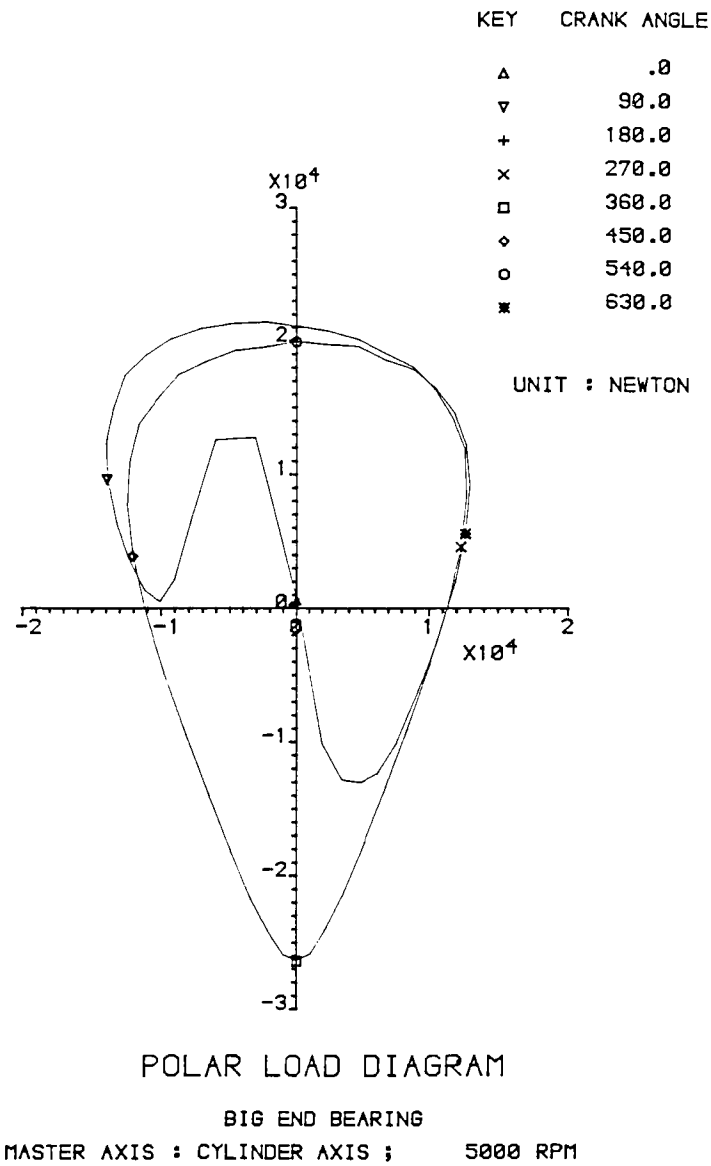


Figure 1 - Polar Load Diagram for the Big-End Connecting Rod Bearing of a 6 Cylinder Automotive Petrol Engine

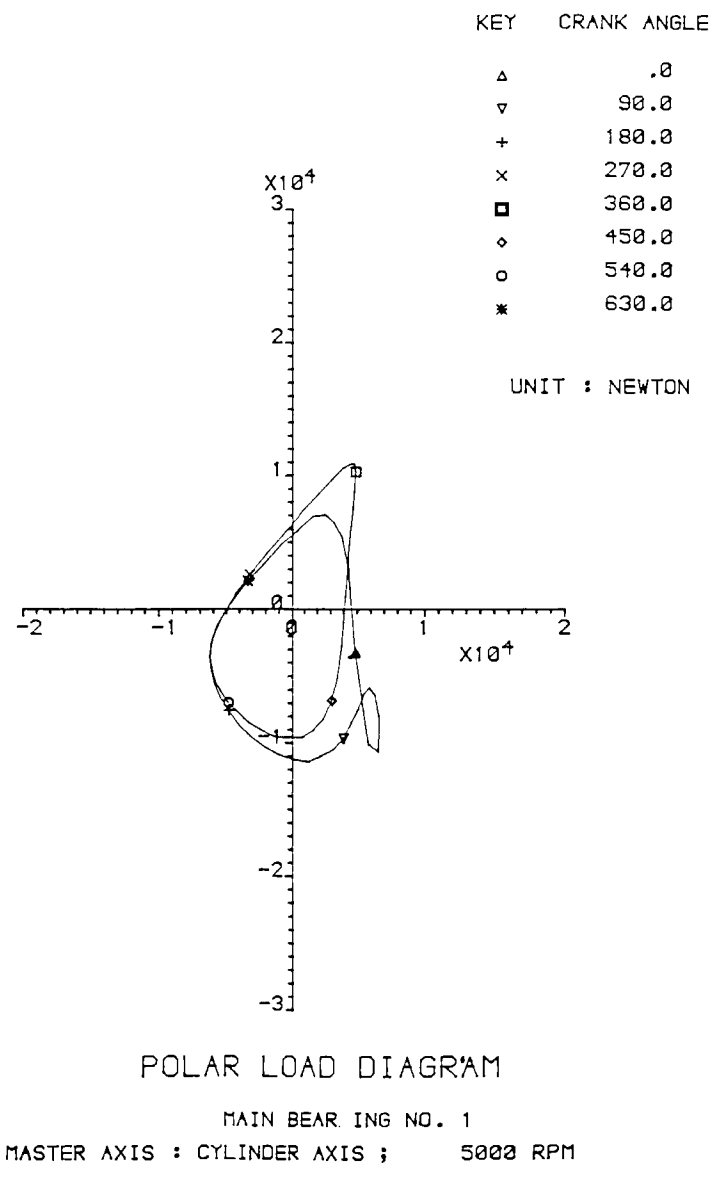


Figure 2 - Polar Load Diagram for the Main Bearing Number 1 of a 6 Cylinder Automotive Petrol Engine

$$\overline{\omega} \qquad \left(\frac{\Omega_s+\Omega_b}{2}-\Omega_L\right)$$

Ω_b, Ω_L, Ω_s angular velocities of bush, load and shaft.

2. LUBRICATION ANALYSIS FOR DYNAMICALLY LOADED BEARINGS

The Reynolds equation in a fixed axis coordinate system for a situation of dynamic loading may be expressed in cylindrical polar coordinates as,

$$\frac{\partial}{\partial\phi}\left(h^3\frac{\partial p}{\partial\phi}\right)+r^2\frac{\partial}{\partial y}\left(h^3\frac{\partial p}{\partial y}\right)$$

$$=12\eta r^2\left(\dot{e}\cos\phi+e\sin\phi\left(\dot{\beta}-\frac{\Omega_s+\Omega_b}{2}\right)\right) \qquad (1)$$

The symbols are presented in the notation and understanding will be clarified by reference to Figure 3. This form of the equation is for an isoviscous and incompressible fluid. This second order partial differential equation enables the film pressure (p) at any location (ϕ,y) to be calculated for a given lubricant viscosity and geometry dependent upon the instantaneous film shape (h = c+ecos ϕ for aligned conditions), the angular velocities of the shaft and bush (Ω_s Ω_b) and the instantaneous radial velocity ($\dot{e}$) and tangential angular velocity $(\dot{\beta})$of the shaft centre.

With a knowledge of the film pressures, the load capacity due to hydrodynamic action can be determined and ($\dot{e}$,$\dot{\beta}$) can be adjusted to ensure the applied load is vectorially balanced by this hydrodynamic generation of pressure. When this balance is achieved, with a knowledge of the shaft centre velocity, a journal movement over a small time period can be determined. A new solution of equation (1) for the new instantaneous operating conditions is then required and the process is repeated until cyclic convergence is obtained.

It is instructive to consider briefly some simplified forms of equation (1) to give confidence in its correctness.

(a) Steady load situations

Normally the bush is stationary (Ω_b=0) and for the case of no shaft movement ($\dot{e}$=$\dot{\beta}$=0). The dynamic form of Reynolds equation reduces to the well known steady form,

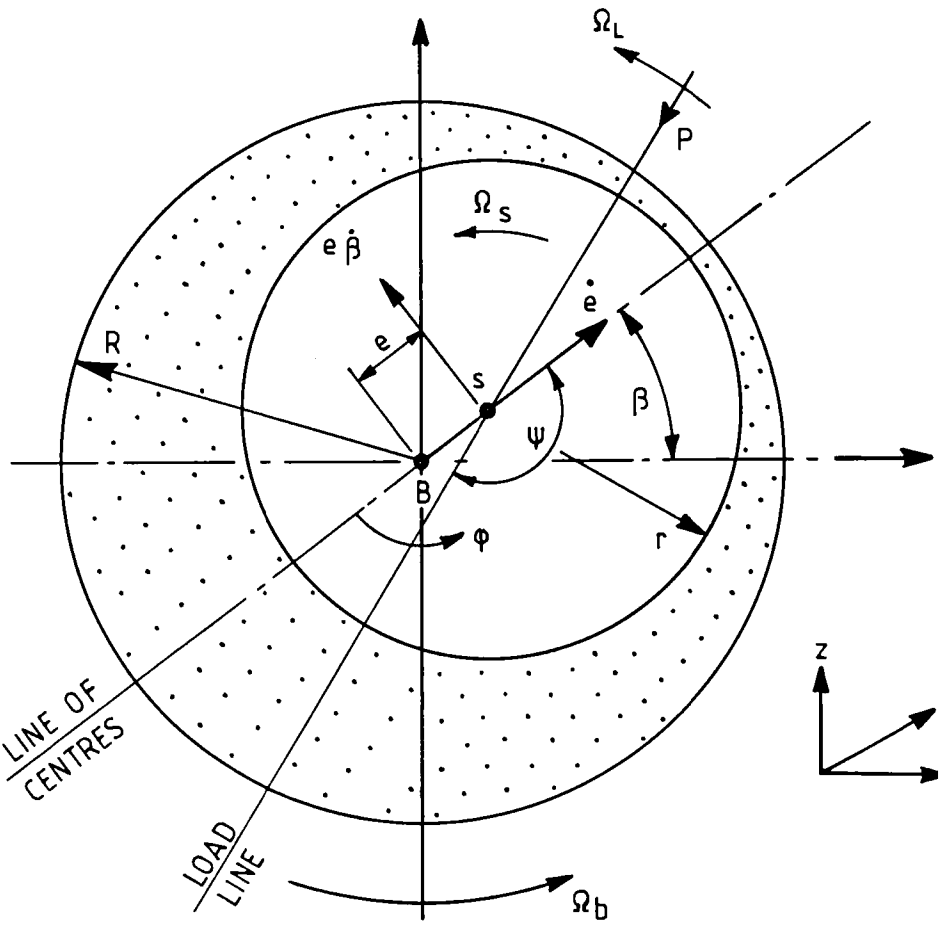


Figure 3 - Journal Bearing Geometry for Dynamic Loading

$$\frac{\partial}{\partial\phi}\left(h^3\frac{\partial p}{\partial\phi}\right)+r^2\frac{\partial p}{\partial y}\left(h^3\frac{\partial p}{\partial y}\right)=-6\eta r^2\Omega_s\, e\sin\phi$$

$$=6\eta r^2\Omega_s\frac{\partial h}{\partial\phi} \qquad (2)$$

(b) In the absence of all rotational motions - no rotation of the shaft or bush about their own centres ($\Omega_s = \Omega_b = 0$) and no precession of the shaft about the bush centre ($\dot{\beta}=0$) - the pure squeeze film Reynolds equation is obtained,

$$\frac{\partial}{\partial\phi}\left(h^3\frac{\partial p}{\partial\phi}\right)+r^2\frac{\partial}{\partial y}\left(h^3\frac{\partial p}{\partial y}\right)=12\eta r^2\dot{e}\cos\phi \qquad (3)$$

Here the shaft is translating along a diametral chord of the bearing or clearance circle.

It is interesting to note that prior to the development of the **Mobility Method** the engine bearing designer achieved an understanding of shaft motion by identifying periods in which hydrodynamic wedge action dominated (with little change in shaft eccentricity) and periods when pure squeeze motion was the primary action responsible for generation of fluid film pressures. (Stone and Underwood (1947)). For the latter situation equation (3) - often simplified - could be used for assessment of motion whilst for the former an 'equivalent' steady load situation, with constant eccentricity ($\dot{e}=0$), could be examined with ($\dot{\beta} = \Omega_L$ e.g. out of balance) and the load capacity proportional to (Ω_L - (Ω_s + Ω_b)/2) - see equation (1). This was termed the 'equivalent speed concept'. This line of thought leads to an interesting physical interpretation. As can be seen, when,

$$\Omega_L=\frac{\Omega_s+\Omega_b}{2}$$

the right hand side of Reynolds equation becomes zero implying no pressure generation or load capacity. This is the so called 'half speed whirl' situation. In general terms if the shaft precesses around the bush centre with an angular velocity equal to the mean of that of the shaft and bush and there is no radial shaft movement,

$$\text{i.e. } \dot{\beta}=\frac{\Omega_s+\Omega_b}{2} \text{ and } \dot{e}=0$$

then there will be no load capability. A line of physical reasoning to explain this is that there is no overall pumping action of lubricant into a convergent hydrodynamic wedge to generate pressures. This situation generally occurs at

some stage during an engine bearing cycle and the importance of squeeze film action at this stage will be evident. The concept expands into 'frictional frequency whirl' in bearing influenced rotor dynamic problems.

2.1 The Basis of the Mobility Method.

The general mobility approach may be applied to bearings of finite width. In practice short bearing theory has commonly been adopted. The reasons for this are various, the main one being that a single formulation for bearings of different width-to-diameter ratios is possible. Broader considerations of solution accuracy will be considered later. The analytical basis with the short journal bearing approximation for dynamic conditions is outlined in the Appendix and the ideas will be developed here.

The pressure distribution in the lubricant film will be given by equation (A2) from the Appendix,

$$p = -\frac{6\eta}{h^3}\left(\frac{b^2}{4} - y^2\right)\left(\dot{e}\cos\phi + e\sin\phi\left(\dot{\beta} - \frac{\Omega_s + \Omega_b}{2}\right)\right) \qquad (A2)$$

For the case of a steady load ($\dot{e} = \dot{\beta} = 0$) and with the bush normally non-rotating (Ω_b=0) this reduces to the well known form

$$p = \frac{3\eta\Omega_s e}{h^3}\left(\frac{b^2}{4} - y^2\right)\sin\phi$$

The significant merit of the short journal bearing approximation is that it is a reasonably accurate simplification for bearings of small width-to-diameter ratio (say < 0.6) at modest eccentricity ratios (say < 0.8). The analytical solution is extremely convenient and can accommodate the occurrence of cavitation (a two phase flow) by neglecting the sub-ambient pressures in the divergent region ($(\pi \leq \phi \leq 2\pi)$ and considering only the superambient pressures in the convergent region ($0 \leq \phi \leq \pi$) (ϕ being measured from the maximum film thickness on the line of centres of the shaft and bush in the direction of average rotation).

The expression for pressure distribution with a dynamically loaded bearing can be integrated to give the components of the load along (P_a) and normal (P_n) to the line of the shaft and bush centres, if the limits of the lubricant film are known (see Appendix). To develop the argument it will be assumed that the fluid film remains full, that is there is no cavitation, and the limits of film extent are from $\phi_1 = 0$ to $\phi_2 = 2\pi$. This assumption will be dropped in presenting realistic data. Equations (A3) of the Appendix then show that,

$$P_a = \frac{\eta r b^3}{c^3}\left(\frac{\pi\left(1+2\varepsilon^2\right)}{\left(1-\varepsilon^2\right)^{5/2}}\right)\dot{e} \qquad \text{(a)}$$

$$P_n = \frac{\eta r b^3}{c^3}\left(\frac{\pi\,\varepsilon}{\left(1-\varepsilon^2\right)^{3/2}}\right)\left(\frac{\Omega_s+\Omega_b}{2}-\dot{\beta}\right) \qquad \text{(b)} \qquad \text{(A3)}$$

Here again simplification to the steadily loaded case gives well known results for **non-cavitating** situations,

$$P_a = 0$$

$$P_n = \frac{\eta r b^3\,\Omega_s}{2\,c^3}\left(\frac{\pi\,\varepsilon}{\left(1-\varepsilon^2\right)^{3/2}}\right)$$

The assumption of a full film leads to the prediction of an attitude angle (the angle between the load line and line of centres) of 90° because of the antisymmetric nature of the circumferential pressure variation about the line of centres). This is the same as saying that the component of load along the line of centres is zero.

Noting that the load components (P_a and P_n) are related to the total load ($P = (P_a^2 + P_n^2)^{1/2}$) through the attitude angle (ψ),

$$P_a = P\cos\psi$$

$$P_n = P\sin\psi$$

then equations (A3) can be simply re-arranged as follows in the style of Booker (1965a) in which the mobility formulation was first introduced,

$$\dot{\varepsilon} = \frac{\dot{e}}{c} = \left(\frac{P(c/r)^2}{\eta b d}\right)\left(\frac{\left(1-\varepsilon^2\right)^{5/2}\cos\psi}{2\pi\left(1+2\varepsilon^2\right)(b/d)^2}\right) \qquad \text{(a)}$$

$$\varepsilon\dot{\beta} = \left(\frac{P(c/r)^2}{\eta b d}\right)\left(\frac{-\left(1-\varepsilon^2\right)^{3/2}\sin\psi}{2\pi(b/d)^2}\right) + \frac{\varepsilon(\Omega_s+\Omega_b)}{2} \qquad \text{(b)} \qquad \text{(4)}$$

Defining mobilities (M_ε and M_ψ) - dimensionless ratios of velocity to force - as follows,

$$M_\varepsilon = \frac{\left(1-\varepsilon^2\right)^{5/2}\cos\psi}{2\pi\left(1+2\varepsilon^2\right)(b/d)^2} \qquad \text{(a)}$$

$$M_\psi = \frac{\left(1-\varepsilon^2\right)^{3/2}\sin\psi}{2\pi\,(b/d)^2} \qquad \text{(b)} \qquad (5)$$

then equations (4) become,

$$\dot{\varepsilon} = \left(\frac{P(c/r)^2}{\eta bd}\right) M_\varepsilon \qquad \text{(a)}$$

$$\varepsilon\dot{\beta} = \left(\frac{P(c/r)^2}{\eta bd}\right) M_\psi + \frac{\varepsilon\left(\Omega_s+\Omega_b\right)}{2} \qquad \text{(b)} \qquad (6)$$

It will be seen from Figure 3 that the vector sum of ($\dot{e}$) and ($e\dot{\beta}$) represents the magnitude and direction of the motion of the shaft centre. Thus in principle the motion can be determined during a transient event from equation (6). The details of this will be considered later, but now the more realistic situation incorporating cavitation will be dealt with.

2.2 The Mobility Method with Cavitation

In practice the occurrence of sub-ambient pressures in the lubricant film will be restricted by the ventilation of air from the surroundings at atmospheric pressure (zero gauge pressure). This is termed gaseous cavitation and should be clearly distinguished from boiling of a liquid at its vapour pressure - vaporous cavitation - which can lead to surface damage due to collapse of the vaporous cavitities. Considering the form of the pressure distribution, equation (A2), the region of positive pressure will extend over π radians between (ϕ_1) and (ϕ_2) such that,

$$\dot{e}\cos\phi_1 + e\sin\phi_1\left(\dot{\beta} - \frac{\Omega_s + \Omega_b}{2}\right) = 0$$

$$\dot{e}\sin\phi_1 - e\cos\phi_1\left(\dot{\beta} - \frac{\Omega_s + \Omega_b}{2}\right) \geq 0 \qquad (7)$$

$$\phi_2 = \phi_1 + \pi$$

The integrals to determine the load and mobility components resolved along and normal to the line of centres (Appendix) will have to be undertaken between (ϕ_1) and (ϕ_2). The process of determining (ϕ_1) and (ϕ_2) involves numerical analysis. Booker (1965b) first carried this out for the short journal bearing approximation and curve fitted his data. This was repeated and shown to be commendably accurate by Clayton (1990). The process can be undertaken in cylindrical polar or rectangular Cartesian coordinates. Booker pointed out that rectangular coordinates eliminated numerical problems near the origin. As will be seen in the next section this curve fitting is carried out with a coordinate system fixed to the load line.

2.3 Practical Aspects of The Mobility Method

The development of the **Mobility Method** in the early 1960's was directed towards the production of a mobility map, with a view to stepping out a shaft orbit by hand. In order to avoid having to 'turn' the map as the direction of the load varied, the load was considered to act in a fixed direction (vertically downwards), and the actual load motion introduced subsequently to locate the shaft position with respect to the bush. This fixed load line direction meant that all angular velocities in the Reynolds equation were with respect to the load angular velocity (Ω_L).

If the curve fits (Booker 1965b) for the mobility number data taking into account cavitation which were mentioned in the previous section are now introduced (see Figure 4 for the coordinate systems with respect to the load line) we have,

$$M_\xi = \frac{(1-\xi)^{5/2}}{\pi(b/d)^2} \qquad \text{(a)}$$

$$M_\eta = \frac{-4\eta(1-\xi)^{3/2}}{\pi^2(b/d)^2} \qquad \text{(b)} \qquad (8)$$

where ($\zeta = \varepsilon\cos\psi$ and $\eta = \varepsilon\sin\psi$). Resolved into the cylindrical polar coordinates of equation (6),

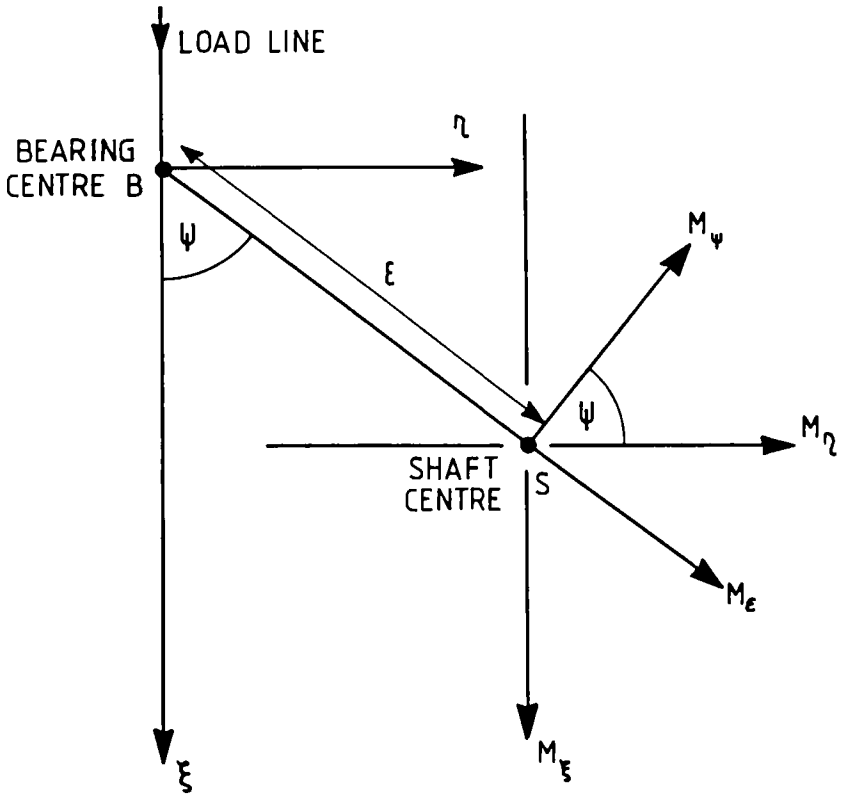


Figure 4 Mobilities in Cylindrical Polar and Rectangular Cartesian Coordinates

$$M_\varepsilon = \frac{(1-\varepsilon\cos\psi)^{3/2}}{\pi^2\,(b/d)^2}\left(\pi\cos\psi(1-\varepsilon\cos\psi) - 4\varepsilon\sin^2\psi\right) \qquad \text{(a)} \tag{9}$$

$$M_\psi = \frac{-(1-\varepsilon\cos\psi)^{3/2}\sin\psi}{\pi^2\,(b/d)^2}\left(4\varepsilon\cos\psi + \pi(1-\varepsilon\cos\psi)\right) \qquad \text{(b)}$$

It will be clear that the scalar magnitude of resultant mobility is given by,

$$M = (M_\xi^2 + M_\eta^2)^{1/2} = (M_\varepsilon^2 + M_\psi^2)^{1/2} \tag{10}$$

From Figure 3 the following identity can be obtained reflecting the shaft centre angular velocity with respect to the load $(\dot{\psi})$,

$$\dot{\psi} = \dot{\beta} - \Omega_L \tag{11}$$

Thus with the fixed load line direction and the mobility values taking into account cavitation (equations (8) and (9)), it is possible to define straightforwardly the shaft centre velocity components relative to the load. Retaining the cylindrical polar coordinates for ease of explanation, these are ($\dot{\varepsilon}$) and ($\varepsilon\dot{\psi}$) which from equations 6(a) and 6(b) can be seen to be,

$$\dot{\varepsilon} = \frac{P(c/r)^2}{\eta bd} M_\varepsilon \qquad \text{(a)}$$

$$\varepsilon\dot{\psi} = \frac{P(c/r)^2}{\eta bd} M_\psi + \varepsilon\left(\frac{\Omega_s + \Omega_b}{2} - \Omega_L\right) \qquad \text{(b)} \qquad (12)$$

It is clearly simple to resolve these into a Cartesian frame of coordinates if so wished.

The procedure for calculating the shaft vector velocity whilst it was at a specific location and taking a step over a time interval was as follows,

(a) The motion over a small time interval was considered to be the same as that obtained by adding two components, one taking all rotational velocities to be zero ($\Omega_S = \Omega_b = \Omega_L = 0$) and one accounting for the angular velocities. The first step can be termed a **'squeeze'** step (although not pure normal squeeze) and the second step a 'whirl' step.

(b) For the **squeeze** step, if **M** is the vector of (M_ε) and (M_ψ), then from equation (12) representing $\mathbf{V_S}$ as the vector sum of ($\dot{\varepsilon}$) and ($\varepsilon\dot{\psi}$) we have

$$V_S = \frac{P(c/r)^2}{\eta bd} . M \left(\text{unit } s^{-1}\right)$$

Now the scalar magnitude of the vector M can be calculated at any point (ε, ψ) from equation (10) using either equations (8) or (9). The direction associated with the vector is given by the tangent of the ratio of the resolved mobility components and is termed the **squeeze path**.

(c) Thus in Figure 5 the clearance circle, a circle representing the possible range of movement of the shaft centre and of radius equal to the radial clearance (c), is shown. The magnitude of mobility M at any point in the clearance circle can be determined and the **squeeze path** direction calculated. Lines of constant mobility may be determined. If the shaft is at a given point (A) as shown its movement over a time period (dt) will be given by

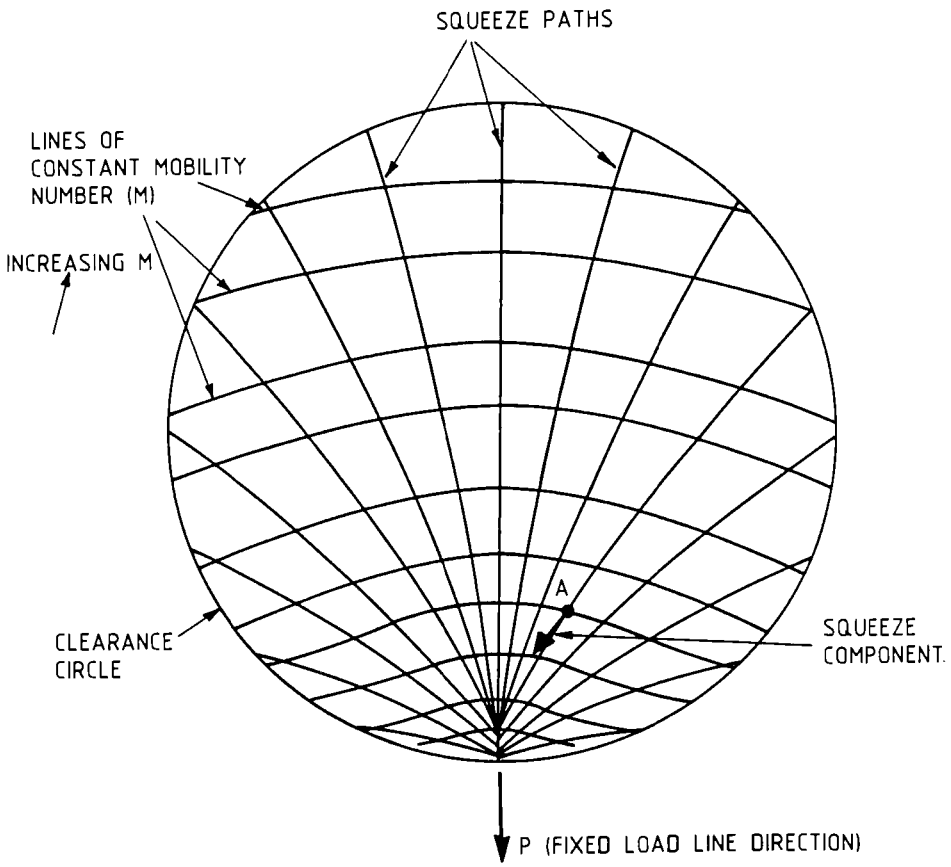


Figure 5 - Short Bearing Mobility Map for a Cavitating Film Showing the 'Squeeze' Component

$$V_s \,.\, c.\, dt = \frac{P(c/r)^2}{\eta bd}\, c.Mdt$$

in the direction of the squeeze path. (see Figure 5). Clearly this will be scaled.

(d) Now the angular velocities can be taken into account. The shaft movement due to these will be a component (see equation 12(b)),

$$\varepsilon.c.\left(\frac{\Omega_s + \Omega_b}{2} - \Omega_L\right) dt = \varepsilon.c.\varpi dt$$

directed tangentially. Here a sign convention is necessary and anticlockwise rotations have usually been taken as positive. This component can be vectorially added to the squeeze component to give the resultant motion, as shown in Figure 6. Strictly speaking since the value of the average mobility number of the resultant will be different from that used in calculating the squeeze step, an iteration to some acceptable tolerance will be required. However, for small time steps when the procedure is carried out on the computer this is not necessary.

(e) The procedure described above has enabled the movement over a small time period to be predicted. The procedure can now be repeated starting from the new position. Normally the average values of quantities such as (P, Ω_b and Ω_L) are used over a step. In this way an orbit can be mapped out. For a four stroke engine cycle convergence of the orbit will take somewhat more than 4π radians of crankshaft revolution from the initial, guessed starting point. For the mobility map, with a fixed load line direction, this orbit will be in the lower portion reflecting the action of the load driving the shaft to the part of the clearance circle in which it is directed. However, this orbit can now be opened out knowing the true instantaneous load direction to give the shaft motion with respect to the bush.

With the advent of digital computers the concept of the graphical technique soon disappeared. Here there is no need for the idea of a fixed load direction and simple transformations can be used to effect the calculations carried out (see Booker (1971)). The process is remarkably quick on a digital computer, the calculation of an orbit for a four stroke cycle being entirely robust and taking a fraction of a second. Figure 7 shows the crankpin orbit of the big-end of a well known benchmark test case, the Ruston and Hornsby 6VEB-X Mk III turbocharged diesel (Campbell et al (1968)).

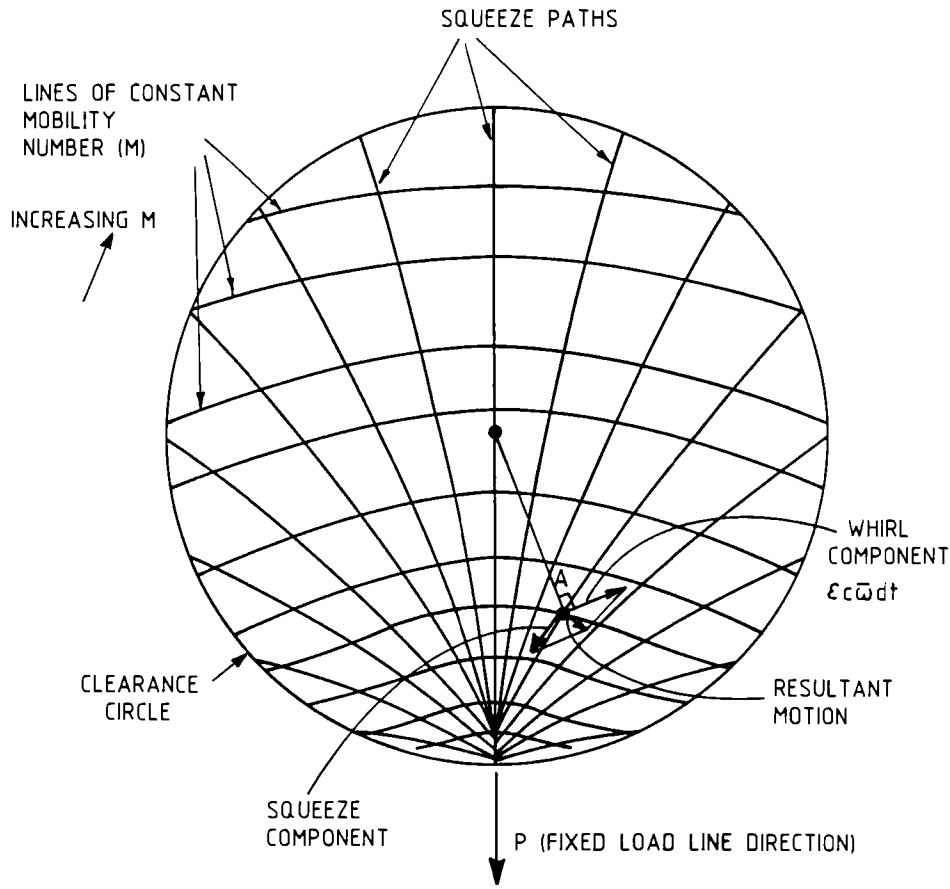


Figure 6 - Short Bearing Mobility Map for a Cavitating Film Showing the 'Whirl' Component and Resultant

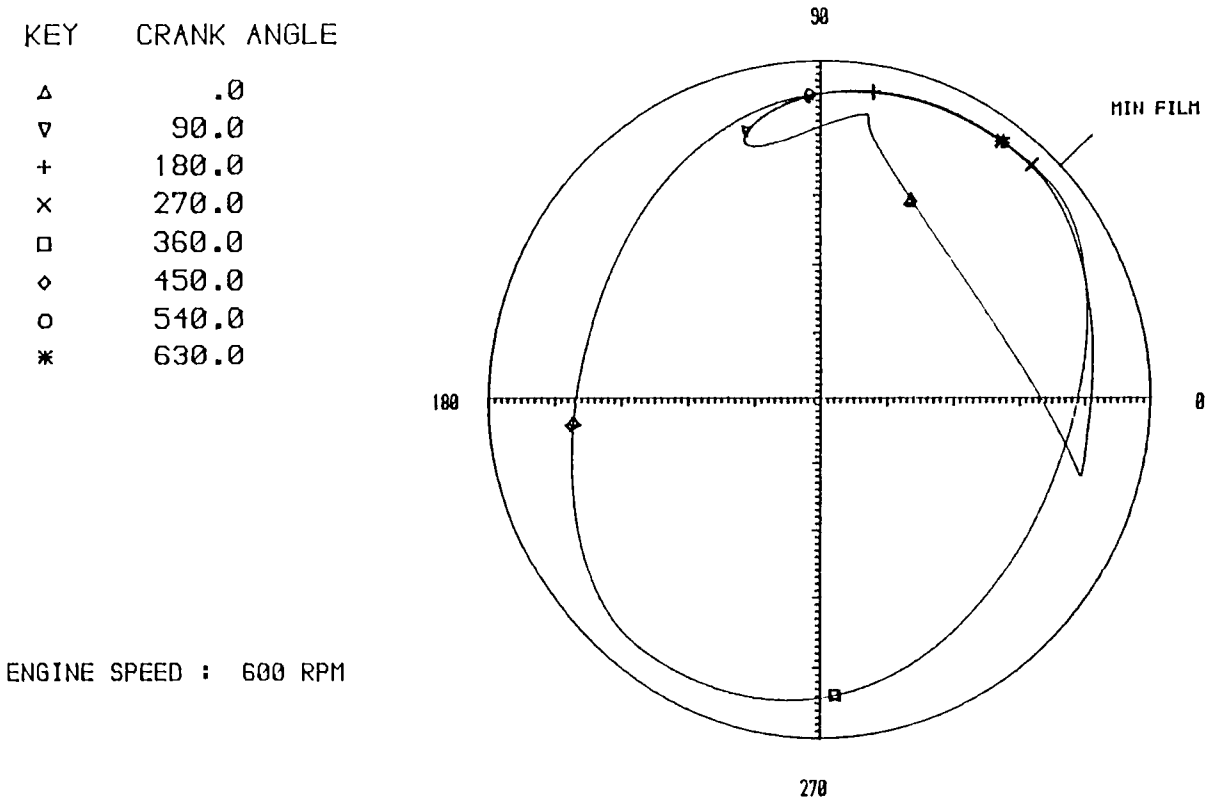


Figure 7 - Big-End Connecting Rod Crankpin Orbit for Ruston and Hornsby 6VEB-X Mark III Turbo-Charged Diesel (Cambell et al (1968)

For the short journal bearing solution it will be noted that,

$$M\alpha \frac{1}{(b/d)^2}$$

Thus a short bearing mobility map can be constructed for (b/d = 1) and simply extended for any other width-to-diameter ratio. The squeeze path directions are independent of (b/d). The simplicity of this approach is now clearly revealed.

3. MAXIMUM PRESSURE IN A DYNAMICALLY LOADED JOURNAL BEARING

The mobility analysis technique which has been described enables the prediction of minimum lubricant film thickness during the cycle. It is worthy of note that Booker (1969) extended the approach to incorporate determination of the maximum film pressure. A map of maximum pressure ratio, the maximum pressure divided by the specific load,

$$\bar{p}_{max} = \frac{p_{max}}{(P/bd)}$$

was developed. The full curve fits to enable the maximum pressure to be calculated in the (ξ,η) coordinates of Figure 4 are,

$$\bar{p}_{max}\ (\xi,o) = \begin{cases} \dfrac{6}{(1-x)^3} M_\xi\ (\xi,o), -\dfrac{1}{2} \le \xi \le +1 \\ -\dfrac{8}{9x} M_\xi\ (\xi,o), -1 \le \xi \le -\dfrac{1}{2} \end{cases}$$

$$\bar{p}_{max}\ (\xi,\eta) = f\ (\xi,\eta)\ \bar{p}_{max}\ (\xi,o)$$

where the function $f(\xi,\eta)$ is a third order polynomial detailed in Booker (1969). The curve fit for M_ξ given in equation (8) may be adopted or Booker gives more precise data. This approach is clearly amenable to simple calculation on a digital computer when the orbit is determined.

It is worthy of note that Goenka (1984) generated mobility data for **finite width** bearing analysis as opposed to short bearing analysis. His curve fits covered mobility number, maximum pressure and its location and the extent and location of the pressurised film.

4. DISCUSSION

It is instructive to consider the magnitudes of influential parameters which are encountered in modern day automobile engine bearings. Although the list below is not complete and is only intended to give an engineering feel, it does focus attention upon the critical nature of the operation.

Minimum lubricant film thickness	1 μm
Maximum shear rate	10^7 s^{-1}
Maximum bearing temperature	120-150°C
Specific loading (up to)	40 MPa
Maximum film pressure	250 MPa
Lubricant dynamic viscosity	0.0025 Pa.s
Power loss per bearing (high speed)	0.25 kW
Lubricant flow rate per bearing (typically)	$15x10^{-6}$ m^3/s (0.015 l/s)

The values will depend upon many factors but it is apparent that operating conditions may well be critical and increased sensitivity in design is needed. Application of the design technique is made in the hope that the simple model will sufficiently reflect reality such that the **trends** in variations with parametric changes will be reflected. As with the analyses of all the major frictional devices in the internal combustion engine, the predicted magnitudes can only be a benchmark and not absolute. Models of increased sensitivity are now required, hopefully without an undue extension of the complexity of application. This can be seen by an itemized consideration of the assumptions and limitations associated with this analytical technique. (Taylor (1991)).

Short journal bearing analysis: Highly inaccurate at high eccentricity ratios and poor flow rate/power loss prediction. However, finite width bearing analysis now straightforwardly possible.

Circumferential symmetry: Not usually the case due to presence of grooves/holes, non-circularity, distortions etc. A simplified approximation to take account of non-axisymmetric effects is possible but a detailed consideration is complex.

Oil film history: Mobility analysis takes lubricant to be available when required. In reality transient continuity considerations may lead to limitations on the film extent. This will be related to the groove/hole topography and supply pressure conditions and a proper consideration is important for good flow rate and power loss prediction. Although analytical techniques to enable assessment of oil film history effects have been developing the computational effort is still somewhat prohibitive.

Elastic/thermal distortions: Not considered either in the local deflection or structural flexibility sense. The current techniques of determining the bearing load decouple the fluid film analysis and introduce a welcome simplicity. In actuality a consideration of elastic distortions will reintroduce the link and although the problem is well posed the analytical procedure and generalized interpretation still pose formidable difficulties. Notwithstanding the problems, the importance of the proper consideration of distortions has attracted growing attention and valuable progress has been made.

Misalignment: This is not addressed and is still a very difficult technical consideration.

Non-Newtonian lubricant behaviour: Mobility analysis assumes a Newtonian lubricant. At the shear rates encountered in engine bearings it is well known that multigrade lubricants exhibit shear thinning and even viscoelastic effects. Analytical considerations of such influences are limited though the background lubricant technology is growing stronger.

Component inertia/vibrations: The pressures generated in the lubricant film are taken to balance the applied load. The neglect of shaft inertia may well be reasonable for most circumstances. The general contribution of bearing stiffness and damping to vibrational behaviour has received little attention.

Thermal analysis: This is rarely undertaken, with an effective viscosity at a specified operating temperature normally being adopted. Good predictions of lubricant flow rate and power loss are required and thus the Mobility Method renders quite unsatisfactory predictions. The background thermal environment of the engine is highly influential and the distribution of heat from viscous dissipation to the lubricant and adjacent solids may be important.

Effect of pressure on viscosity: Not considered although the values of maximum pressures with rigid surfaces would indicate that it is important.

The above catalogue of assumptions demonstrates clearly the challenge to those with an interest in engine bearing fluid film lubrication. It is not just a matter of Reynolds equation but a complex problem when properly put invokes challenging fundamental, physical and analytical arbitrations. It is beyond the current capabilities of designers to address the analysis with all the complications identified. A step-by-step approach to make predictions more sensitive is required, with the exercise of judgement in making reasonable simplifications being an important ingredient.

5. CONCLUSIONS

(i) The **Mobility Method** approach for the calculation of the orbit of a shaft within the clearance space of a dynamically loaded journal bearing has been described. Design data has been presented and its nature justified analytically for the case of the short journal bearing approximation.

(ii) The technique when programmed onto a digital computer offers a remarkable robust and exceptionally quick approach to predict engine bearing journal orbits. Graphical facilities to show clearly the orbit, pressure maps, full film and cavitation regions can assist the design process considerably.

(iii) The technique as currently used offers benchmark predictions which hopefully reflect trends as parameters are varied. The wide range of assumptions and limitations have been pointed out and the challenge for the future will be to make predictions more sensitive.

REFERENCES

Booker, J.F. (1965a), "Dynamically Loaded Journal Bearings : Mobility Method of Solution", Trans ASME, Jrl. Basic Eng., Sept. pp 537-546.

Booker, J.F. (1965b), "A Table of the Journal-Bearing Integral", Trans ASME, Jrl. Basic Eng., June, pp 533-535.

Booker, J.F. (1969), "Dynamically Loaded Journal Bearings : Maximum Film Pressure", Trans. ASME, Jrl. Lub. Tech., July, pp 534-543.

Booker, J.F. (1971), "Dynamically Loaded Journal Bearings : Numerical Application of the Mobility Method", Trans ASME, Jrl. Lub. Tech., Jan, pp 168-176.

Campbell, J. et al, (1968), "Bearings for Reciprocating Machinery : A Review of the Present State of Theoretical, Experimental and Service Knowledge", Proc. I.Mech.E., 182, Pt 3A (Conf. on Lub. Wear), pps. 34.

Clayton, G.A. (1990), "Engine Bearing Analysis and Design", Ph.D. Thesis, Mech. Eng. Dept., Univ. of Leeds, U.K., pps 228.

Ocvirk, F.W. (1952), "Short Bearing Approximation for Full Journal Bearings", NACA, Tech. Note 2808.

Goenka, P.K. (1984), "Analytical Curve Fits for Solution Parameters of Dynamically Loaded Journal Bearings", Trans. ASME, Jnl. Lub. Tech, October (106), pp 421-428.

Martin, F.A. (1974), "Design Procedures for Dynamically Loaded Journal Bearings", Glacier Metal Co. Ltd. CG48/74, pps 32.

Stone, J.M. and Underwood, A.F. (1947), "Load Carrying Capacity of Journal Bearings", Qt. Trans. SAE, 1, p.56.

Taylor, C.M. (1991), "Fluid Film Lubrication in the Internal Combustion Engine", Frontiers of Tribology Conf., IOP, J.Physics D: Applied Physics, 25, 1992, pp 91-100.

APPENDIX

The Short Journal Bearing Solution for Dynamic Loading

If the bearing is short in the axial (y) direction the pressure flow in this direction is more significant than that in the circumferential (ϕ) direction which can be neglected (Ocvirk (1952)). This implies that the first term on the left hand side of the Reynolds equation (1) may be omitted giving,

$$\frac{\partial}{\partial y}\left(h^3 \frac{\partial p}{\partial y}\right) = 12\eta\left(\dot{e}\cos\phi + e\sin\phi\left(\dot{\beta} - \frac{\Omega_s + \Omega_b}{2}\right)\right) \qquad \text{(A1)}$$

Integration, twice, for a well aligned bearing (h not f(y)) gives,

$$p = \frac{6\eta}{h^3}\left[\dot{e}\cos\phi + e\sin\phi\left(\dot{\beta} - \frac{\Omega_s + \Omega_b}{2}\right)\right] y^2 + Ay + B$$

With the boundary conditions p = o at y = ±b/2 to determine the constants of integration A and B we obtain,

$$p = -\frac{6\eta}{h^3}\left(\frac{b^2}{4} - y^2\right)\left(\dot{e}\cos\phi + e\sin\phi\left(\dot{\beta} - \frac{\Omega_s + \Omega_b}{2}\right)\right) \qquad \text{(A2)}$$

The hydrodynamic force components due to this pressure distribution acting along (P_a) and normal (P_n) to the line of centres can be determined from,

$$P_a = -\int_{\phi_1}^{\phi_2}\int_{-b/2}^{b/2} pr\cos\phi\, dy\, d\phi$$

$$P_n = \int_{\phi_1}^{\phi_2}\int_{-b/2}^{b/2} pr\sin\phi\, dy\, d\phi$$

To identify the principle involved in short journal bearing analysis application via the Mobility Method consider a complete or non-cavitating film for demonstration purposes ($\phi_1 = 0$, $\phi_2 = 2\pi$). This gives,

$$P_a = \frac{\eta r b^3}{c^3} \int_0^{2\pi} \left[\frac{\dot{e}\cos^2\phi}{(1+\varepsilon\cos\phi)^3} + \frac{e\sin\phi\cos\phi}{(1+\varepsilon\cos\phi)^3} \left(\dot{\beta} - \frac{\Omega_s + \Omega_b}{2} \right) \right] d\phi$$

$$P_n = \frac{\eta r b^3}{c^3} \int_0^{2\pi} - \left[\frac{\dot{e}\sin\phi\cos\phi}{(1+\varepsilon\cos\phi)^3} + \frac{e\sin^2\phi}{(1+\varepsilon\cos\phi)^3} \left(\dot{\beta} - \frac{\Omega_s + \Omega_b}{2} \right) \right] d\phi$$

The well known Sommerfeld substitution gives (Booker (1965)),

$$\int_0^{2\pi} \frac{\cos^2\phi\, d\phi}{(1+\varepsilon\cos\phi)^3} = \frac{\pi(1+2\varepsilon^2)}{(1-\varepsilon^2)^{5/2}}$$

$$\int_0^{2\pi} \frac{\sin\phi\cos\phi}{(1+\varepsilon\cos\phi)^3} = 0$$

$$\int_0^{2\pi} \frac{\sin^2\phi\, d\phi}{(1+\varepsilon\cos\phi)^3} = \frac{\pi}{(1-\varepsilon^2)^{5/2}}$$

such that

$$P_a = \frac{\eta r b^3}{c^3} \left(\frac{\pi(1+2\varepsilon^2)}{(1-\varepsilon^2)^{3/2}} \right) \dot{e} \qquad \text{(a)}$$

$$P_n = \frac{\eta r b^3}{c^3} \left(\frac{\pi\,\varepsilon}{(1-\varepsilon^2)^{3/2}} \right) \left(\frac{\Omega_s + \Omega_b}{2} - \dot{\beta} \right) \qquad \text{(b)}$$

(A3)

Engine Tribology / C.M. Taylor (Editor)

Engine bearing design: Design studies, wider aspects and future developments

F.A.Martin

Consultant to T & N Technology Ltd.

1. INTRODUCTION

In the previous chapter by C.M.Taylor, the fundamental background and lubrication analysis of engine bearings was considered. It is intended to expand on this here, relating some of the basic principles to bearing design. Predictive techniques generally fall into two categories, rapid methods for day-to-day application engineering and rigorous methods for product development and diagnostic purposes. A design example for a typical automotive big end bearing is used to illustrate the various aspects of performance prediction (from rapid methods), including journal centre orbits, minimum film thickness, film extent map, oil flow and power loss.

With the increased availability of computing power more realistic assumptions about bearing conditions can be made using rigorous methods. These relate to oil feed features, oil film history (flow), connecting rod bearing deformation and main bearing load sharing. (A bibliography is supplied for further reading).

1.1. Notation

1.1.1. Dimensional quantities

Engine :

A_{cyl}	=	cylinder area	(m^2)
F_{cyl}	=	load due to cylinder pressure	(N)
F_{rec}	=	load due to reciprocating mass	(N)
F_{rot}	=	load due to rotating mass	(N)
L	=	connecting rod length	(m)
M_{rec}	=	reciprocating mass	(kg)
M_{rot}	=	rotating mass	(kg)
N	=	rotational speed	(rev/s)

P_{cyl}	=	cylinder pressure	(N/m^2)
P^*_{cyl}	=	maximum cylinder pressure	(N/m^2)
R	=	crank throw	(m)
α	=	crank angle	(°)

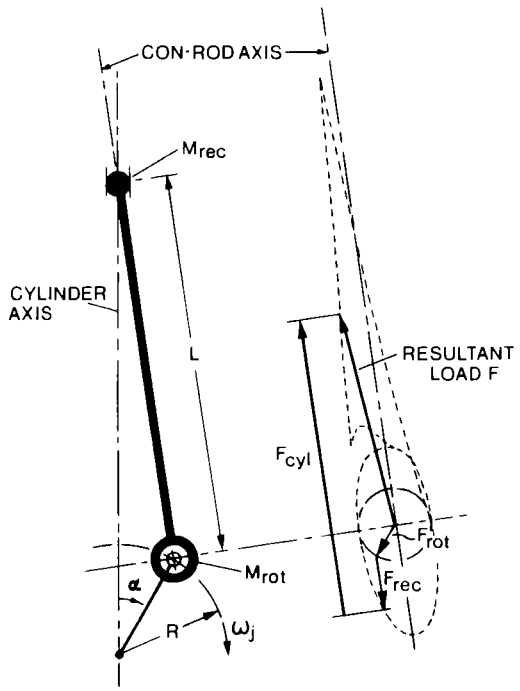


Fig.1 Engine system and bearing load notation

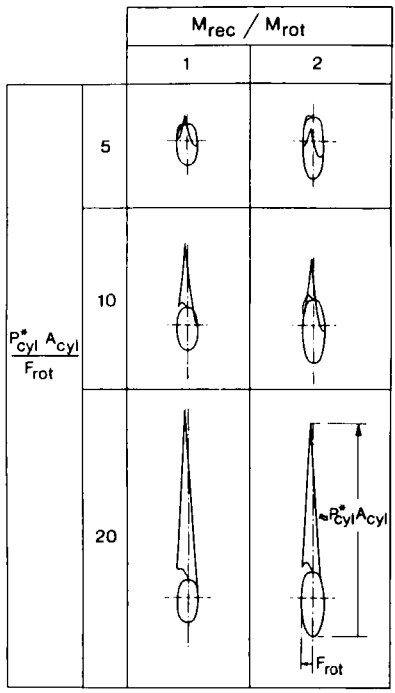


Fig.2 Big end bearing load notation defining load diagram shapes

The resultant instantaneous force acting on the connecting rod bearing is derived from the following vector components :-

$$F_{cyl} = P_{cyl}.A_{cyl} / [1 - (R/L)^2 \sin^2\alpha\,]^{0.5}$$

$$F_{rot} = M_{rot}\, R\, \omega_j^2$$

$$F_{rec} = M_{rec}\, R\, \omega_j^2\, [\cos\alpha + (R/L)\cos 2\alpha] / [1 - (R/L)^2 \sin^2\alpha]^{0.5}$$

The component F_{cyl} is directed along the connecting rod axis. The other two components are directed along the crank and connecting rod axes respectively.

Bearing :

b	= bearing length	(m)
c or c_r	= radial clearance	(m)
c_d	= diametral clearance	(m)
d	= bearing diameter	(m)
d_h	= feed hole diameter	(m)
e	= journal centre eccentricity	(m)
F	= bearing load	(N)
h_g	= film thickness at feed hole position	(m)
h_{min}	= minimum oil film thickness	(m)
h_{max}	= maximum oil film thickness	(m)
$H_{PRESSURE}$	= pressure power loss constituent	(W)
H_{SHEAR}	= shear power loss constituent	(W)
$H_{SQUEEZE}$	= squeeze power loss constituent (fictional)	(W)
$H_{TRANSLATORY}$	= translatory power loss constituent	(W)
H	= total power loss	(W)
P_f	= oil pressure at feed hole	(N/m^2)
P_{max}	= maximum hydrodynamic pressure	(N/m^2)
P_n	= specific load (F/bd)	(N/m^2)
Q_F	= film history flow	(m^3/s)
Q_H	= hydrodynamic flow (instantaneous)	(m^3/s)
Q_P	= feed pressure oil flow (rapid solution)	(m^3/s)
Q_{MARTIN}	= simulation of Q_{RBC}	(m^3/s)
Q_{RBC}	= Reynolds boundary condition flow (rigorous method)	(m^3/s)
r	= bearing radius	(m)
v	= journal centre velocity (note appropriate reference frame used) (rel. to x y frame or down squeeze path)	(m/s)
v_ξ	= squeeze path velocity component (parallel to load line)	(m/s)
β	= angle measured from F to v	(rad)
β_h	= oil hole position on journal	(°)

ϕ	=	angle measured from F to h_{min} position	(rad)
η	=	oil viscosity	(Ns/m^2)
θ	=	angle from h_{max}	(rad)
τ_j	=	Torque exerted on the oil film by the journal	(Nm)
τ_b	=	Torque exerted on the oil film by the bearing	(Nm)
ω or ω_j	=	journal angular velocity	(rad/s)
ω_b	=	bearing angular velocity	(rad/s)

1.1.2. Dimensionless design/operating parameters

$(F_{rot} / \eta N d b) (c / r)^2$ = rotating load number

P^*_{cyl} A_{cyl} / F_{rot} = cylinder load ratio

M_{rec} / M_{rot} = reciprocating mass ratio

P_{cyl} / P^*_{cyl} = normalized cylinder pressure

R / L = slider crank ratio

b / d = slenderness ratio

ε = eccentricity ratio (e/c)

2. DESIGN AIDS RELATED TO JOURNAL CENTRE MOVEMENT

An example for a typical automotive engine big end bearing is introduced to show the load pattern and journal orbit viewed from different reference points (observer axes): the significance and use of such orbits is explained. The same particular example is later used when discussing other aspects of bearing performance. The engine and bearing data, for a 4-cylinder 1300cc engine, is shown in Table 1. Typical variation of cylinder pressure with time (normalized to a maximum unit of one for further use in parameter studies) is shown in Fig. 3 .

Table 1
Engine and bearing data (particular example)

Engine data			Bearing data		
M_{rot}	= 0.32	(kg)	d	= 0..042	(m)
M_{rec}	= 0.48	(kg)	b	= 0.0168	(m)
R	= 0.036	(m)	c_d	= 0.00004	(m)
L	= 0.120	(m)	η	= 0.004	(Ns/m^2)
P^*_{cyl}	= 7.0	(MN/m^2)	P_f	= 0.4	(MN/m^2)
A_{cyl}	= 0.0043	(m^2)	d_h	= 0.006	(m)
N	= 4000/60	(rev/s)	β_h	= 30	(°)

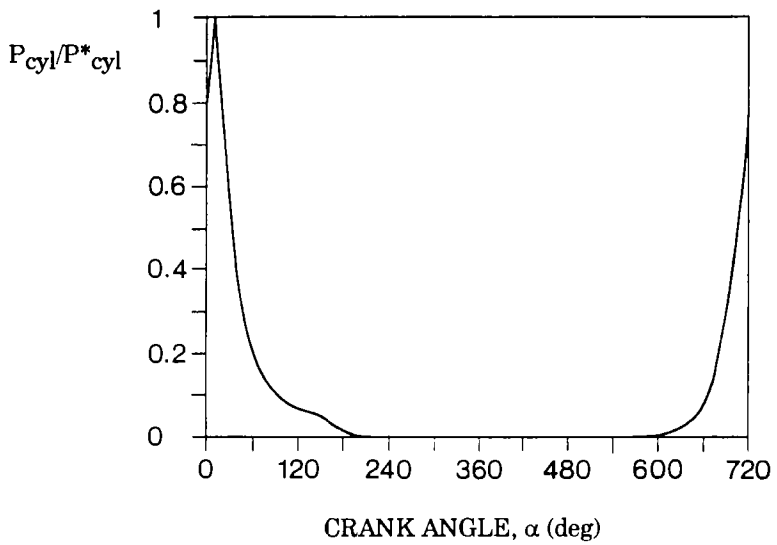


Fig.3 Normalized cylinder pressure (typical - used in example case).

A simplified two mass system is considered here, (see notation Fig.1). M_{rot} reflects the rotating portion of the connecting rod mass and M_{rec} reflects the remainder of the connecting rod mass plus that of the piston assembly. The two mass system is adequate for general parameter studies where one is seeking trends in performance. However, in practice a slightly more accurate method involving the moment of inertia of the connecting rod is generally used.

2.1. Journal Centre Orbit Within Clearance Circle

The journal centre orbits shown in Fig. 4 were computed using oil film dynamic characteristics for a finite bearing based on the curve fitted finite element data given by Goenka 1984. This data is available in a general 'mobility' (rapid solution) format so that journal centre orbits can be computed in a few seconds, rather than the several hours necessary for a finite element method (rigorous solution without the benefit of curve fit routines). Another finite bearing rapid solution is that of Moes, see Booker 1971. The minimum film thickness for both of these solutions will be much lower than that predicted by short bearing theory (another commonly used rapid solution).

Load diagrams and journal centre orbits 'observed' from three different reference axes are shown in Fig. 4 , (considering clockwise rotation of the shaft). The numbers on the load diagrams and journal centre orbits relate to crank angles (representing time).

2.1.1. Observed from load line. The upper orbit shown in the figure represents the journal position relative to the load line, the reference axis for applying mobility data. This type of orbit, however, is not usually produced in design studies (but is a means to an end); the computed output data generally being transformed to a more meaningful reference axis, either that of the bearing (observer on con-rod axis) or that of the journal (observer on crank pin axis). Whilst considering this orbit (i.e. with reference to the load line), the following observations are of interest.

i) The dotted curve on the left hand side in the clearance circle represents the steady load locus.

ii) The dotted curve on the right hand side represents the locus for a pure rotating load.

iii) The minimum and maximum film thickness positions for dynamically loaded bearings will occur on these locus lines (i) & (ii).

iv) Inside this locus (boundary) the film thickness will decrease (decay region).

v) Outside this locus (boundary) the film thickness will increase (recovery region).

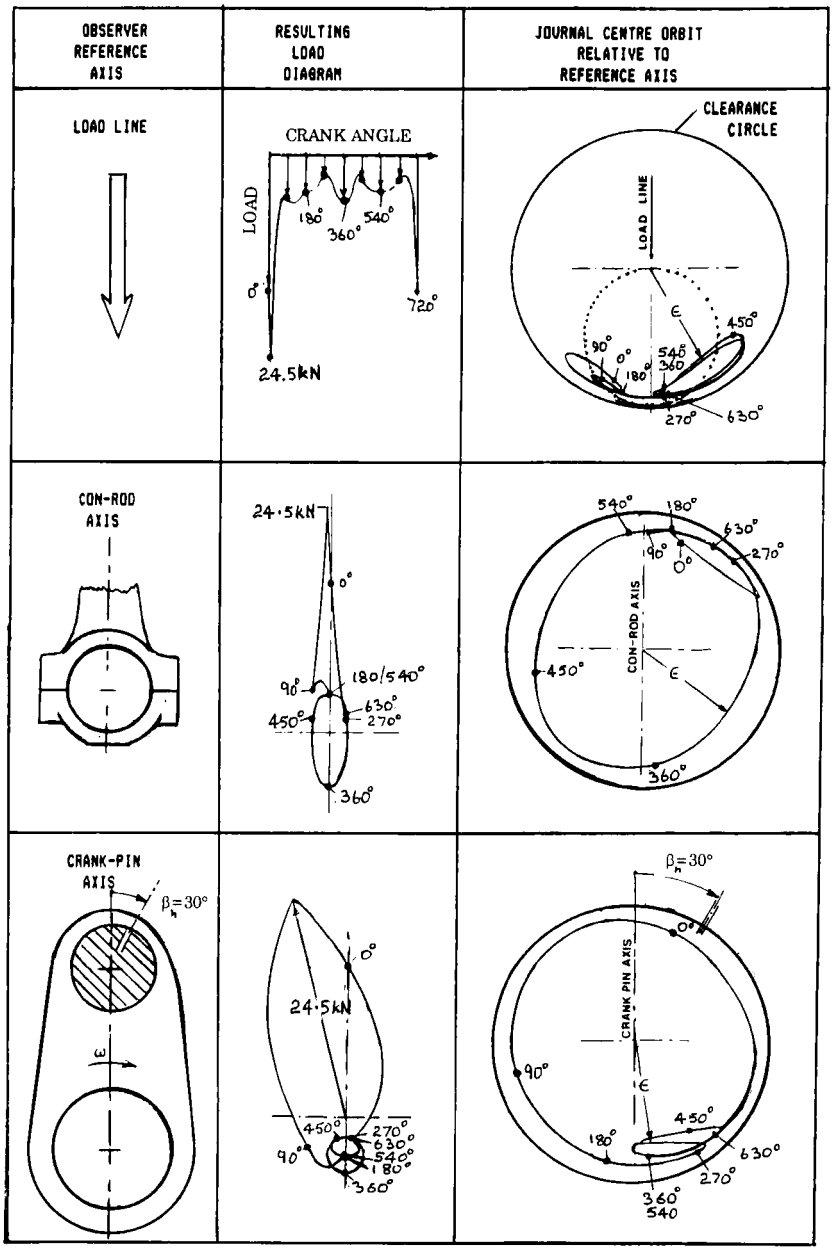


Fig. 4 Example big-end bearing case: journal centre orbits within the clearance circle, observed from different reference axes.

2.1.2. Observed from con-rod axis. In Fig. 4 the centre row shows the polar load diagram and journal centre orbit relative to the con-rod axis (ie. with an observer situated on and oscillating with the bearing). The loop in the load diagram is due to inertia forces and the remaining load is due to the cylinder pressure. The part of the journal centre orbit due to inertia loading is clearly identified by following the path 270° to 630° crank angle. The nearness of the orbit trace to the clearance circle is a measure of oil film thickness. The load at 0° (crank at top dead centre) is seen to be many times higher than the load at 180° and yet the minimum film thickness at 0° is larger than at 180°. The reason for this is that the load vector in the region of 0° is rotating in the opposite direction to the shaft which greatly enhances the load carrying capacity of the bearing.

Journal centre orbits for 'main' crank shaft bearings plotted in this way (with respect to bearing) are useful in assessing the best place for oil feed features (away from high load and thin film regions). For big end bearings, however, the oil feed is generally from a feed hole (or holes) emerging from the crank pin and therefore it is helpful to observe the load and orbit relative to the crank pin as shown at the bottom of Fig. 4.

2.1.3. Observed from crank pin axis. One can see from the orbit (lower diagram) that for most of the time the journal centre hovers around a small region of moderately thin films (see positions 270° to 630° crank angle). This occurs on the main bearing side of the crank pin about 150° from the top of the crank pin (in the direction of rotation). An oil feed hole, to maintain an adequate supply of oil for hydrodynamic action, would generally be in a predominantly thick film area away from this 150° region. A survey of hole positions (β_h) met in current practice (single hole case measured from the top of the crank pin in the direction of rotation) typically ranged from 25° to 70°. The hole position β_h for the example case is shown at 30° (Fig. 4) which is generally away from thin film regions for most of the time. Experimental evidence from the NEL bearing rig (Cooke 1983) shows that when simulating a large diesel engine big end bearing, the flows on a unit base were:-

Hole position β_h°	0	30	60	90
flow	1 unit	0.70	0.45	0.18

This trend in flow is also reflected when studying the change in general film thickness at the feed hole, when moving the feed hole to different positions of β_h in Fig.4. Also it should be noted that the oil hole should preferably not be placed in regions of continuing high developed oil film pressure and a 'film (or pressure) extent map', explained next, will be useful to make judgements on this.

2.2. Better Presentation of Data to Aid Designers

2.2.1 Film extent map. Film extent maps show what happens around the bearing during two revolutions of crank rotation. The information on them, and the realization of what can be obtained from them, has steadily improved over the years, Figs. 5 and 7 show part of that development.

Fig. 5 shows the film extent map for the example big-end bearing case (The "eye" of the con-rod shows direction of viewing). The horizontal axis gives a 'picture' of what is happening around the bearing surface circumferentially (follow the eyes) and the vertical axis represents changes in these events with time.

The film extent here is seen to cover a region of 180° corresponding to the hydrodynamic flow region. It should be noted that short bearing theory would give a 180° film extent. However for the finite bearing theory this is not strictly the case, although shown here for simplicity and general guidance.

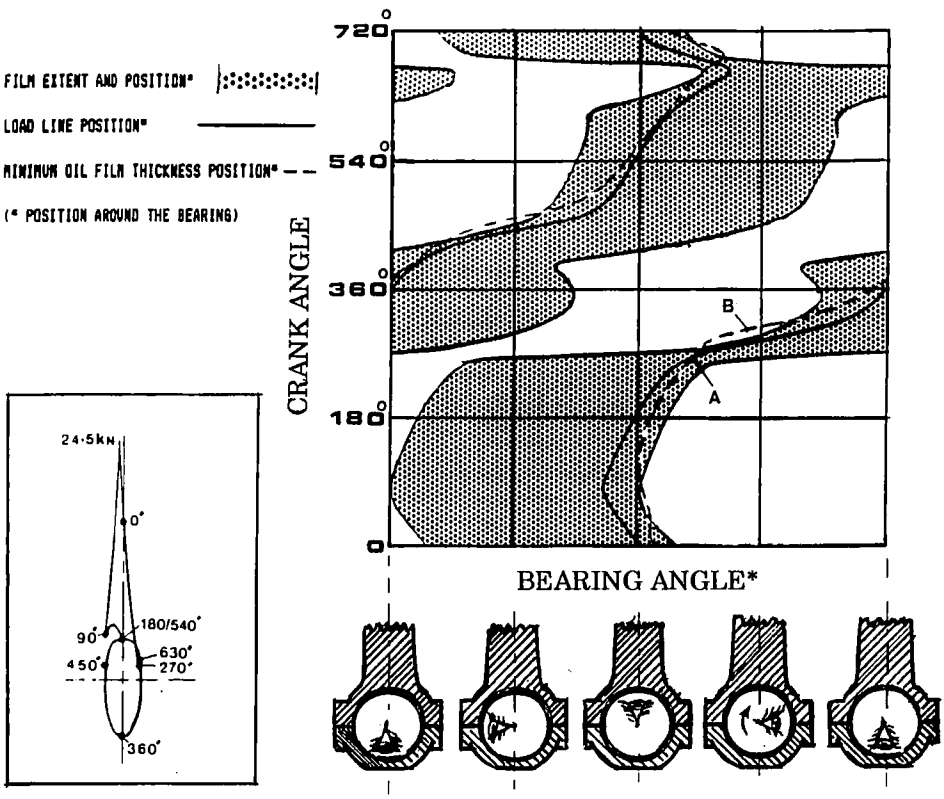


Fig. 5 Film extent map - Example big end bearing case.

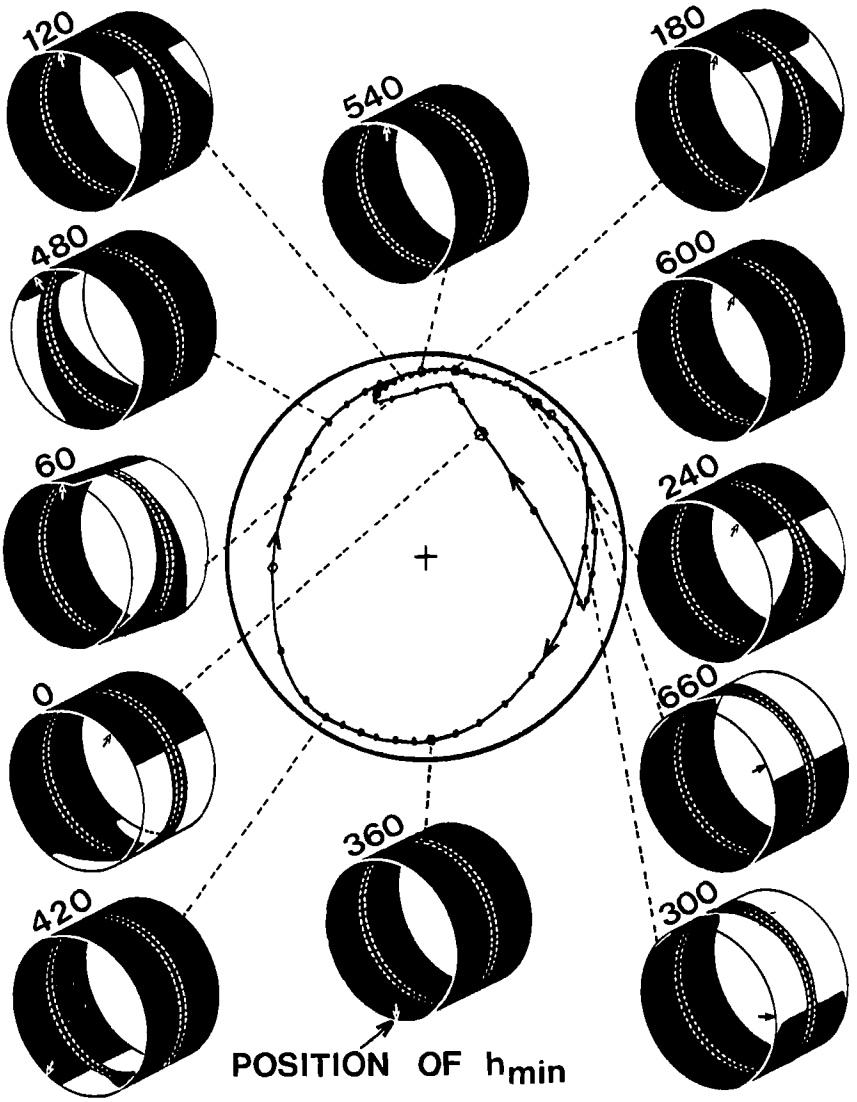


Fig. 6 Film extent for the VEB Mk III big end bearing study case (Martin 1985)

For other types of bearing with a circumferential groove, there is generally a copious flow of oil due to the pressurised feed groove and a complete 360° film extent is often present as illustrated in Fig 6. [Such grooving is generally associated with main bearings rather than the big end case (VEB MkIII engine) shown].

Returning to the example case in Fig. 5, the position of the load is shown by the full line within the load carrying film extent (dotted area). The trace of the load path can easily be identified by studying the accompanying load diagram. For example, looking on the rod half of the bearing (upwards), the load on the film extent map appears at 0/720,180 and 540° crank angle (consistent with the load diagram).

The dashed line on the map represents the journal centre path (i.e. the local minimum film position at the crank angle considered). When the journal centre position and load line are coincident (point A) then the journal is on the central squeeze path. When the journal centre path (dashed line) is outside the film extent region (point B) then the film thickness will be increasing and there may be a tendency towards negative film pressures. (NEL film pressure experiments on a circumferentially grooved bearing under dynamic loading showed such trends - reviewed by Martin 1983). These maps and film rupture intensity maps are useful in identifying and solving cavitation erosion damage problems in engine bearings; Garner, James, Warriner 1980 -(marine diesel example).

2.2.2. Film pressure map. Another option showing the angular extent of selected film pressure levels is generally more useful in identifying the active part of the film. This together with a trace of the oil hole position, Fig. 7, gives a valuable aid to the designer when assessing the best position for the crank drilling. This map (for the example case) shows the region in the bearing where the developed hydrodynamic pressure is greater than the feed pressure of 4.1 bar (at the emergence of the crank drilling). The oil hole should be placed so as to avoid this region if possible and also avoid local thin film and high pressure regions (judged by P_{max}, $P_{max}/2$ and $P_{max}/4$ zones).

The path of the crank drilling position, for β_h equal to 30°, appears satisfactory. With a β_h of 150° the crank drilling path would be subjected to hydrodynamic oil film pressures, generally greater than the feed pressure of 4 bar, during most of the cycle; this would be likely to restrict the flow from the crank drilling. Also with a β_h of zero the path would cut across the high pressure zone and (with a large hole) may derate the bearing locally, possibly causing bearing metal fatigue. This chart using a rapid finite bearing solution shows how to avoid this situation. A more rigorous full 2 D solution (inherently allowing for oil feed features) but taking hours rather than seconds to run, is capable of showing the increase in developed oil film pressure for such a situation which can then be more readily assessed, Conway-Jones, Martin, Gojon 1991. The range of β_h from 25 to 70° based on practice, appears to coincide satisfactorily with the conclusions in the theoretical study considered here. It may , however, be prudent to keep β_h greater than 30°. Some uses of these design aids (maps) are given above and further

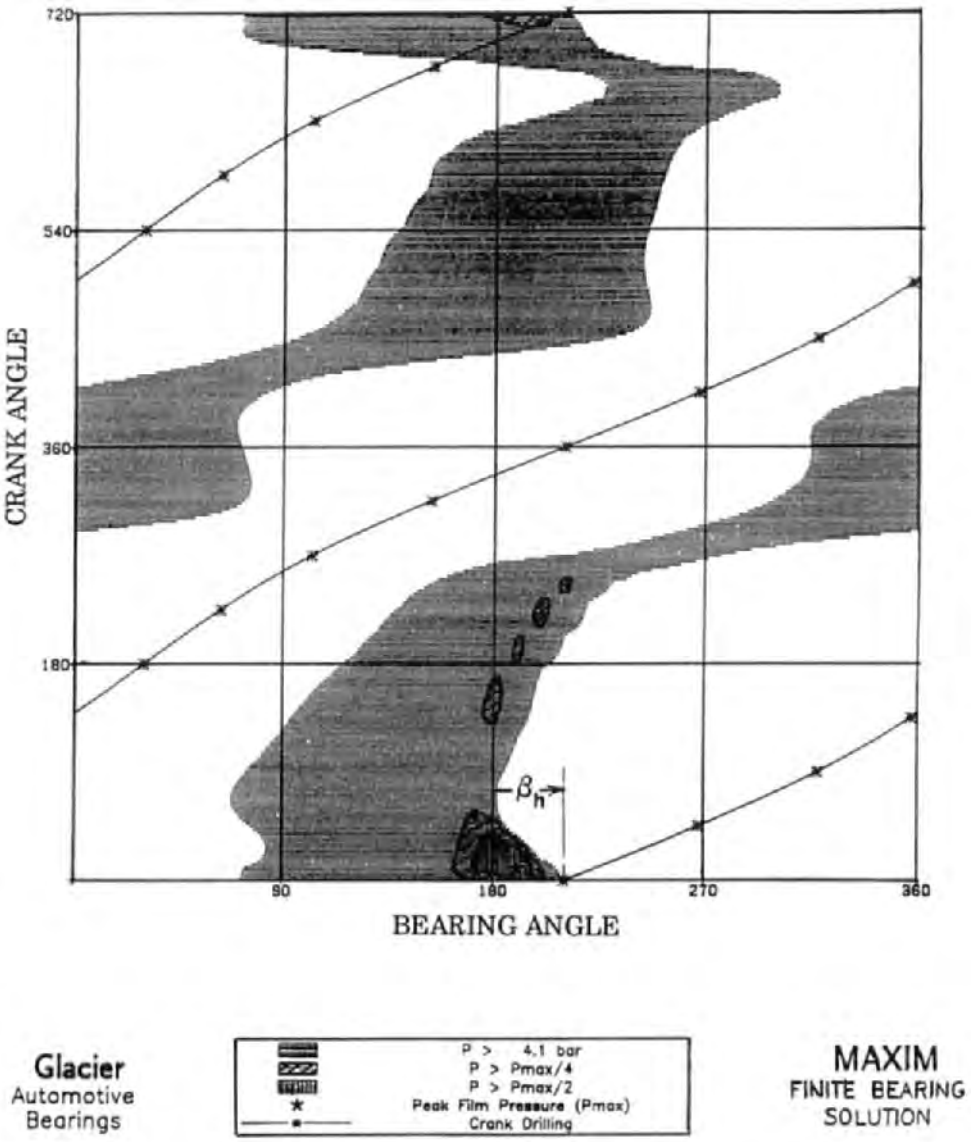


Fig 7 Film pressure map - big end bearing example case
4 cylinder 1300 cc engine at 4000 rev/min
(courtesy of The Glacier Metal Co Ltd/T & N Technology Ltd)

information on other applications can be found in the literature - Spikes and Robinson 1982 illustrate the connection between two film extent maps with the big end fed from the main bearing. Campbell, Gojon, Serre and Warriner 1986 show how these maps are used to examine the angular extent of the load concentration on the journal of a medium speed diesel engine and how improvements to the bearing performance can be made by better balancing (avoiding local thin films and high temperatures concentrated on a small area of the journal).

3. MINIMUM OIL FILM THICKNESS

The minimum oil film thickness during a complete cycle of operation is one of the most significant parameters on which to judge bearing performance, obtained directly from the maximum eccentricity ratio (as $h_{min} = C_r [1-\varepsilon_{max}]$). It is generally used as a comparator and represents a major factor in relating predicted performance with existing bearing experience on similar type engines. It is difficult to give precise values of minimum film thickness at which bearing damage might occur, as other factors such as high bearing temperature, misalignment, inadequate oil feed arrangements and adverse environmental conditions will all have an effect. Also the position and dwell time of the small film region may be important. Another reason why judgements on film thickness must generally be on a comparative basis is that different theories will give different answers. For instance the minimum film thickness predicted from the well established short bearing theory will be greater (possibly by 1.3 to 2.5 times) than the more accurate finite bearing theory, depending on operating conditions. As long as one associates the particular theory used with ongoing practical feed back of performance then one can build up a useful framework on which to make judgements on design.

Table 2.
Guidance on danger levels for film thicknesses predicted by short bearing film model for connecting-rod bearings.

	d (typical) mm (in)	h_{min} (dangerous) μm (μin)
Automotive (gasoline)	50 (2)	1.0 (40)
Automotive (Diesel)	75-100 (3-4)	1.75 (70)
Industrial (Diesel)	250 (10)	2.5 (100)

Booker 1979, after carrying out a survey with various bearing manufacturers, gives some guidance on danger levels for film thickness. These, shown in Table 2, are for use with short bearing theory.

3.1. Con-rod Big End Bearing Film Thickness Considering The Inertia Load Loop

Martin and Booker 1967 found (using short bearing theory) that the minimum film thickness during the inertia load loop, Fig. 8, was generally smaller than that influenced by the firing load, even when the firing load F_2 was greater than the inertia load F_1. Under these conditions it was possible to predict the minimum film thickness ratio h_{min}/C_r, considering the inertia load loop alone for many bearing cases. The design aid shown in Fig. 9 summarizes these results, where the minimum oil film thickness ratio is found knowing the reciprocating and rotating mass parameters. The effect, on film thickness, of changing any of the variables can be readily assessed from this chart (applicable for cases where $F_2 < 5F_1$).

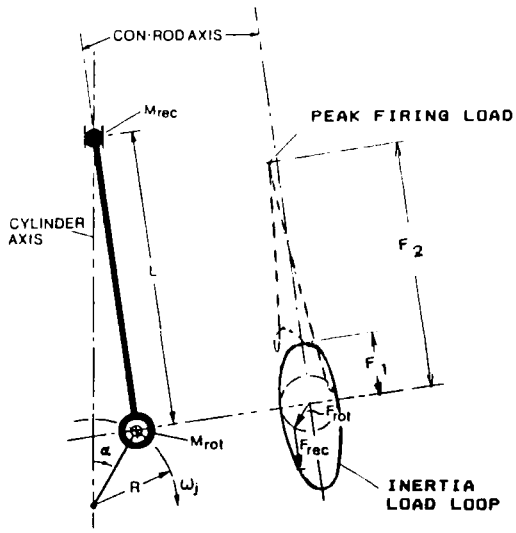


Fig 8 Inertia load loop - from reciprocating and rotating forces

3.2. Con-rod Big End Bearing Film Thickness Considering the Full Load Diagram and an Improved Film Model

A more accurate model (based on Goenka 1986 finite element bearing data in Mobility form) has been used in a parameter study to develop the design chart, Fig. 10. This shows film thickness based on finite bearing theory and also considers the cylinder pressure term. The use of this chart is illustrated by the example case where b/d equals 0.4, the F_{rot} term equals 10, the P^*_{cyl} term is 15 and M_{rec}/M_{rot} is 1.5. The lower and upper extremes of each band width for the P^*_{cyl} terms represent values of M_{rec}/M_{rot} equal to 1 and 2 respectively. The resulting h_{min}/C_r (by interpolation) is seen to be about 0.05 so with a radial clearance of 0.00002 the minimum film will be the product of these i.e. 1μm.

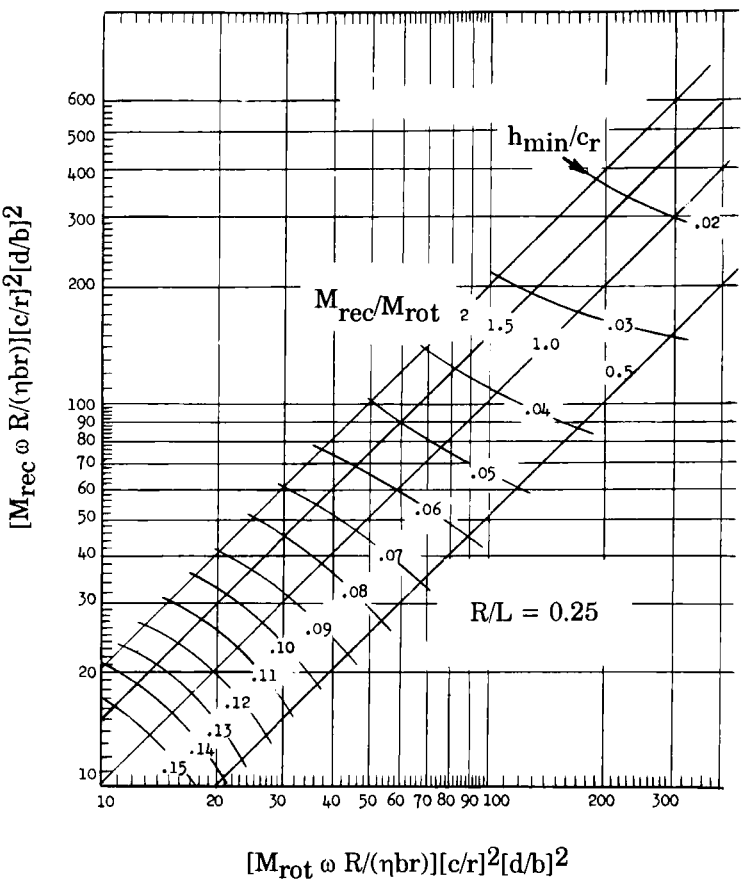


Fig 9 Minimum oil film ratio (inertia loading only) for big end bearings (based on short bearing theory)

Figs.9 and 10 are useful to assess the effects of all the pertinent variables and compare new designs with existing engine bearing performance. The minimum film thickness generally occurs at the start of the inertia load loop after firing (at about 270° crank angle). The film thickness at this point is influenced by the previous path of the journal centre (leading up to this point) which is affected by the cylinder pressure pattern (Fig. 3) and the load diagram shapes (Fig. 2). The size of the load diagram used in Fig. 10 is defined by the rotating load term involving F_{rot}.

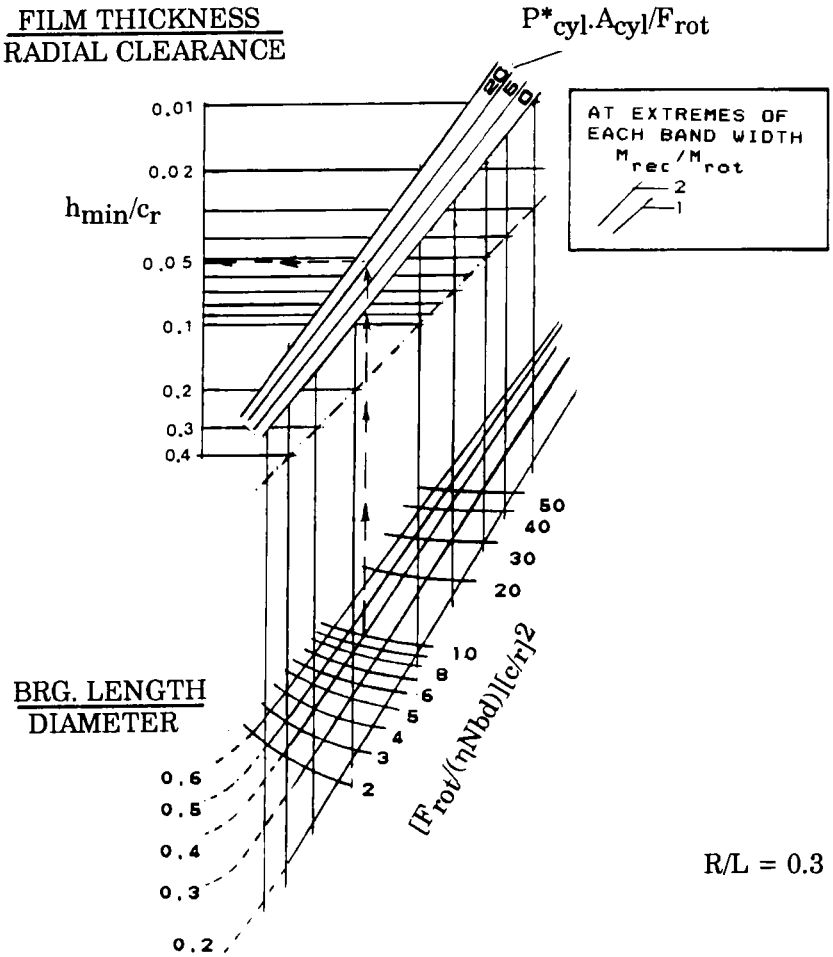


Fig. 10 Minimum oil film ratio (parameter study)
for big end bearings (based on finite bearing theory)

The zero P^*_{cyl} band in Fig.10 represents the effect of considering inertia load alone and corresponds with that of Booker 1979 who uniquely relates ε_{max} with ε_0 where Booker defines ε_0 as "the steady-state response to a steadily rotating inertia load applied to a steadily rotating journal in a non-rotating sleeve". An interpretation of Booker's relationship between ε_{max} and ε_0 (read from his graphs) is shown in Fig. 11 in terms of h_{min} and h_0. This relationship generally confirms the overall trend shown in the guidance chart Fig. 10 considering the effect of different values of P^*_{cyl}. (Note $h_{min}/c_r = 1 - \varepsilon_{max}$)

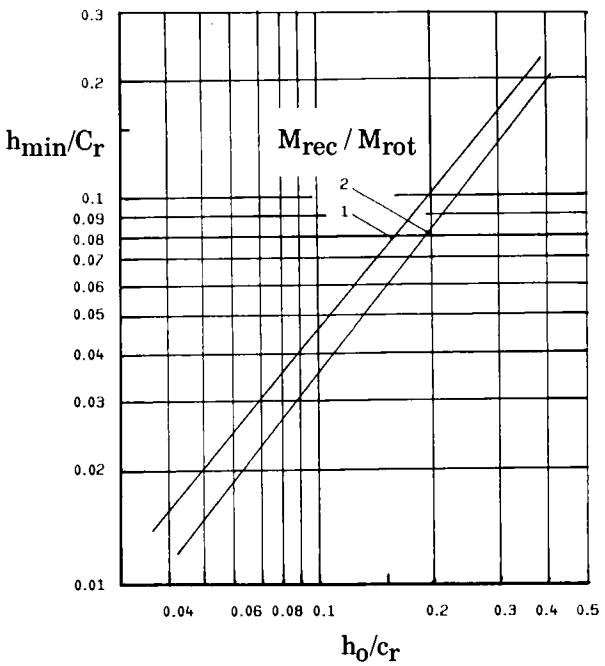


Fig. 11 h_{min}/C_r (inertia loading alone and R/L = 0.30) versus h_0/C_r (for pure rotating load, synonymous with steady load and R/L = 0)

3.3. Considering Oil Feed Features when Predicting Film Thickness from Rapid Methods.

Main bearings generally have partial or full 360° grooves which feed one or more adjacent big end bearings via the crank drillings. Sometimes an intermain bearing is completely independent of feeding any big end bearing in which case it may be fed by a single hole. When calculating film thickness from rapid methods (e.g. Goenka 1984 curve fit finite bearing solution), these oil feed features can be allowed for by considering the following 'rules': a) If the lubricant is supplied via an oil hole (away from high load regions) the bearing has to be considered plain. b) For a circumferential groove, the bearing can be treated as two plain bearings sharing the load. c) For a partial circumferential groove (which usually extends about 180°), an approximation can be used whereby the bearing is treated as a full width plain bearing when the load line is in the plain part, and as a circumferentially grooved bearing when the load line is in the grooved part. This approximation is less accurate for grooves of small circumferential extent. Film thickness results and journal centre orbits, using these 'rules', are generally similar to those obtained from the rigorous (not film history) finite bearing solution. (see Jones, Lee, Martin 1982).

4. OIL FLOW PREDICTIONS

Oil flow from an engine bearing is an important factor in the design procedure. It is required in the heat balance to determine the operating temperature and oil viscosity, and is helpful when estimating oil pump capacity. Unfortunately it is a very difficult parameter to predict accurately.

Rigorous methods are available (Jones 1983, Paranjpe and Goenka 1990) which give reasonably accurate predictions of the total flow through the bearing but often take several hours to compute for one bearing case. Therefore, attempts have been made to simulate these results using rapid methods so that the equivalent of the rigorous results can be predicted in just a few seconds. This involves considering the hydrodynamic flow Q_H caused by the shaft rotation and the resulting film pressures, together with the feed pressure flow Q_P the direct result of oil being forced through the bearing by the supply pressure. The suitable combination of Q_H and Q_P depends on the type of oil feed configuration used (e.g. circumferential groove, partial groove, oil hole) and the accuracy required. Let us examine these flows in more detail.

4.1. Hydrodynamic Flow Q_H

The hydrodynamic flow Q_H used in the rapid methods is independent of the oil feed pressure and is dependent on the oil leaving the load carrying film. Consider the steady load case, Fig. 12, as an analogy. The journal centre O_j is displaced from the bearing centre O_b due to the load W by an amount e (e divided by radial clearance gives the eccentricity ratio ε).

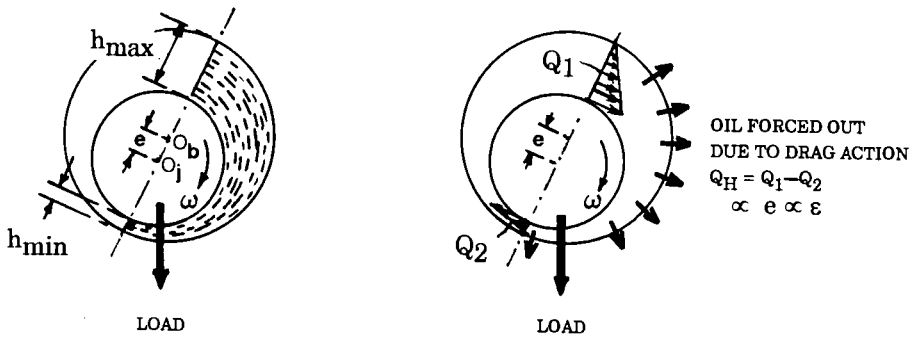


Fig. 12 Hydrodynamic flow due to wedge action under steady load conditions

For a steadily loaded bearing the film from h_{max} to h_{min}, Fig.12, forms a wedge of oil which is then forced out due to the drag action of a rotating shaft. For a full 180° film, the drag flow Q_1 at h_{max} (represented by the circumferential linear velocity gradient) is proportional to h_{max} and the flow Q_2 at h_{min} is proportional to h_{min}. The hydrodynamic flow leaving the bearing is proportional to the

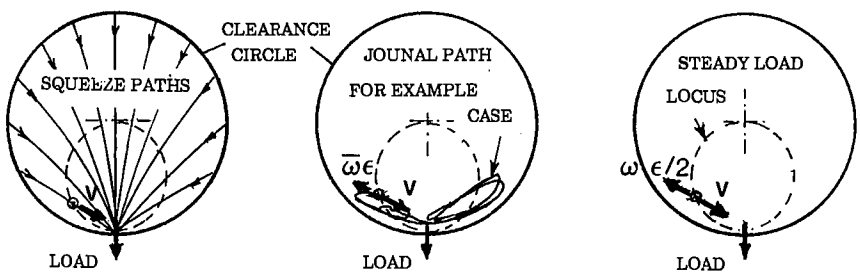
difference between these flows (i.e. proportional to ε). For the steady load, short bearing solution, the hydrodynamic flow in dimensionless terms is given by

$$Q_H / b\,d\,N\,C_d = \pi\,\varepsilon/2 \qquad (1)$$

This is discussed further in section 4.1.2 considering the general dynamic case based on squeeze velocities. For finite bearing theory, the velocity gradients in Fig 12, are slightly modified, allowing for the film pressure effect on oil film velocity.

4.1.2. General dynamically loaded bearing case - Hydrodynamic flow Q_H

The hydrodynamic flow Q_H for a dynamically loaded bearing may be obtained by considering what happens on a Mobility Map. In a previous chapter it was shown that the formation of the load carrying film is dependent on the velocity component V directed along the squeeze path. The other velocity component ω'ε, around a whirl path, does not contribute to the load but to the resultant velocity of the journal centre path. The squeeze path concept of a non rotating journal, Fig.13a, with known velocity V (from Mobility values) will define the oil film and flow characteristics anywhere in the clearance space. What may seem surprising is that this is unique irrespective of whether we are considering a non rotating journal and squeeze (Fig.13a), a journal centre orbit (Fig.13b) or a steady load case (Fig.13c).



(a) NON-ROTATING JOURNAL (b) DYNAMIC LOAD (c) STEADY LOAD

Fig.13 Same 'squeeze' velocity for given load and journal position

In all the three cases shown the dimensionless velocity component V (i.e. velocity/radial clearance) is given by:-

$$V = [(F/\eta bd)(C_d/d)^2].M \qquad (2)$$

where M is the Mobility value at the journal position under consideration and F the bearing load. The hydrodynamic flow Q_H is given by the displaced oil while the journal centre travels along a squeeze path as shown in Fig.14.

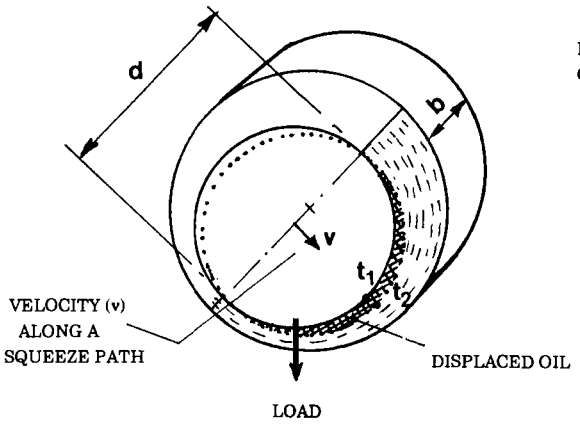


Fig.14 Hydrodynamic flow on squeeze action basis.

For all cases (including the three in Figs 13(a) (b) and(c)) Q_H is simply equal to bearing projected area times the velocity along the squeeze path. i.e.

Hydrodynamic flow Q_H = v.b.d (3)

[In the steady load case (on the steady load locus) the squeeze and whirl velocities (see Fig.13c) are equal and opposite in order to maintain its position in the clearance space. Hence $V = \omega\varepsilon/2$ or $2v/C_d = 2\pi N.\varepsilon/2$. Substituting for v in equation (3) gives Q_H = v b d = $[(\pi N C_d \varepsilon /2) b d]$ resulting in the same equation as when considering wedge action alone. This illustrates the uniqueness of the basic hydrodynamic (squeeze) flow equation (3) which aptly describes the steady load case in the realm of the general dynamic case].

A general equation for Q_H at any instant during cyclic loading (at any crank angle) and at any position in the clearance space is given in terms of Mobility number (based on equations (2) and (3)) :-

$$(Q_H \eta)\,(d/b) / (c_r^3 P_n) = 4\,M \qquad (4)$$

where P_n equals F/bd and where the Mobility data M is for the actual b/d (for a circumferentially grooved bearing, b is taken as the bearing land width).

It should be noted that this equation is strictly for a 180° film extent as considered in the short bearing theory (Booker 1965/79) . Finite bearing data (Goenka 1984) may be used in a similar way, apparently with little loss in accuracy.

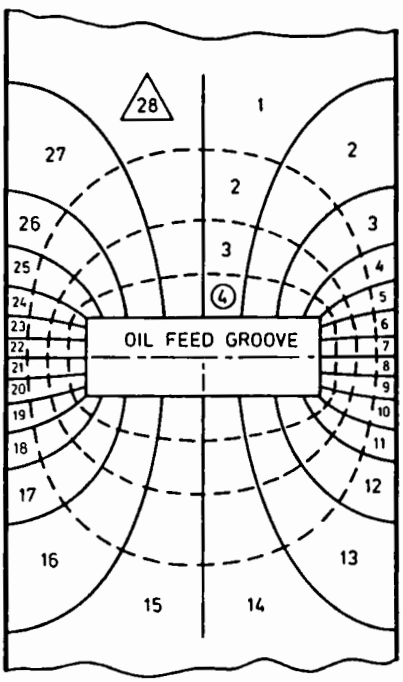
4.2. Feed Pressure Flow Q_p

The feed pressure flow Q_P is the flow due to feed pressure at the bearing feed groove or hole (considering a stationary journal and bearing). The basic concept of feed pressure flow is that of viscous flow through thin slots where, for a rectangular slot, the flow Q_P is proportional to film thickness cubed h^3, the width of the flow path, and the pressure difference P_f along the length of the flow path: it is inversely proportional to viscosity η and the length of the flow path. For a parallel flow path the flow in dimensionless terms is given by:-

$$Q_P \eta / h^3 P_f = (1/12) \text{ x width of flow path / length of flow path} \qquad (5)$$

In a bearing situation non-parallel flow paths are considered, where the lines of flow and lines of constant pressure form curvilinear 'squares' as shown in Fig.15. For a concentric journal, the radial clearance c_r replaces h and the flow is a function of the number of curvilinear squares n_L along the flow paths and number of squares n_w along constant pressure lines,shown in Fig.15.

Equivalent WIDTH OF FLOW PATH / LENGTH OF FLOW PATH $= n_w / n_L$



$$Q_P \eta / h^3 P_f = (1/12)\, n_w / n_L$$
$$= (1/12)\,(28/4)$$
$$= 0.58$$

Fig.15 Example showing constant pressure lines and flow path for an oil feed groove, forming curvilinear squares.

4.2.1. For Circular feed holes. The same principle of geometric plotting applies to circular feed holes (appropriate to some main bearings and also to the feed hole emerging on the crank-pin in connecting rod bearings.) The ratio of 'width of flow path' completely around the feed hole divided by the 'length of flow path' (considering maps of curvilinear squares) is related to the feed hole diameter d_h and the bearing length b. For a small diameter feed hole and an eccentric bearing the radial clearance term, for this case, is replaced by the local film thickness h_g at the feed hole position. The general flow equation for a circular feed hole is :-

$$(Q_P \eta)/(h_g^3 P_f) = 0.675 (d_h/b + 0.4)^{1.75} \quad (6)$$

Equation (6) compares well with experimental results (Hirano, Shodai 1958), (Martin, Lee 1983).

4.2.2. For rectangular grooves. Main bearings sometimes have partial circumferential grooves extending at least 180° around the bearing. For such cases it is necessary to have an accurate curve-fitted equation for Q_P based on computed numerical data where flow is related to changes in film thickness adjacent to the groove. The following general comprehensive curve-fitted equation (7) considers any groove extent, any groove aspect ratio, journal position and groove position. (see Fig. 16 for notation).

$$Q_P \eta / (c_r^3 P_f) = [f_1/6] [(1.25 - 0.25a/b) / (b/a - 1)^{0.333}] + [f_2/6] [(d/b) / (1 - a/b)] \quad (7)$$

where $f_1 = [(1 + \varepsilon \cos \theta_1)^3 + (1 + \varepsilon \cos\theta_2)^3]$

and $f_2 = [\theta + 3\varepsilon \sin\theta + \varepsilon^2(1.5\theta + 0.75\sin2\theta) + \varepsilon^3(\sin\theta - 0.333\sin^3\theta)]_{\theta_1}^{\theta_2}$

If used for a concentric journal then $f_1=2$ and $f_2=2w/d$)

Equation (7) can be used with confidence for any groove extent up to 270° for bearings of length /diameter ratio up to 0.5.

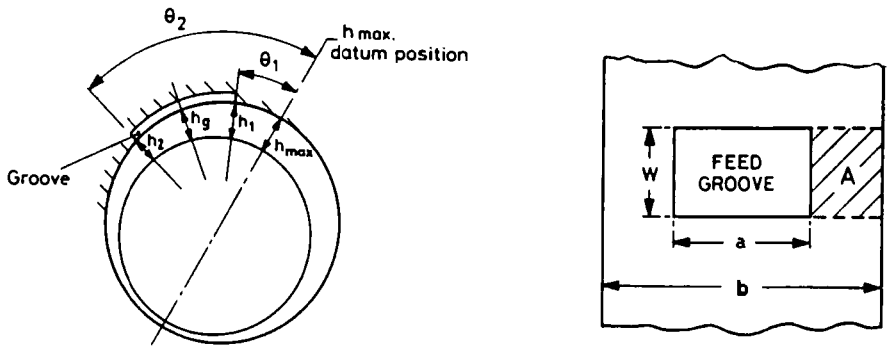


Fig. 16 Rectangular groove notation

4.2.3. For a complete (360°) circumferential groove

$$Q_P \eta/(C_r^3 P_f) = \pi d\,(1+1.5\varepsilon^2)/[3(b-a)] \qquad (8)$$

where b is the overall length of the bearing and a the axial width of the groove. It should be noted that all the above Q_P equations are for the feed pressure effect on flow independent of sliding velocities.

4.3. Improved Finite Bearing Models - Rigorous Solutions

The object of considering the rapid solutions which give flow components Q_P and Q_H (apart from an understanding of basic concepts) was to use them to simulate the results of the more rigorous solutions which are very time consuming on the computer (several hours run time rather than seconds for quick methods). More attention to the development of a simulation process has been made over recent years (Jones, Lee, Martin 1982),(Clayton, Taylor 1988) and (Martin, Stanojevic 1990) by studying the predicted performance results of many bearing cases using rapid and rigorous solutions.

4.3.1 Finite bearing model using Reynolds boundary conditions. Here an integrated flow from the bearing is calculated which includes both the effect of the hydrodynamic action and the oil feed pressure. Allowance is also made for the derating effect of oil feed features when they are within the load carrying film. At any instant in time the actual loaded film will consist of a certain volume of oil, however no allowance is made to check that there is an adequate source of oil to enter that load carrying film space prior to the time it is required. Situations can therefore arise where there is inadequate oil in the load carrying film leading to smaller oil film extents and in turn affecting the oil film chacteristics such as the load capacity and oil flow.

4.3.2. Finite bearing model considering oil film history. Oil film models have been developed which can keep track of the transport of oil within the bearing. These calculate the degree of filling of the clearance space at any point as the solution proceeds. Generation of hydrodynamic pressure is then inhibited in any region of the bearing where the clearance space is not completely filled, thereby taking account of the history of the oil film (Jones 1983). In this model the oil entering the bearing at any instant is not generally the same as the oil leaving the bearing at that instant, although the averaged flow throughout the load cycle must be the same (averaged oil in and out).

4.4. Simulation Of Rigorous Solution By Rapid Methods

The feed pressure flow Q_P may be used in a basic framework to translate results from rigorous solutions to more generally usable rapid prediction procedures. For instance in engine bearing studies experience has shown that:-

Flow (rigorous solution) / Q_P(rapid solution)
is a function of
Q_H(rapid solution) / Q_P(rapid solution)

From experience it has been established that the instantaneous flow based on Reynolds boundary conditions can be related to the instantaneous (rapidly derived) Q_H and Q_P for many groove configurations, including the single hole and partial circumferential groove. The empirical equation for instantaneous flow (considering single groove configurations) is given by :-

$$Q_{MARTIN}/Q_P = 1 + Q_H/Q_P - 0.3\,(Q_H/Q_P)^{0.5} \qquad (9)$$

This equation generally simulates the total flow for Reynolds boundary conditions (after averaging) and also gives good trends in the variation of outflow with time. This is shown in Fig.17 for the big end connecting rod bearing example considered previously, note the closeness of the Reynolds boundary condition (RBC) curve and the Q_{MARTIN} curve throughout the whole 720° of crank angle. There is ample evidence to suggest that this equation gives a good guide to flow based on RBC for other types of groove arrangements (Jones, Lee, Martin 1982). The more realistic cycle averaged oil film history flow Q_F, Fig. 17, is seen to be much less than the averaged value of Q_{MARTIN}. Generally bearings fed by a small hole are more likely to be influenced by film history effects (flow reduction) than ones with a copious supply (as from a full circumferential groove).

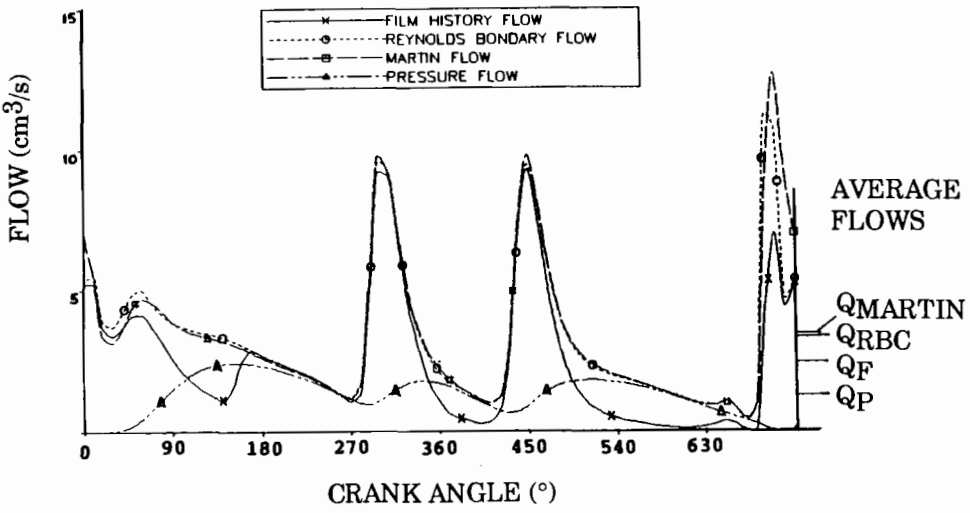


Fig. 17 Comparison of flows for the automotive bearing example case.

For big end bearings Martin, Stanojevic 1991 suggest the following form of equation to simulate film history results :-

$$Q_F/Q_P = [Q_{MARTIN}/Q_P]^S \qquad (10)$$

where S is typically 0.7 to 0.9, when $P^*_{cyl}\ A_{cyl}/F_{rot}$ is greater than 10. When calculating the flow the hole size should include chamfer effects where the hole emerges in the crank pin. Allowance should also be made for any additional feed pressure due to inertia effects on the column of oil in the crank drilling. A sensitivity chart Fig. 18, for big end connecting rod bearings, shows the general effect of change in bearing and engine system parameters on 'film history flow'. It gives a quick overall view of trends (where $Q_H > 0.7\ Q_P$). The dashed line shows how such charts are used. If the feed hole size is increased by 50% (a ratio of 1.5:1) then the resulting film history flow would be increased by 20% (a ratio of 1.2:1). Apart from engine speed, the other engine system parameters have little influence on oil flow. Considering the bearing parameters, clearance can be seen to have a major effect on flow .

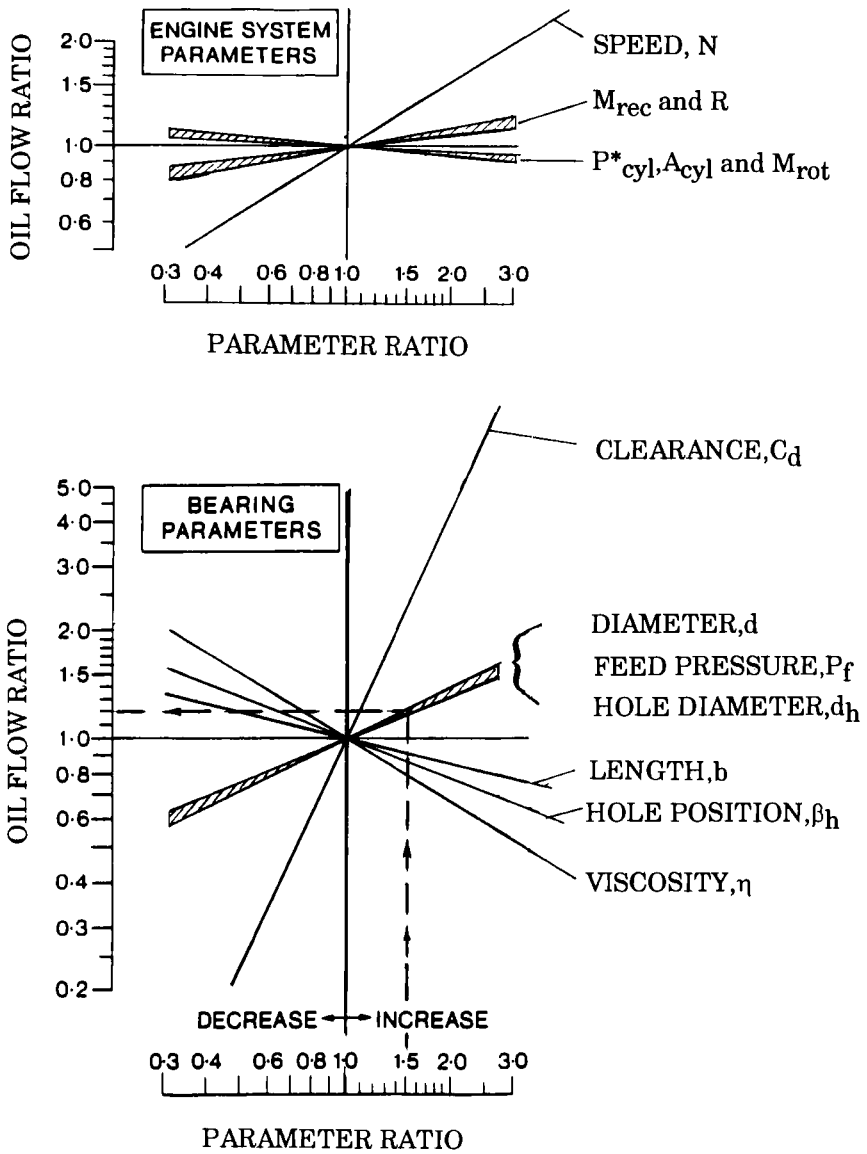


Fig. 18 Typical oil flow sensitivity chart based on film history flow for a big end bearing fed by a single hole.

If several variables are changed simultaneously then the resulting change in oil flow will be the product of the resulting individual ratios.

5. BEARING FRICTION AND POWER LOSS

In the automotive field it has been estimated that a 10% reduction in the direct mechanical losses within an engine gives perhaps 1 to 3 % improvement in fuel consumption. Such savings are very significant when viewed on a worldwide basis, in both economic and conservation terms. A typical energy distribution is shown in Fig.19 for a passenger car engine under part throttle conditions where one can see the 'Engine Mechanical Friction' as a proportion of the indicated engine output.

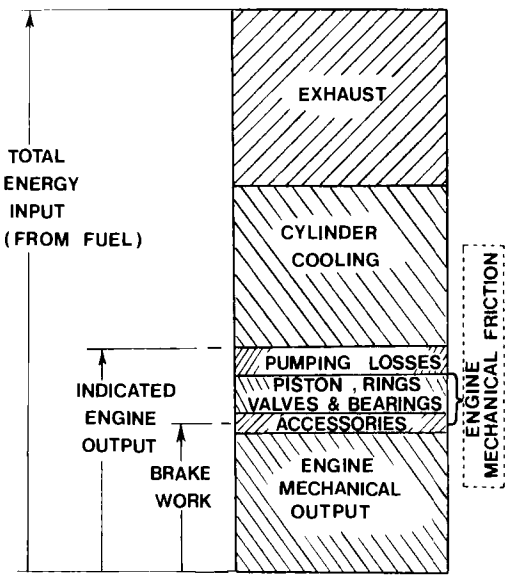


Fig.19 Typical energy distribution in an automotive engine (gasoline engine - part open throttle)

The mechanical friction losses in an engine are developed in the piston (skirt) and rings, the valve system, the bearing system, and those accessories directly associated with the engine. According to Lang 1982, friction losses which could be improved or optimized tribologically amount to about 72.5% of the mechanical losses, consisting of -

valve train	6%
connecting rod bearings	10%
main bearings	12.5%
piston rings	19%
pistons	25%

Many manufacturers, particularly in Japan (e.g. Hisatomi, Hiroshi 1982) and USA, (Kovach, Tsakris, Wong 1982) have gained improvements in fuel economy by reducing the mechanical friction losses within the engine by careful re-design. Further information is given in a comprehensive review on friction in internal combustion engine bearings - Martin 1985.

5.1. Power Loss in Engine Bearings

The hydrodynamic power loss in engine bearings is required for estimating the bearing operating viscosity through a heat balance routine (and establish its load capacity). It may also be used directly in studies of friction minimization with resulting economies in fuel consumption. Care should be taken however, as usually with engine bearings, reducing friction may result in less reliable performance.

5.1.1. General friction torque and power loss equations for the dynamic load case. One of the most comprehensive set of equations for predicting power loss in bearings is given by Booker, Goenka, van Leeuwen 1982 in vector and cross-product format. Power loss constituents (Martin 1985), based on this equation, are given below in scalar form relating to the geometric notation and xy frame given in Fig. 20. These power loss equations predict instantaneous losses (i.e. at a single crank angle condition). The total power loss for the bearing is obtained from cycle averaged values (as shown in Fig. 21 later).

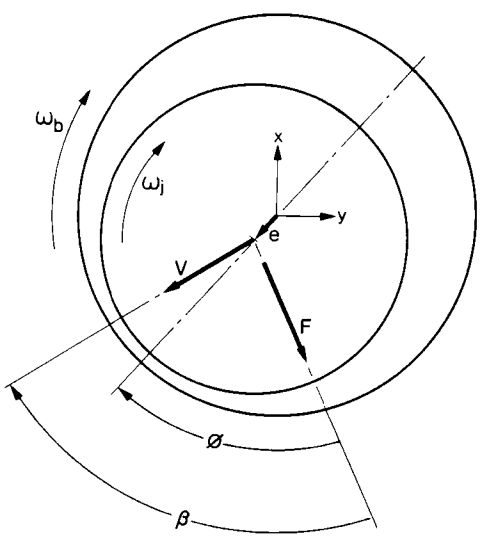


Fig. 20 Notation used for power loss equations

The power loss constituents are given by:-

a) SHEAR (or Couette) CONSTITUENT

$$H_{SHEAR} = \omega_j [\tau_j] + \omega_b [\tau_b]$$

$$= \omega_j [(\eta r\, {}^3 b/c)\, J_1{}^{00} (\omega_j - \omega_b)] + \omega_b [-(\eta r\, {}^3 b/c)\, J_1{}^{00} (\omega_j - \omega_b)]$$

$$= (\eta r\, {}^3 b/c)\, J_1{}^{00} (\omega_j - \omega_b)^2 \qquad (11)$$

where $J_1{}^{00}$ allows for the film extent and change in film thickness around the bearing.

$$J_1{}^{00} = c \int_{\theta_1}^{\theta_2} (1/h)\, d\theta = \int_{\theta_1}^{\theta_2} [1/(1+ \varepsilon \cos\theta)]\, d\theta$$

For a complete 2π film extent, which often occurs in a centrally grooved bearing (see Fig.6) $J_1{}^{00} = 2\pi / (1 - \varepsilon^2)^{0.5}$

For a 'π' film extent this value will not necessarily be halved as it is dependent on the film extent relative to the minimum film position (see Martin 1985; also Table of Bearing Integrals, Booker 1965).

b) PRESSURE CONSTITUENT

$$H_{PRESSURE} = \omega_j [F\, e \sin\varphi]/2 + \omega_b [F\, e \sin\varphi]/2$$

$$= [(\omega_j + \omega_b) / 2]\, F\, e \sin\varphi \qquad (12)$$

c) TRANSLATORY CONSTITUENT

$$H_{TRANSLATORY} = F\, v \cos\beta \qquad (13)$$

The pressure constituent of power loss at some instance in the load cycle may appear as a negative value. A more rational approach, combining the shear and pressure constituents, will give all positive constituents. Furthermore this combination relates to work done due to journal rotation and the remaining term relates to work done due to the translatory movement of the journal centre, relative to the xy frame, Fig. 20.

Table 3.
Summary of alternative power loss equations

COMMONLY USED TERMINOLOGY FOR INSTANTANEOUS POWER LOSS
(Martin 1985 based on Booker, Goenka and van Leeuwen 1982)

POWER LOSS	=	"SHEAR" CONSTITUENT ('COUETTE')	+	"PRESSURE" CONSTITUENT	+	"TRANSLATORY" CONSTITUENT
H	=	$\eta r^3 J_1^{00} (\omega_j - \omega_b)^2$	+	$[(\omega_j + \omega_b)/2]F e \sin\varphi$	+	$F v \cos\beta$
		Rate of work done due to angular 'rotation' of journal and bearing				Rate of work done due to 'translation' of journal centre (ref. xy frame)

ALTERNATIVE GROUPS AND TERMINOLOGY - 1
(Using Mobility data - Martin 1985)

POWER LOSS	=	"SHEAR" CONSTITUENT	+	"SQUEEZE" CONSTITUENT
H	=	$\eta r^3 J_1^{00} (\omega_j - \omega_b)^2$	+	$F v_\xi$
		From differences in journal and bearing rotation (relative to any 'observer')		'Observer' placed rotating at mean entrainment velocity $(\omega_j + \omega_b)/2$

Note v_ξ can be related directly to Mobility vector M_ξ (parallel to the load line) by:-

$$v_\xi = F\ (c/r)^2\ c\, M_\xi / (\eta\, b\, d)$$

ALTERNATIVE GROUPS AND TERMINOLOGY - 2
(Vector and scalar form Martin, Booker, Lo 1987; Vector form Booker 1989)

POWER LOSS	=	"SHEAR" (OR "COUETTE") CONSTITUENT	+	"SQUEEZE" (OR "POISEUILLE") CONSTITUENT
H	=	$\underline{T}_{shear} \cdot \Delta\omega$	+	$\underline{F}_{pressure} \cdot \underline{V}_{squeeze}$ (vector form)
	=	$\eta r^3 J_1^{00} (\omega_j - \omega_b)^2$	+	$F[\{(\omega_j + \omega_b)/2\} \sin\varphi + v\cos\beta]$ (scalar form)
		From differences in journal and bearing rotating relative to any 'observer'.		From journal translation relative to an 'observer' rotating at the mean 'entrainment' velocity of journal and bearing.

5.1.2. Other forms of power loss equations. Several terms make up the general power loss equation above and these can be made more suitable to particular situations by choosing the appropriate frame of reference. The friction losses due to shear will be the same irrespective of where an observer is placed, since the shear term involves the *difference* of the two angular velocities. The total power loss is not affected by the choice of the reference frame and therefore the *sum* of the remaining terms, labelled 'pressure' and 'translation' will be unchanged, although they will vary individually depending on the motion of the 'observer' frame. Such changes to the 'pressure' and 'translation' terms can be illustrated by an example case where the 'observer' is rotating at the mean 'entrainment' velocity $(\omega_j + \omega_b)/2$. These more recent equations are summarized in Table 3 together with the earlier ones detailed above. The various groups and terminology are given for completeness. The first alternative group shows a fairly concise power loss equation for use with Mobility data. The last group shown (in vector form) is derived from a unique and novel set of basic equations for fluid films (for further details see Booker 1989). As mentioned previously all the equations in Table 3 should produce the same value of power loss (as seen from the first and last set of equations).

5.2. Power Loss in Big End Connecting Rod Bearings

The use of the above equations is illustrated considering the example big end bearing case .

5.2.1. Power loss - big end bearing example The power loss in the big end bearing example case for the automotive 1300cc engine has been calculated using the alternative terminology ('shear' and 'squeeze') shown in Table 3. The engine and bearing data, load diagram and journal centre orbit for this example are given in Table 1 and the middle of Fig. 3.

The predicted power loss distribution throughout 0° to 720° of crank angle, using finite bearing theory, is shown in Fig. 21 together with cycle averaged values. The upper line shows the total power loss assuming a 2π (void free) film extent. The slightly lower line through the hatching shows the total power loss assuming a completely void-filled cavitated zone in the non-load carrying part of the bearing. In reality, the total power loss with a 'striated' film will probably lie between these lines. To err on the safe side, higher friction, a completely void free film is sometimes considered (as in the parameter study to follow). The major component of the cycle averaged power loss, for the example case, is due to the 'shear' term, 70 W when using the short bearing solution and 84 W when using the finite bearing solution; the remaining component termed 'squeeze' (including the effect of journal centre translation velocity) accounts for about 10 W for both solutions.

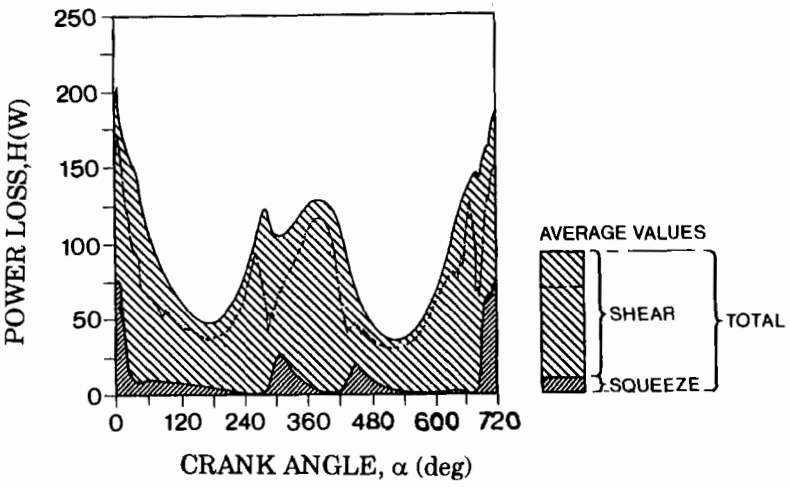


Fig. 21 Power loss - big end bearing example (1300cc engine).

5.2.2. General design chart for power loss in big end bearings Cycle averaged power loss results for over 500 other cases are summarized in dimensionless form in Fig.22. This design aid covers a wide range of design parameters and considers the various shaped polar load diagrams given in Fig. 2 and involves both bearing and engine system parameters.

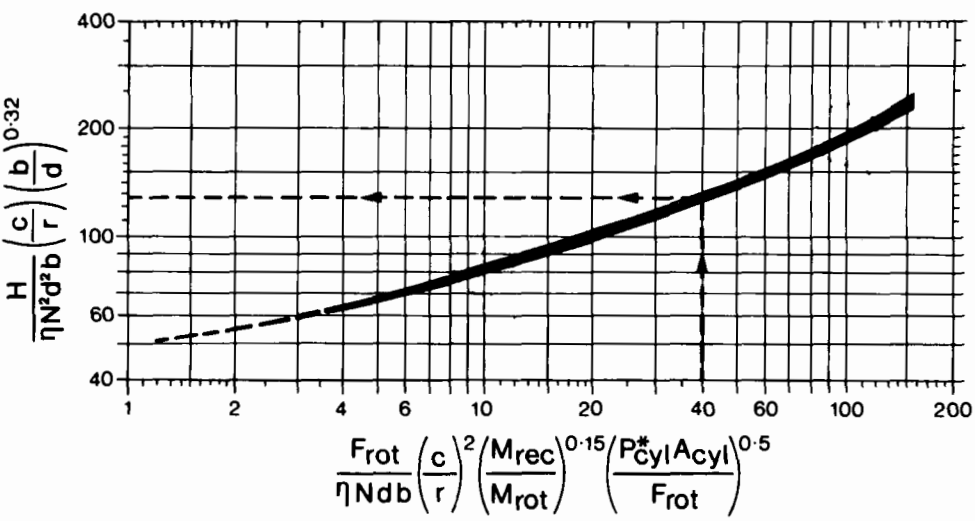


Fig. 22 Total power loss - finite difference solution

Alternatively the power loss can be obtained (numerically) from a curve fit of these results:-

$$H_{total} = [(\ 11\ X^{0.57} + 40\ )\ (\eta\ N^2\ d^2\ b)\ ]/\ [(c/r)\ (b/d)^{0.32}]\ \text{watts} \qquad (14)$$

where the variables are in consistent SI units and where X equals all the terms (with appropriate indices) on the horizontal axis, Fig.22, involving the dimensionless load number, reciprocating mass ratio and the cylinder pressure term. It is assumed in equation 14 that a complete full film of oil is present. (To allow for the presence of striated films one could take 85% of this value).

To complement these results general guidance sensitivity charts, Fig.23, have been developed separating out the engine and bearing parameters so that the effect on power loss of individual variables can be readily assessed. For example if the diameter is increased by 50% (i.e. increased by a ratio of 1.5:1) then the power loss will be increased by 2.4:1. Similarly it can be seen that changing the reciprocating mass has little effect on power loss in the big end bearing (although it will affect friction losses in the piston assembly). If two variables are changed simultaneously then the multiplying factor on power loss will be the product of the two resulting individual factors.

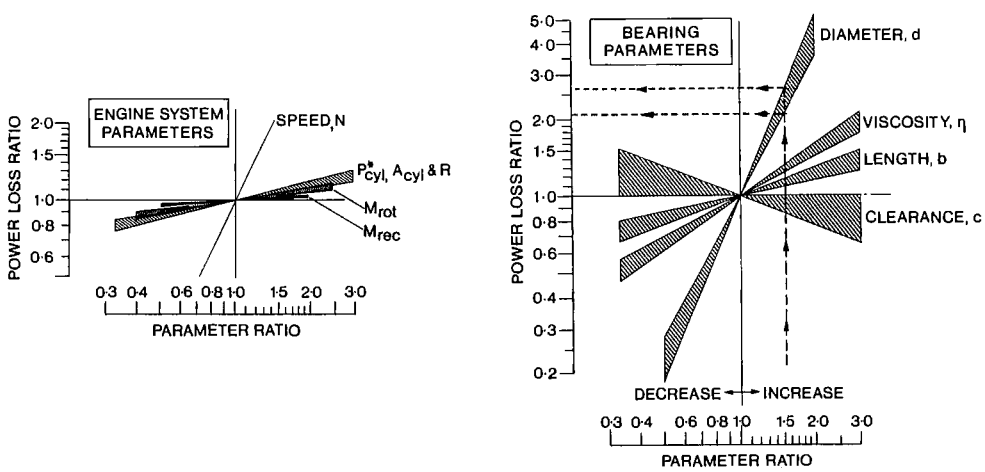


Fig. 23 Power loss sensitivity - parameter study

The power loss graph, Fig.22 and sensitivity charts, Fig.23, are useful to show trends when changes are made.

5.3. Improved Power Loss Predictions

With the general trend of increasing power to weight ratios in automotive engines, bearing sizes are becoming smaller and are required to operate with thinner oil film thicknesses. This increases the chance of having asperity contact between shaft and bearing at some brief moment during the load cycle. Such intermittent contact will result in higher local friction in the asperity contact zone. This consequently produces a larger friction torque (averaged over the load cycle) than would be predicted by the hydrodynamic theory discussed previously. A method for calculating the effects of this intermittent contact (called 'Direct Contact Reaction') is given in a paper by Conway-Jones and Gojon 1991. In practice additional heat generated due to direct contact frictional heating, when present, is added to that generated by the hydrodynamic action (mainly due to shearing of the oil film). This is then considered in the reiterative heat balance to determine the operating temperature of the bearing by considering the heat generated in the oil film and the heat carried away by both the oil flow and by the conduction through the crankshaft. As the oil flow from automotive big end bearings is small compared with that from main bearings up to 50% of the heat generated in an automotive connecting rod bearing may be conducted down the crankshaft. This has been recognized by some engine builders and has been substantiated in the detailed analysis of 'Heat flow in crankshaft bearings' by Conway-Jones and Gojon, 1991. In the same paper it is also shown that due to the conduction through thin oil films the oil cannot get more than a few degrees hotter than the shaft which acts as a recuperator.

6. ELASTIC DISTORTION OF BEARING HOUSING AND SHAFT FLEXIBILITY

There are two areas where improved modelling, considering elastic distortion, can give a more realistic prediction of bearing performance. One relates to improved crankshaft modelling allowing for the better prediction of main bearing load sharing. The other relates to gross deformation of the bearing housing and has particular significance in connecting rod bearing studies, where the big end bearing may stretch considerably under load.

6.1. Main Bearing Load Sharing

The loads on the big end bearing are relatively simple to calculate . The main bearing loads must react against the big end loads and traditionally a statically determinate system is generally considered (rapid method) in which the crankshaft is treated as if it were pin jointed at the axial mid-position of each main bearing. In practice however both crankshaft and crankcase have finite stiffness, so that very complex interactions can be set up throughout the entire engine.

Researchers have now attempted to take into account engine flexibility, and to couple this with the bearing analysis. Work at Cornell University (USA) and independently at Perkins Engines Ltd.(UK) was reviewed Martin 1983. Welsh and Booker 1983 improved the modelling of the shaft to include 'super element' substructures. Their procedure has been adapted by T&N Technology with the assistance of Professor Booker for use in the Glacier program LOADSHARE. The effect of crankshaft and crankcase stiffness on bearing load and minimum oil film thickness is illustrated in an example in reference Conway-Jones, Martin, Gojon 1991.

6.2. Elastic Distortion of Bearing Housing.

The Elastic distortion of bearing housings particularly at the connecting rod big end is an important factor in design. Experimental evidence of distortion, on the big end bearing of the Ruston VEB Mk III 600 h.p. diesel engine, was reported by Campbell, Love, Martin & Rafique 1967. A way of presenting these results plotting the journal position and the limiting clearance shapes* (for various crank angles) is shown in Fig. 24. The predicted journal positions for the rigid bearing are shown at the centre of Fig. 24 with the limiting clearance shapes spaced around the outside. When the load is on the cap half of the bearing (downwards), stretching of the bearing is clearly evident (at crank angle 360°), where the vertical clearance is about twice the horizontal clearance. Also the relatively small local clearance near the journal centre indicates the cap half of the bearing wrapping around the crank pin.

The procedure for computing the elastohydrodynamic solution is very complex and apparently it was not until the early 1980's (see special references - Engine bearings with elastic deformation) that researchers predicted film thickness for such bearings. The timing of this probably coincided with the availability of more computing power and an awareness of the need for such procedures. The predicted limiting clearance shapes in the rod and cap halves of an automotive engine, shown in the discussion on the Oh and Goenka paper 1985, generally show the same trends as the experimental shapes in Fig. 24.

Aitken and McCallion, 1991, include in their paper a concise and interesting review of elastohydrodynamic models for dynamically loaded big end bearings.

*Footnote

Physically these shapes can be interpreted by considering the bearing shape 'frozen' at that instant (crank angle ref.) and the shape can then be described by the trace of the journal centre if the journal were rolled around the bearing surface. This gives a limiting clearance shape within which the journal centre remains (analogous to the usual clearance circle for a rigid circular bearing). This type of presentation complements rather than replaces the usual linear plot of film thickness against angle around the bearing.

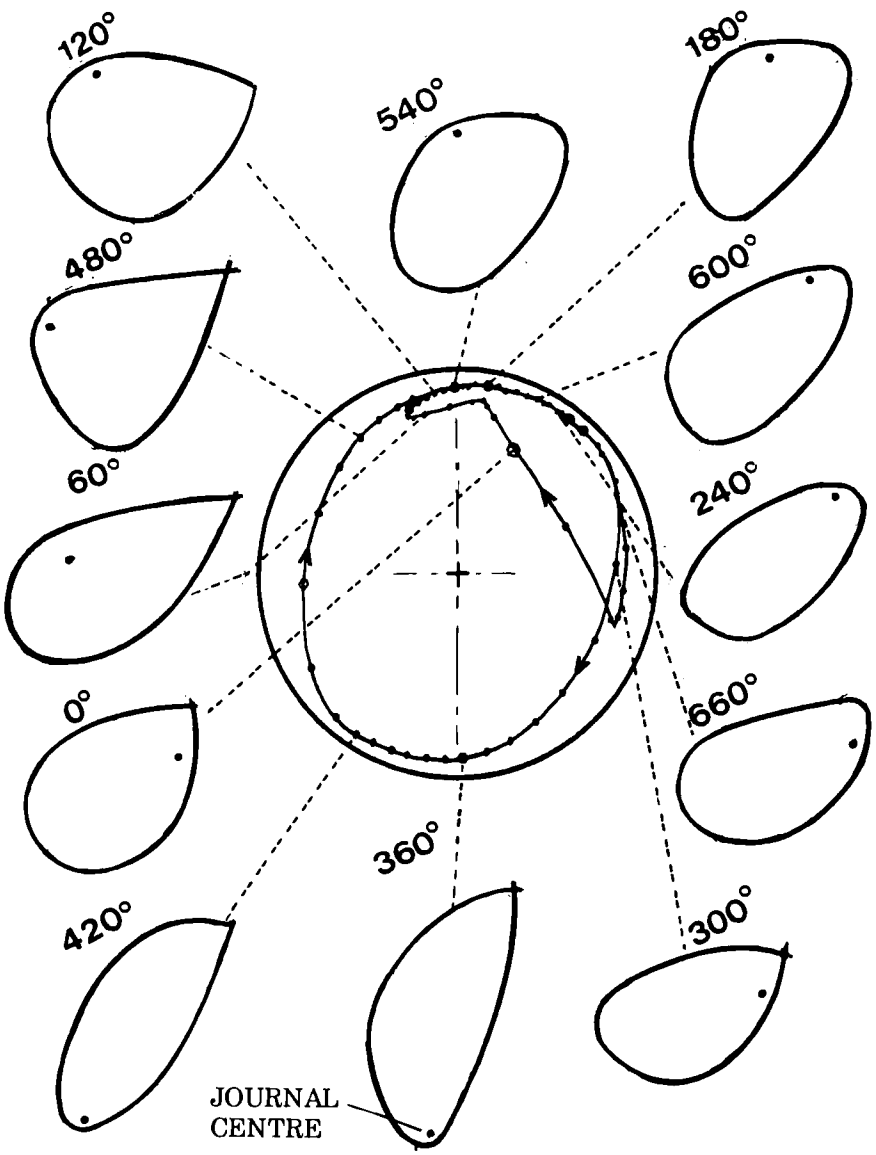


Fig. 24 Continually changing limiting clearance shapes and journal position from actual VEB Mk III engine tests (compared with rigid bearing)

7. EXPERIMENTAL SUPPORT

Experimental support is a necessary part of the development and verification of bearing design techniques. Experimental results are helpful:

a) to verify existing procedures (giving confidence to designers).
b) to modify design procedures to bring them nearer to practice.
c) to show the significance of various parameters
and d) to point out gaps in knowledge and a need for further research (either experimental or theoretical).

A few such work examples pertinent to the notes on 'Oil flow' (section 4) and 'Bearing friction and power loss' (section 5) illustrate the usefulness of such experiments in relation to design techniques.

7.1. Experiments Relating to Oil Flow.

7.1.1. Intermain bearing flow on a 1.8 litre engine. Oil flow results from an intermain bearing of a 1.8 litre 4 cylinder gasoline engine were measured and compared with various predictive methods, Jones, Lee, Martin. 1982. The intermain bearing was chosen as there was no onward feed to the big end and flows could be measured directly. Details and operating conditions are given in the paper and a summary of the experimental flows for different feed arrangements is shown in Fig.25, together with results from various flow predictive methods.

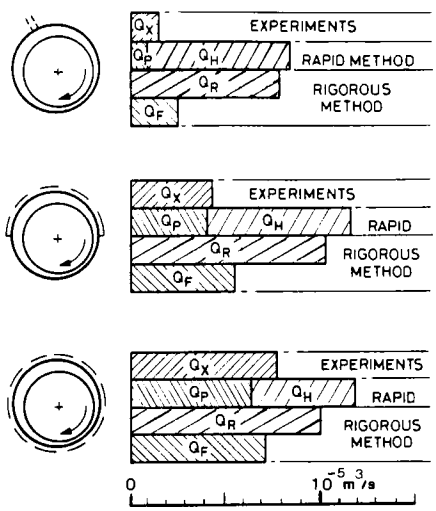


Fig. 25 1.8 litre engine - intermain bearing: Oil flows from various predictive techniques compared with engine test results

One can see that for this particular example case, the 180° partial groove and 360° groove have experimental flows (Q_X) about three and five times that of the single hole respectively. Furthermore, it can be seen that the more realistic film history flow (Q_F) is very close to the experimental values. For the partial and circumferentially grooved bearing the feed pressure flow (Q_P) also gives a good indication of actual flow, but not so for the single hole case. As the film history flow is verified experimentally, the general film history mathematical model is now viable as an 'experimental' tool for further research and parameter studies. The theoretical results shown in Fig.25 have also been confirmed by others. (see Authors closure to ref. Paranjpe and Goenka 1990)

7.2. Experiments Relating to Friction.

7.2.1. Friction torque for main bearings from motored engines.

Sometimes experimental data from various sources becomes more meaningful if grouped together. A survey of how friction torque in main bearings depends on the bearing variables is shown in Fig. 26. (references quoted are from Martin '85).

Key No.	Engine size	Literature Reference
1	Ford large range	Bishop'64
2	Vauxhall 1250cc	Vickery'75
3	Mazda 1500cc	Miyachika'80
4	Ford 4-cyl.	Kovach et al'82
5	Magnum 2000cc	Nagao et al'83
6	Subaru 1300cc	Hoshi'84

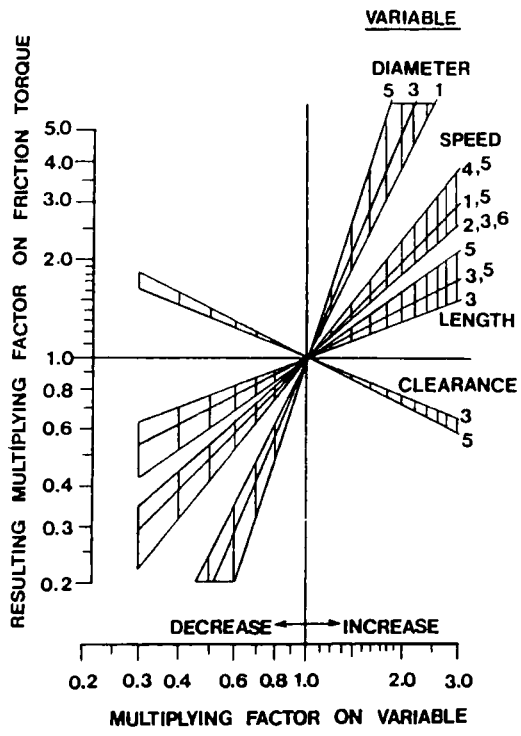


Fig. 26 Friction torque sensitivity chart for main bearings (data from motored engines)

It is reassuring that both Fig. 23 from theory and Fig.26 from practice have similar characteristic trends, albeit one relates to big end bearings and the other to main bearings. (The power loss and torque trends considering the variable speed are naturally different).

7.2.3. Torque measured under thin film conditions in a dynamically loaded test bearing. Fig. 27 shows torque measurements carried out on a friction test rig at T & N Technology, Conway-Jones, Martin, Gojon 1991. The bearing was subjected to dynamic loading. Thin oil film conditions were achieved by increasing the oil temperature to the bearing. Note the significant difference between predicted torque and measured torque which levels off at high temperatures (thin film conditions). These results are supportive evidence to justify the consideration of intermittent contact when the highest asperities of each surface make contact through the oil film. (discussed in section 5.3.)

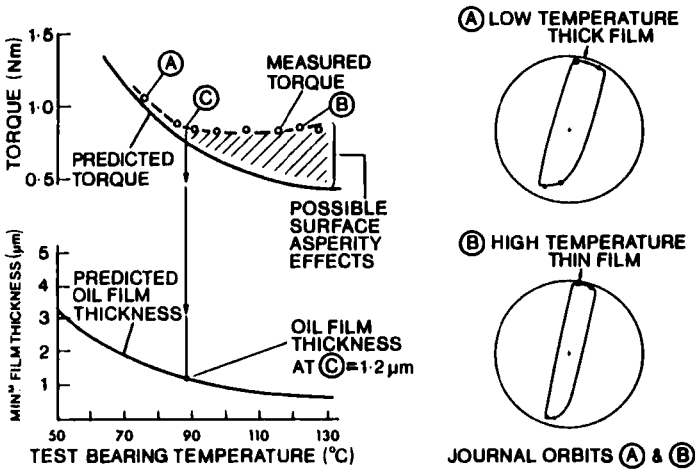


Fig. 27 Comparison between measured and predicted friction torque in dynamically loaded test bearing

From the above examples it can be seen that experimental results are invaluable in pointing out more realistic predictive methods as well as validating existing design procedures.

8. CLOSING COMMENTS

Some of the present advances in bearing design procedures occurred in the 1980's when presumably more computing power became available and there was an awareness of a need for more realistic solutions. The notable advances include oil film history programs (more realistic oil flow predictions): main bearing load sharing (more realistic bearing loads) and the consideration of elastic deformation particularly in big end bearings. Oil film history and load sharing programs are currently used by bearing specialists (application and research engineers) only on a 'non-routine' basis as they take several hours computer time (as against a few seconds for the rapid interactive day-to-day programs). The more advanced programs are often used as diagnostic tools or for examining new designs outside the range of current experience.

As rigorous programs become more complex and time consuming, certain aspects of these are now considered in processed form. Resulting design aids, curve fitted equations or simply 'rules' result from detailed studies of many (sometimes hundreds of) bearing cases using rigorous solutions. These purposely processed results are used as design aids to obtain 'rigorous' results from rapid methods (graphical design aids, data banks, curve-fit equations etc). Examples of such processing to give improved rapid design methods are:-

a) The finite bearing solution curve-fitted equations in Mobility format, (equations, Goenka 1984).

b) Allowing for groove configurations (rules, section 3.3) when using the finite bearing curve fit solution (also applies to short bearing Mobility method).

c) Minimum film thickness in connecting rod bearings based on finite bearing theory (design chart, section 3.2).

d) Oil flow in big end bearings from oil film history program (equations, section 4.4). This uses the rapid predictions of pressure flow Q_P and hydrodynamic flow Q_H merely as a framework to curve-fit the rigorous film history flow Q_F.

e) Power loss in connecting rod bearings from finite bearing solution (design chart and equations, section 5.2.2)

Other current improvements in bearing design techniques relate to the heat flow in crankshaft bearings (Conway-Jones & Gojon 1991) involving oil flow and power loss (with the introduction of 'direct contact reaction'). Thus we have a relatively new design tool for predicting more realistic operating viscosities and bearing temperatures, allowing for heat conduction from big-end to main bearing and taking account of surface asperity contact in the bearings.

This is an important step forward since all the bearing characteristics are dependent on effective operating viscosity in the bearing. Experimental results from real engines (different types and sizes) could be helpful to further verify and develop these improved techniques.

Other effects which are important in estimating the effective operating viscosity and should be included in the heat balance (at some stage of development in the future) are the elasticity effects of multigrade oils which contain polymers to improve the viscosity index. For such oils the viscosity is not only dependent on temperature but there is a thinning effect at high rates of shear. Studies are continuing on this and also on the thickening effect of viscosity at high film pressures (see for example Bates, Williamson, Spearot & Murphy 1988).

In the relatively new field of predicting elastohydrodynamic effects in engine bearings (particularly big end bearings), the limiting clearance shapes and the local wrap round effects are clearly important from the bearing performance point of view. The effect, on bearing performance, of local design changes to the con-rod section (e.g. cutting pieces away in the rod half) are now capable of being studied by some companies (Goenka and Oh 1986b) using rigorous solutions. Aitkin and McCallion 1992 use an inertial ring model to characterize elastic big end bearing film thickness performance. Such aids (simulating the eye of the con-rod in a relatively simple way) could be helpful in giving a general feel for the inter-relationship between structure and bearing parameters when considering minimum film thickness.

9. ACKNOWLEDGEMENTS

Thanks are extended to T & N plc for permission to reproduce Fig. 27 (from Conway-Jones, Martin, Gojon 1991) and to T & N Technology Ltd for permission to use their finite bearing data to enable the P*cyl. term to be included in the film thickness chart (Fig.10). Figs. 1, 2, 3, 20, 21 and 23 (from Martin, Booker, Lo 1987) and Figs. 6, 19 and 26 (from Martin 1985) are reprinted by permission of the Council of the Institution of Mechanical Engineers.

10. REFERENCES

10.1. General

Bates T.W., Williamson B., Spearot J.A. and Murphy C.K. 1988. 'The Importance of Oil Elasticity'. Industrial Lubrication and Tribology, Nov/Dec. pp 4-19.

Booker J.F. 1965a. 'Dynamically Loaded Journal Bearings - Mobility Method of Solution'. J. Basic Eng., Trans ASME, Sept. Series D. 187, pp 537-546.

Booker J.F. 1965b. 'A Table of the Journal Bearing Integral'. Trans. ASME Journal of Basic Engineering. June. Vol.87 pp 533-535.

Booker J.F. 1971. 'Dynamically-Loaded Journal Bearings: Numerical Application of the Mobility Method'. Trans. ASME, Journal of Lubrication Technology, Vol.93, Series F, No. 1, January.pp 168-176. Errata: Vol. 93, Series F, No.2, April. p.315.

Booker J.F. 1979. 'Design of Dynamically Loaded Journal Bearings', Fundamentals of the Design of Fluid Film Bearings, ASME publication.

Booker J.F., Goenka P.K. and van Leeuwen H.J. 1982. 'Dynamic Analysis of Rocking Journal Bearings with Multiple Offset Segments'. Trans. ASME Journal of Lubrication Technology. Oct. Vol. 104 pp 478-490; addendum: 1983 April. Vol 105 p.220.

Booker J.F. 1989. 'Basic Equations for Fluid Films with Variable Properties'. Trans. ASME Jour. of Trib. 111. pp 475-483.

Campbell B.D., Gojon R., Serre Y. and Warriner J.F. 1986. 'Developments in Engine Bearing Design Techniques' AE Tech. Symposium (April) AE plc. Cawston House, Rugby, Warwickshire. UK.

Clayton G.A. and Taylor C.M. 1968. 'Thermal Considerations in Engine Bearings'. Proc. of the 15th Leeds-Lyon Symposium on Tribology, The Tribological Design of Machine Elements, Elsevier. pp 333-342.

Conway Jones J.M. and Gojon R. 1991. 'Heat Flow in Crankshaft Bearings'. Proc. of the 17th Leeds-Lyon Symposium on Tribology. Elsevier Science Publishers.

Conway-Jones J.M., Martin F.A. and Gojon R. 1991. 'Refinement of Engine Bearing Design Techniques'. Tribology International.Vol.24 no.2, April, pp 119-127.

Cooke W.L. 1983. 'Performance of Dynamically Loaded Journal Bearings. Part 2: Measurement of Oil Film Pressures.' NEL Report No.688. National Enginering Laboratory, Glasgow.

Garner D.R., James R.D. and Warriner J.F. 1980 'Cavitation Erosion Damage in Engine Bearing: Theory and Practice'. ASME Trans. J.Eng.Power, 102, pp 847-857.

Goenka P.K., 1984, 'Analytical Curve-fits for Solution Parameters of Journal Bearings', Journal of Tribology, Trans. ASME, Series F, Oct.'84, pp 421-428.

Hirano F. and Shodai N. 1958. 'Oil Flow Coefficient of Pressure Fed Journal Bearing'. Bull. Japan Soc.Mech.Eng. 1,2 pp 184-188.

Hisatomi T. and Hiroshi L. 1982. 'Nissan Motor Company's New 2 Litre Four Cylinder Gasoline Engine'. SAE Transactions. SAE paper 820113.

Jones G.J., Lee C.S. and Martin F.A. 1982 'Crankshaft Bearings: Advances in Predictive Techniques Incorporating the Effects of Oil Holes and Grooving'. AE Tech Symposium (April) AE plc. Cawston House, Rugby, Warwickshire, UK.

Jones G.J. 1983, 'Crankshaft Bearings: Oil Film History', Proc. of the 9th Leeds-Lyon Symposium on Tribology - Tribology in Reciprocating Engines, Butterworths, pp 83-89.

Kovach J.T., Tsakris E. A. and Wong L.T. 1982. 'Engine Friction Reduction for Improved Fuel Economy'. Society of Automotive Engineers SAE paper 820085.

Lang O.R. 1982. 'Reibungsverluste in Verbrennungsmotoren' (Friction losses in combustion engines). Schmiertechnik + Tribologie 29. Presented at conference 'Limits of Lubrication' London July 1981.

Lloyd T., Horsnell R. and McCallion H. 1966. 'An Investigation into the Performance of Dynamically Loaded Bearings: Design Study'. Symp. on Journal Bearings for Reciprocating and Turbo Machinery, Proc.I.Mech.E. 181 part 3B.

Martin F.A. and Booker J.F. 1967. 'Influence of Engine Inertia Forces on Minimum Film Thickness in Con-Rod Big-End Bearings'. Proceedings I.Mech E Vol.181, Part 1, pp 749-764.

Martin F.A. 1983, 'Developments in Engine Bearings', Proc. of the 9th Leeds - Lyon Symposium on Tribology, Tribology in Reciprocating Engines, Butterworth. Also Tribology International, June'83, Vol. 16, No.3, pp 147-164.

Martin F.A.,Lee, C.S., 1983, 'Feed-Pressure Flow in Plain Journal Bearings', Trans. ASLE, Vol.26, July'83, pp 381-392.

Martin F.A. 1985 'Friction in Internal Combustion Engine Bearings'. I.Mech E. Conf. Proc., Combustion Engines - Reduction of Friction and Wear. London, March I Mech E. MEP publication. pp 1-18.

Martin F.A, Booker J.F. and Lo P.M. 1987 'Power Loss in Connecting Rod Bearings'. I.Mech.E Conf.on Tribology Friction and Wear, London. Paper C167/87. pp 701-708.

Martin F.A. and Stanojevic M. 1991. 'Oil Flow in Connecting Rod Bearings'. Proceedings of the 17th Leeds-Lyon Symposium on Tribology, Elsevier Science Publishers.

Paranjpe R.S., Goenka P.K. 1990, 'Analysis of Crankshaft Bearings using a Mass Conserving Algorithm', STLE Tribology Transactions, Vol.33, 3, pp 333-344.

Spikes R.H. and Robinson S.M. 1982. 'Engine Bearing Design Up-to-Date'. IMechE. Conf. 'Tribology - Key to the efficient engine'. MEP publication pp 1-8.

Welsh W.A. and Booker J.F. 1983. 'Dynamic Analysis of Engine Bearing Systems'. SAE paper 830065 also Proc. of the 9th Leeds-Lyon Symposium on Tribology, 'Tribology of Reciprocating Engines' Eds. Dowson D. et al. Butterworth's London. pp 29-36.

10.2. Engine Bearings with Elastic Deformation.

Aitken M.B. and McCallion H. 1991.'Elastohydrodynamic Lubrication of Big-end Bearings. Part 1: Theory and Part 2: Ratification.' Proc. I.Mech.E. Part C: J.Mechanical Eng.Science. Vol.205. pp 99-119.

Aitken M.B. and McCallion H. 1992. 'Parametric Minimum Film Thickness Performance of an Elastic Big-end Bearing under Inertial Load'. Proc. I.Mech.E Part C: J Mechanical Eng.Science. Vol.206. pp 3-12

Campbell J., Love P., Martin F. and Rafique S. 1967-8. 'Bearings for Reciprocating Machinery: A Review of the Present State of Theoretical, Experimental and Service Knowledge.' Proc. I.Mech.E. 182(3A), 51-74

Fantino B., Godet M. and Frene J. 1983. 'Dynamic Behaviour of an Elastic Con-rod Bearing - Theoretical Study'.SAE/SP.539. SAE paper 830307. pp 23-32.

Fantino B. and Frene J. 1985. 'Comparison of Dynamic Behaviour of Elastic Con-rod Bearings in Both Petrol and Diesel Engines.' Trans. ASME Journal of Tribology 107(1), pp 87.

Fenner D.N., McIvor J.D.C., Conway-Jones J.M. and Xu H 1993. 'The Effect of Compliance on Peak Oil Film Pressure in Connecting Rod Bearings' Proc. of 19th Leeds-Lyon Symposium of Tribology. Elsevier Science Publishers.

Goenka P.K. and Oh K.P. 1986a. 'An Optimal Short Bearing Theory for Elastohydrodynamic Solutions of Journal Bearings.' Trans. ASME. Journal of Tribology, 108(2), pp 294-299.

Goenka P.K and Oh K.P. 1986b. 'An Optimal Connecting Rod Design Study - A Lubrication Viewpoint'. Trans. ASME. Journal of Tribology, July, 108(3), pp 487-496.

Goodwin G. and Holmes R. 1978. 'On the Continuous Monitoring of Oil-film Thickness in an Engine Bearing.' Proc. I.Mech.E. Vol.192. pp 371.

Kumar A., Goenka P.K. and Booker J.F. 1990. 'Modal Analysis of Elastohydrodynamic Lubrication: A Connecting Rod Application.' Trans. ASME. Journal of Tribology, 112(3), July, pp 524-534.

La Bouff G. and Booker J. 1985. 'Dynamically Loaded Journal Bearings: Finite Element Treatment for Rigid and Elastic Surfaces.' Trans. ASME. Journal of Tribology 107(4), pp 505-515.

Martin F. 1983. 'Developments in Engine Bearing Design.'
Tribology International., 16(3), July, pp 147-164.

McIvor J.D. and Fenner D.N. 1989. 'Finite Element Analysis of Dynamically Loaded Flexible Journal Bearings: A Fast Newton-Raphson Method.' Trans. ASME. Journal of Tribology, 111(4), Oct. pp 597-604.

Moes H., Ten Hoeve P.B.Y and Van derHelm J. 1989. 'Thermal Effects in Dynamically Loaded Flexible Journal Bearings.' Trans.ASME. Journal of Tribology 111(1), pp 49.

Oh K.P. and Goenka P.K. 1985. 'The Elastohydrodynamic Solution of Journal Bearings Under Dynamic Loading'. Trans. ASME. Journal of Tribology, July, 107(3), pp 389-395.

Smith E.H. 1985, 'Elastohydrodynamic Lubrication in Dynamically Loaded Journal Bearings'. Proc. of the 12th Leeds-Lyon Symposium on Tribology. (Eds. D.Dowson, C.M.Taylor, M.Godet and D.Berthe), pp 375-380.

van der Tempel L., Moes H. and Bosma R. 1985. 'Numerical Simulation of Dynamically Loaded Flexible Short Journal Bearings.' Trans. ASME. Journal of Tribology, 107(3), pp 396.

van der Tempel L., Moes H. and Bosma R. 1985. 'Starvation in Dynamically Loaded Flexible Short Journal Bearings.' Trans. ASME. Journal of Tribology. Vol.107. pp 516-521.

Xu H. and Smith E.A. 1990. 'A New Approach to the Solution of Elastohydrodynamic Lubrication of Crank-shaft Bearings.' Proc. I.Mech.E. Journal of Mechanical Enginering Science. Vol.204(C3) pp 187-196.

This Page Intentionally Left Blank

Engine Tribology / C.M. Taylor (Editor)

VALVE TRAIN - CAM AND FOLLOWER : BACKGROUND AND LUBRICATION ANALYSIS

C M TAYLOR

Department of Mechanical Engineering, University of Leeds, Leeds S2 9JT

1 INTRODUCTION

In the 1970's the western world became increasingly aware of the fact that the earth's reserves of fossil fuels were not limitless. Political pressures from environmental groups and, much more significantly, the pressures brought about by the Middle Eastern countries upon the economies of their consumers during the 'Oil Crisis', caused many governments to campaign for the merits of saving energy. The motor manufacturers were put under increasing pressure from the consumer to design vehicles that not only cost less than their competitors for a similar specification, but also returned better fuel consumption figures.

Coincidentally, around this time, the use of an overhead camshaft (OHC) using pivoting followers to drive two banks of valves had just become very popular within the motor industry. The design was favoured as it allowed the cylinder head to be assembled separately from the block (good from a production point of view), it operated better at high speeds than other single camshaft designs driving two banks of valves (as it was more rigid) and it utilised fewer components and was therefore cheaper. Unfortunately, the design was found to be inherently poor from a tribological view point - many manufacturers suffering from early failures of their (OHC) valve train systems due to excessive wear of the cams and followers.

The problem was therefore defined; the designer was required to design valve trains that had very small frictional losses and that would not wear out within the life of the rest of the vehicle. Regrettably the tools that the designer needed to fulfil the task were not available and similar types of design kept emerging.

Much work has been carried out upon ways in which the efficiency of the internal combustion engine could be increased. Many workers investigated the sources of losses within the internal combustion engine and ways in which they could be reduced (for example, Parker and Adams (1982), Hoshi (1984) and Martin (1985)). Although the majority of the losses were found to be due to thermal inefficiencies, there are still large benefits to be gained from reducing the

mechanical losses which account for approximately 15% of the total fuel energy input. Hoshi suggested that the valve train frictional losses account for between 7.5% - 21% of the total engine frictional loss, and so there is still a large scope for improvement.

Whilst there have been major developments in the tribological understanding of the behaviour of many of the engine components such as the dynamically loaded bearings and the piston assemblies, the amount of work upon the valve train has been relatively small. This may be due to the complexity of the problem as the cam and follower operate in the most arduous tribological conditions within the internal combustion engine. There is still a need for theoretical studies of the cam and follower contact to be undertaken, for the tribology of such contacts is still far from being fully understood. There is also an increasing need for existing knowledge to be put into a form that can be used by designers within industry.

1.1 Notation

a	follower acceleration
b	half width of Hertzian line contact
de	small change in e (Figure 4)
e	eccentricity of contact point from line of action
E	power loss associated with the contact
E'	equivalent elastic modulus
F	friction force
G	dimensionless materials parameter
h	film thickness
H	dimensionless film thickness
L	cam lift
p	pressure
p_{max}	maximum Hertzian pressure
P'	load per unit width
R_b	base circle radius
R_c	radius of curvature of contact point on cam ($=R_e$)
R_e	equivalent radius of curvature of contact point
s	curve tangential co-ordinate direction
t	time (also by 'dot' notion, ie e = de/dt)
u,v,w	absolute velocities in cartesian co-ordinate directions
u_e	mean entraining velocity
U_e	dimensionless speed parameter
W_e	dimensionless load parameter
x,y,z	cartesian co-ordinate directions
α	pressure-viscosity coefficient
$\dot{\gamma}$	angular velocity of tangent to a curve
η	dynamic viscosity of lubricant

η_o	reference dynamic viscosity at ambient pressure and chosen temperature of lubricant
θ	cam angle
λ	film thickness parameter
μ	coefficient of friction
ρ	lubricant density
$d\phi$	small angular variation (Figure 4)
Ω	cam rotational frequency

Subscripts

1,2	solids in contact (Figure 3)
C_1	point on camshaft (Figure 3)
cen	central
max	maximum
min	minimum

2. BACKGROUND

Valve train design philosophy has been changing rapidly over the past decade. This has been due mainly to research carried out after many manufacturers experienced serious valve train wear problems with new engines in service. The cost of research therefore paled into insignificance compared with the cost of warranty claims and the loss of sales.

Although throughout this century various different types of engine valving have been tried, the poppet valve has been almost universally adopted by the major automobile manufacturers. Other types of valving such as sleeve or rotary valves have been deemed to have lubrication difficulties, allow excessive engine oil consumption, provide poor sealing and have excessive frictional losses (Buuck (1982)).

Over the years there has been a demand for every increasing engine speeds in the search for more energy efficient engines. This has caused the rise in popularity of overhead camshaft mechanisms (OHC) at the expense of push-rod systems. The push-rod system had been favoured in the past due to its many virtues; ease of adjustment, the availability of the camshaft to drive accessories such as the oil pump and distributor and good lubrication and wear characteristics. As the camshaft is located close to the sump it receives a plentiful supply of oil in the form of oil mist and splash from the crankshaft. Also the tappets are free to rotate, thus improving lubricant entrainment and decreasing wear by ensuring that any two points on the cam and tappet surfaces do not continually meet cycle after cycle.

The main disadvantage of push-rod systems is their flexibility brought about by the use of long thin push-rods. This makes them unsuitable for use in very high speed engines. OHC mechanisms are inherently much stiffer. Modern production techniques have also added to the decline in popularity of the push-rod system. OHC mechanisms utilise fewer parts and allow the cylinder head assembly to be built up as a separate unit. Production engineers see these as great advantages (Polak and Letts (1987)). OHC mechanisms appear in two basic forms; either direct acting or via a pivoted follower. One of the major manufacturing and assembly problems with these systems is in the alignment of the camshaft and followers. If the cam lobes and followers are not properly aligned then severe edge loading can occur resulting in damage to the contacting parts. Mercedes patented a novel solution to this problem in 1959. Their design utilised end pivoted followers which located upon spherical ended posts rather than the more conventional rocker shaft. This allowed the followers to self align.

Whilst direct acting OHC mechanisms have proved to be very successful both from a performance and wear point of view, the same cannot be said of pivoted OHC mechanisms. Many manufacturers have experienced major wear problems with these types of valve trains. This has been attributed to many causes. One suggested cause is the higher temperature seen in the cylinder heads of modern engines due to the adoption of thermostatically controlled electric fans and the use of more selective coolant channels through the whole of the engine. These higher temperatures not only increase the bulk temperature of the contacting parts, thus increasing the probability of scuffing (Dyson and Naylor (1960)), but also serve to lower the viscosity of the lubricant. As the followers do not run in bores they have limited means of conducting away heat generated in the contact region - this again leads to high bulk temperatures in the followers. As the camshaft is at the very top of the engine in an OHC arrangement it is also at the end of the lubricant feed path. In many early OHC designs the cam and follower contact had to rely upon oil splashed from the camshaft bearings to provide adequate amounts of lubricant. Other reasons suggested for the untimely demise of such systems have been fuel dilution of the lubricant and oil starvation at engine start-up.

Most manufacturers have solved the problem of excessive wear in pivoted OHC systems by the adoption of high specification materials and the use of spray bars and even holes in the cam lobes to supply sufficient lubricant to the contact region. The additional supply not only serves to lubricate the contact but also acts as a coolant. One manufacturer has reverted to the use of manual chokes which perhaps suggests that they blamed fuel dilution for their problems. Possibly if attention were to be directed towards the geometry and kinematics of such systems significant gains in lubricant entrainment could be achieved. This point will be illustrated by the use of parametric studies.

OHC mechanisms, however, do have several drawbacks apart from their frequently poor wear characteristics. The camshaft is often used to drive auxiliaries such as the fuel and oil pump and the distributor. This means that these

components must be mounted very high on the engine if they are to be driven from the camshaft. It is therefore often necessary when using an OHC valve train to drive the oil pump from the crankshaft so as to keep the distance from the pump to the sump as small as possible. This unfortunately aggravates overall engine length and may also lead to lubricant aeration problems. The majority of the flexing of the valve train system takes place in the camshaft in an OHC system, which can cause serious ignition timing problems if the distributor is driven from the camshaft.

It can be seen that the choice of a valve train is very complex and that many other engine components and characteristics must be taken into account in making a selection. It is apparent that the valve train designer cannot be isolated from the overall design of the engine. This is becoming more and more apparent as engine specifications improve and engine speeds increase.

3. VALVE TRAIN DESIGN PHILOSOPHY

Several design parameters have usually been decided upon before the valve train of an engine is designed: the engine's displacement, the bore and stroke, the overall engine height, the maximum allowable manufacturing and assembly cost, the desired engine performance, etc. These parameters set restraints upon the final choice and design of valve train.

The performance of the engine is closely related to how efficiently the charge of air and fuel can be drawn into the combustion chamber and how efficiently the exhaust gases can be expelled after combustion. In order to admit and expel the various gases effectively, great care is taken over the design of the combustion chamber, the valve lift curve, the choice of the number of valves per cylinder and their angle of entry into the combustion chamber. The choice of valve train can often be virtually dictated by these conditions.

If, for example, the performance requirement for an engine dictates that the valve must be opened and closed very rapidly this implies that very high accelerations are required along the cam flanks. It is certain that push-rod systems would be eliminated from the choice of available mechanisms at this point due to their flexibility. OHC pivoted followers may also be eliminated from the choice as the maximum allowable flank acceleration is limited by the permissible concavity of the cam flanks. This may leave a direct acting OHC mechanisms as the only choice. The designer must then decide whether the large diameter followers required for this type of system can be fitted into the space available, and whether the luxury of two camshafts can be afforded if the valves are not in line.

The alternative types of valve train mechanisms obviously have differing advantages and disadvantages and the designer must weigh the requirements of his design in order of preference in order to get the right compromise. Below many of the good and bad points of various valve train mechanisms are summarised:

3.1 Advantages and Disadvantages of Pivoted Follower OHC Valve Train Systems

- Not as stiff as direct acting OHC mechanisms. This along with cam profile limitations due to cam concavity necessitates the use of modest valve accelerations. These restraints obviously limit the engine breathing. Maximum operating speed is less than that for direct acting OHC.

- Poor wear characteristics.

- Allows the use of only one camshaft even when there are two banks of valves.

- The cam lobes are smaller than those of a direct acting system as a mechanical advantage can be used by having rocker ratios greater than unity. This allows smaller camshaft bearings to be used.

- Adjustment is usually very easy.

3.2 Advantages and Disadvantages of Direct Acting OHC Valve Train Systems

- Good engine breathing as valve train is very rigid and therefore high valve accelerations and operating speeds are possible.

- Symmetrical cam profiles give symmetrical lift curves.

- Good lubrication conditions - low wear.

- If the valves are not in line then two camshafts are required unless a direct acting cam and follower is used in conjunction with a cam and pivoted follower.

- The cam lobes for a direct acting system are much larger than any other OHC system. This necessitates the use of large camshaft bearings.

3.3 Advantages and Disadvantages of Push-rod Valve Train Systems

- Very flexible. This allows only very low valve accelerations to be used thus limiting engine breathing. The low rigidity also limits the maximum allowable operating speeds.

- Low wear. The cam followers receive a plentiful supply of lubricant mist and splash from the main engine bearings due to their close proximity to the sump.

- A mechanical advantage may be used by employing a rocker ratio of greater than unity. This leads to small cam lobes.

- Easy adjustment.

- Engine auxiliaries driven by the camshaft suffer less detrimental effects than with any other system.

Once the type of mechanism has been decided upon the geometry of the system must be calculated. If a direct acting OHC system is used then the geometry can be readily decided. The diameter of the follower is dictated by the maximum eccentricity of the cam during the operating cycle. The maximum eccentricity is given directly by the maximum valve velocity (see Dyson and Naylor (1960)). The designers can then calculate whether there is enough space between the valves to accommodate the followers. Once the followers have been sized, an approximate value for the equivalent reciprocating mass of the valve train components can be calculated. This, along with the maximum required engine speed and valve acceleration, gives the minimum allowable spring stiffness to prevent valve bounce at the maximum rated engine speed. The spring stiffness is then chosen by allowing a given amount of spring cover across the nose region of the cam. The actual valve spring dimensions and number of coils are decided by available standard wire diameters, space limitations, an acceptable number of working coils, satisfactory fatigue life and dynamic considerations. Valve spring design is discussed in greater detail by Beard and Hempson (1962).

The cam size is chosen by satisfying the requirements of maximum allowable Hertzian stress, engine height, sliding speed at the contact, cam bearing diameter, and also, hopefully, by studying the lubrication conditions around the cam cycle for the range of operating speeds.

If a pivoted follower system is chosen the decision processes are similar but with additional variables, such as rocker ratios and symmetry to be considered. Having chosen a particular valve train the designer must then choose the correct materials for the contacting parts. This must be done with a knowledge of the stressing of the components and the sliding speeds and lubrication conditions expected at the contact. The choice of materials for cam and follower pairs is a complex subject worthy of study in its own right, and indeed many learned society papers have been written on the subject.

4. CAM AND FOLLOWER LUBRICATION

The received wisdom as to the mechanism of operation of cams and followers is that boundary lubrication pertains. It is, however, becoming clearer that lubrication analysis as a design tool has much merit (Taylor (1991)). A quarter of a

century ago Müller (1966) considered both theoretically and experimentally the characteristics of the two cam profiles shown in Figure 1.

On the basis of the value of maximum Hertzian stress at the nose the judgement would clearly be that cam 1 should have superior performance to cam 2. However, considering the predicted variation of oil film thickness the indication would be that cam 2 might be better. Engine performance tests using a steel camshaft and cast iron tappets showed that cam 1 suffered scuffing followed by pitting whilst the performance of cam 2 was satisfactory. The message was quite clear: a consideration of the lubrication performance of cams and followers, that is potential oil film generation, could be a fruitful aspect of tribological design. Although there was a limited further investigation of cam and follower lubrication before 1980, Müller's point seems to have been lost. The design of the most important element of the valve train continued to rely tribologically on materials

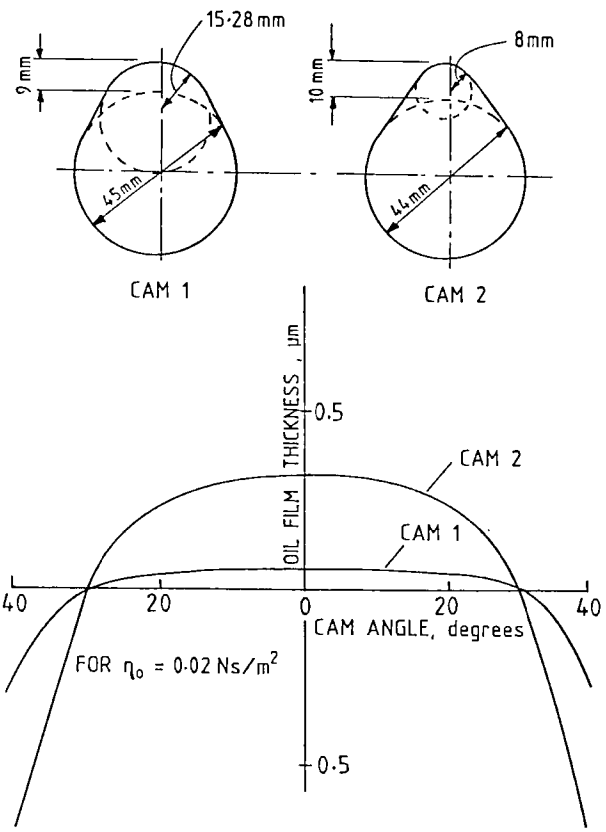


Figure 1 Cam Profiles Considered by Müller (1966)

specification, lubricant additive package choice and Hertzian stress calculations. However, the last decade has seen a renewed interest in lubrication analysis as applied to cam and follower performance. In this regard theoretical studies have addressed the prediction of film thickness, friction, wear and flash temperature with consideration of transient effects and mixed lubrication. In addition there has been substantial experimental work covering amongst other things the measurement of film thickness, friction, temperature and wear.

The remaining parts of this paper will therefore be given over to development of a lubrication design analysis for an automotive cam and follower. The general principles of such an analysis apply to all cam and follower forms (Figure 2) and the case of the direct acting (or cam and bucket follower) will be considered. It is pertinent to note that the interest in this particular geometry is widespread at this time as operational demands and environmental issues become increasingly pressing. For application of the analysis to alternative geometries, Ball et al (1989) is a suitable starting point.

5. THEORETICAL BASIS FOR LUBRICATION OF A DIRECT ACTING CAM AND FOLLOWER

5.1 Cam and Follower Kinematics

Consider the situation shown in Figure 3 with the contact point on the rising flank of the cam. The cam is designated component number one and the follower, which is taken not to rotate, number two. The cam rotates about (0) whilst (0') is the centre of curvature of the cam surface at the instantaneous contact point (C). The absolute velocities of the follower in the (x,z) co-ordinate directions are determined simply as:

$$u_2 = 0 \tag{1}$$

and

$$w_2 = e\Omega \tag{2}$$

The acceleration (a) of the follower may be calculated as:

$$a = \frac{dw_2}{dt} = \dot{e}\Omega \tag{3}$$

The x-velocity of the point on the cam in contact with the follower is:

$$u_1 = (R_b + L)\Omega \tag{4}$$

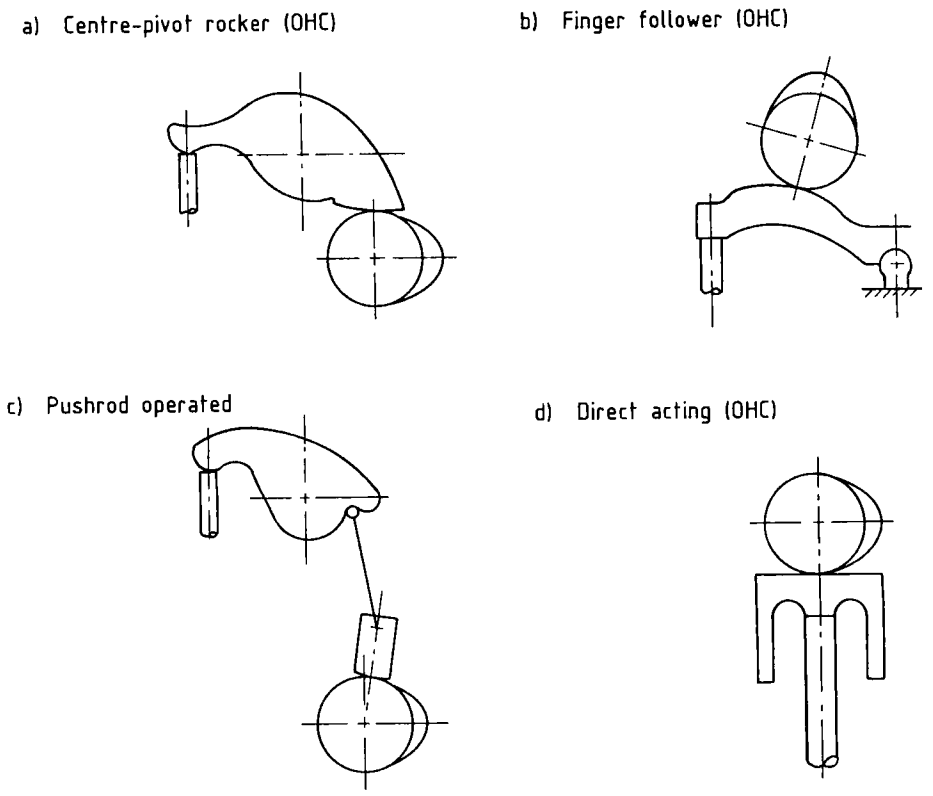


Figure 2. Alternative Cam and Follower Forms

The velocity of the contact point along the cam surface (s) may be calculated as follows with due regard for sign:

$$\dot{s} = u_1 + \dot{e} = (R_b + \dot{L})\Omega + \dot{e} \qquad (5)$$

The instantaneous radius of curvature of a point moving along a curve with velocity ($\dot{s}$) such that the angular velocity of the tangent to the contact point is ($\dot{\gamma}$) given by:

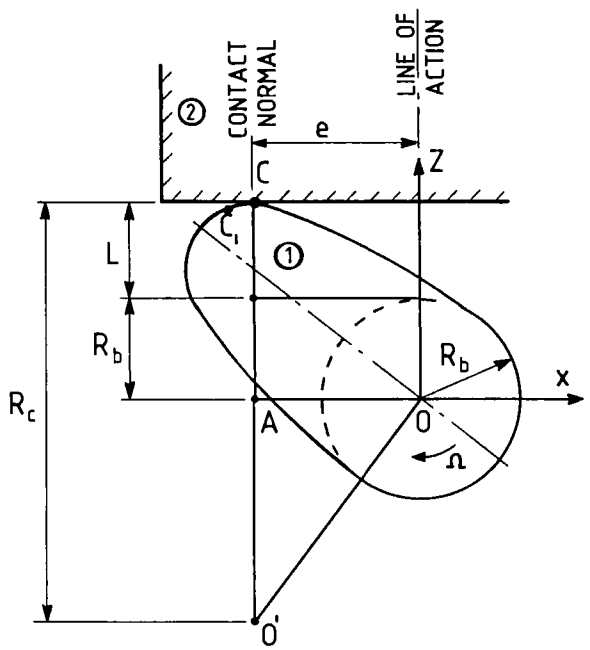


Figure 3 The Geometry of a Direct Acting Cam and Follower

$$R_c = \frac{\dot{s}}{\dot{\gamma}} = \frac{ds}{dt} \cdot \frac{dt}{d\gamma} = \frac{ds}{d\gamma}$$

For the flat faced follower:

$$R_c = \frac{\dot{s}}{\Omega} \tag{6}$$

Combining equations (5) and (6) we find:

$$R_c = R_b + L + \frac{\dot{e}}{\Omega} = R_b + L + \frac{a}{\Omega^2} \tag{7}$$

It will be noted that this analysis is valid for a general cam profile with a flat faced follower. For the purpose of application with the Reynolds' equation the absolute velocities of the cam surface in the (x,z) co-ordinate directions at a point (C_1) close to the nominal contact point (C) are required. To sufficient accuracy:

$$u_{C_1} = u_1 = (R_b + L)\Omega = \left[R_c - \frac{a}{\Omega^2}\right]\Omega \tag{8}$$

For the (z) co-ordinate direction (see Figure 4 also):

$$w_{C_1} = (e + de)\Omega = e\Omega + \Omega de \tag{9}$$

where the (z) velocity of the instantaneous contact point (C) may be seen to be given by:

$w_1 = e\Omega$

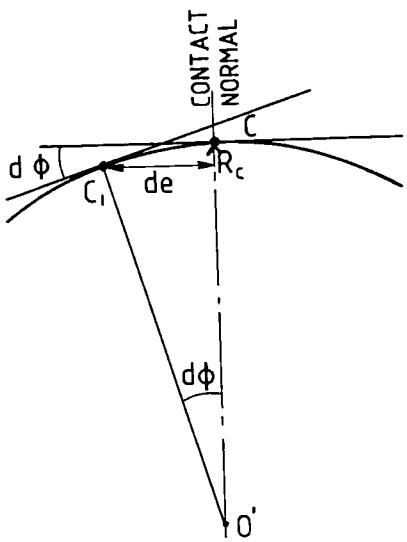


Figure 4 Direct Acting Cam and Follower - Details Close to the Nominal Contact Point.

From Figure 4 it may be seen that:

$R_c\, d\phi = de$

and $d\varphi = \frac{dz_1}{dx} = -\frac{dh}{dx}$

such that

$$w_{C_1} = e\Omega = R_c\,\Omega\,\frac{dh}{dx} \tag{10}$$

5.2 Contact Loading

The precise determination of the load to be carried by a cam and follower contact during a cycle is a complex matter. In general the forces associated with the operation of the mechanism are the inertia forces, the spring force, the friction forces arising as a result of relative motion between the contacting parts and the forces resulting from the stiffness and damping characteristics of the overall structure.

As a simplification to facilitate lubrication analysis the frictional forces might be neglected in the determination of contact loading and by taking the mechanism to be rigid, the stiffness and damping characteristics can also be omitted. The determination of cam load from calculation of inertia and spring forces therefore becomes straightforward. With the determination of contact load during the cycle of cam operation the calculation of the maximum Hertzian stress to which the solid materials are subjected is a simple matter. If the load per unit contact width at any instant is (P') then the corresponding maximum Hertzian stress (p_{max}) is:

$$p_{max} = \left[\frac{P'E'}{2\pi R_c}\right]^{\frac{1}{2}} \tag{11}$$

5.3 Lubrication Analysis

A general form of Reynold's equation for a fixed axes co-ordinate system may be written:

$$\frac{\partial}{\partial x}\left(\frac{\rho h^3}{12\eta}\frac{\partial p}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\rho h^3}{12\eta}\frac{\partial p}{\partial y}\right) = \frac{\partial}{\partial x}\left(\frac{\rho(u_1+u_2)h}{2}\right) + \frac{\partial}{\partial y}\left(\frac{\rho(v_1+v_2)h}{2}\right)$$

$$+\rho\left(w_2 - w_1 + u_1\frac{\partial z_1}{\partial x} - u_2\frac{\partial z_2}{\partial x} + v_1\frac{\partial z_1}{\partial x} - v_2\frac{\partial z_2}{\partial x} + h\frac{\partial p}{\partial t}\right)$$

For the line contact situation being considered here and for quasi-steady state conditions at a particular instant in time this equation reduces to:

$$\frac{d}{dx}\left(\frac{\rho h^3}{12\eta}\frac{dp}{dx}\right) = \frac{d}{dx}\left(\frac{\rho(u_1+u_2)h}{2}\right) + \rho\left(w_2 - w_1 + u_1\frac{\partial z_1}{\partial x} - u_2\frac{\partial z_2}{\partial x}\right) \tag{12}$$

Much published data is available reflecting solutions of this equation for a variety of situations. For the case of rigid counterformal surfaces with an isoviscous lubricant the classical work of Martin (1916) enables prediction, for example, of the minimum lubricant film thickness as a function of the operating conditions and geometry. With elastohydrodynamic lubrication predictions of minimum and central film thickness in a Hertzian contact as a function of the influential variables are also readily available (Dowson and Toyoda (1979)).

The use of the formulae for film thickness prediction necessitates the determination of the appropriate mean entraining velocity (u_e) of the lubricant into the contact at each instant under consideration. Care must be taken in establishing the correct value of this. The term $(u_1 + u_2)/2)$ appearing in the first term of the right hand side of equation (12) does not represent the mean entraining velocity for the situation under examination as shown in Figure 1. This is because the equation is written for a fixed co-ordinate system, with in this case the curved cam surface moving through space. Such situations can lead to confusion about the relative importance of pressure generation actions due to 'entrainment' and 'squeeze'.

The true mean entraining velocity must be established by considering both terms on the right hand side of equation (12). Using the values of surface velocity established in equations (1), (2), (8) and (10) and noting that $(\partial z_2/\partial x = 0)$ it can readily be shown that for an incompressible lubricant the Reynolds' equations reduces to:

$$\frac{d}{dx}\left[\frac{h^3}{12\eta}\frac{dp}{dx}\right] = \left[R_c + \frac{a}{\Omega^2}\right]\frac{\Omega}{2}\frac{dh}{dx} \qquad (13)$$

such that the mean entraining velocity (u_e) for any cam profile is:

$$u_e = \left[R_c + \frac{a}{\Omega^2}\right]\frac{\Omega}{2}$$

The latter expression may be deduced more directly in this case by consideration of the surface velocities in the x-co-ordinate direction relative to the contact, which is moving in this direction with velocity $(-a/\Omega)$ as can be seen from equation (3). Thus the surface velocities of the cam and follower relative to the contact become respectively, from equations (8) and (1), $(R_c\Omega)$ and $a/\Omega)$. The mean entraining velocity is quite clearly that rendered by the more formal approach presented previously.

5.3.1 Operational Characteristics

For steady state, elastohydrodynamic lubrication of a line contact the following expressions have been developed for the prediction of minimum and central film thickness respectively, in non-dimensional form (Dowson and Toyoda (1979)).

$$H_{min} = 2.65\, U_e^{0.7}\, G^{0.54}\, W_e'^{-0.13} \tag{15}$$

$$H_{cen} = 3.06\, U_e^{0.69}\, G^{0.56}\, W_e'^{-0.10} \tag{16}$$

The predictions of equation (15) can be used directly as an estimate of the instantaneous minimum lubricant film thickness during the cycle of cam operation. From this the film thickness parameter (λ), the ratio of this lubricant film thickness to the composite surface roughness, may be calculated at any point of contact. This parameter has proved of great value in assessment of the severity of lubrication of counterformal contacts.

The prediction of the lubricant film thickness at the centre of the nominal Hertzian contact given by equation (16) can be adopted in the assessment of contact friction and hence overall torque required to drive the cam. Approximating the lubricated contact between the cam and the follower as a region of constant film thickness (h_{cen}) and taking the friction to be determined by the sliding in this region (thus neglecting the contribution of any rolling friction either in the contact or the inlet zone) the friction (F) may be estimated as:

$$F = \int_{-b}^{b} \frac{\eta(u_1 - u_2)}{h_{cen}} \tag{17}$$

Substituting the Barus relationship for pressure dependent viscosity ($\eta = \eta_0 e^{\alpha p}$) and noting that the pressure in a Hertzian line contact is given by:

$$p = p_{max}\left[1 - \frac{x^2}{b^2}\right]^{\frac{1}{2}}$$

we obtain:

$$F = \frac{\eta_0 (u_1 - u_2)}{h_{cen}} \int_{-b}^{b} e^{\alpha p_{max}\left(1 - x^2/b^2\right)^{\frac{1}{2}}}\, dx \tag{18}$$

with (p_{max}) given by equation (11). This integral may be solved numerically. For most practical situations predictions of friction obtained using equation (18) will be several orders of magnitude greater than those obtained in practice. This is because the predictions of viscosity rendered by the Barus formula are unrealistic

at higher pressures. Rheological studies of the behaviour of lubricants in highly loaded non-conformal contacts have indicated a limiting traction coefficient (or coefficient of friction).

If the prediction of equation (18) for friction according to classical thin film fluid analysis results in a value of coefficient of friction (μ) greater than the limiting value, then the limiting value should be taken to apply.

The determination of frictional power loss (E) may be obtained numerically from:

$$E = \int_{cycle} F R_c \Omega dt \qquad (19)$$

6. CYCLIC FILM THICKNESS VARIATION

Dowson et al (1985) have described the development of a computationally based analysis to predict the tribological performance of a cam and flat faced follower. Here the discussion will be limited to the interpretation of predictions of the lubricant film thickness between the components. The general principles enunciated for this form of mechanism vis-a-vis film thickness will also apply to other cam and follower forms. The typical variation of film thickness is shown in Figure 5.

For the four power polynomial cam in question a separation of 0.4mm was maintained mechanically over the base circle region. The 'blips' in the traces of Figure 5 represent therefore the commencement and cessation of contact between cam and follower which were associated with ramp regions designed for a constant velocity of lift. The particular sector of interest is from -60°(the beginning of the opening flank) through 0° (the cam nose) to +60° (the end of the closing flank). The arrangement under examination was symmetric and the healthy film thickness over the flank regions, the modest film over the nose and the two critical areas of 'zero' lubricant film-thickness are evident in Figure 5(a).

An alternative presentation of cyclic film thickness variation which gives and improved physical interpretation is the 'toad' or 'butterfly' diagram shown in Figure 5(b) whilst the substantial film thickness over the flank regions of the cam is demonstrated in Figure 5(a), what is not evident is the extensive proportion of the cam surface which enjoys this satisfactory lubrication. Thus in Figure 5(a) the peaky nature of the film thickness variation on the flanks is indicative of the small time period associated with contact in these regions. Correspondingly the lengthy time period over the nose region with much small predicted film thickness, which is virtually constant, is linked with only a restricted surface extent of the cam. The latter area is clearly vulnerable to surface distress.

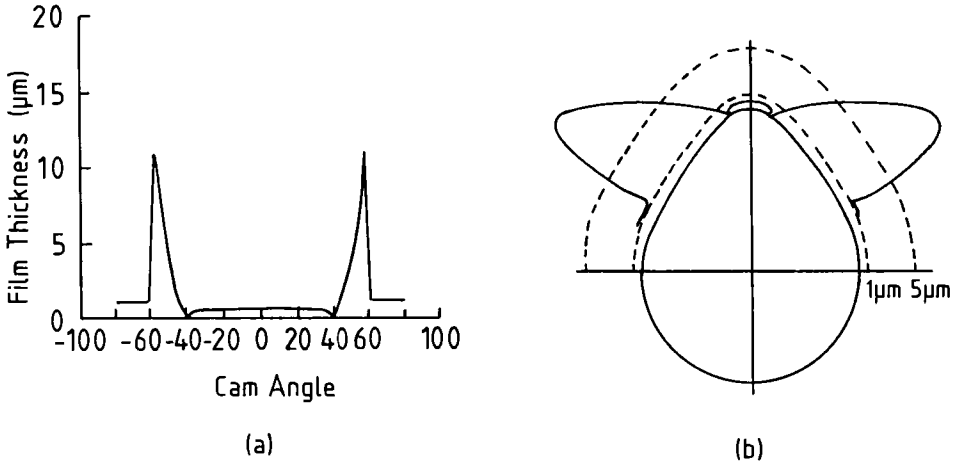


Figure 5 Typical Variation of Film Thickness Between Direct Acting Cam and Follower (a) As a Function of Cam Angle (b) Around the Cam Circumference.

The interpretation of the variation in lubricant film thickness during the cycle and identification of the factors which may influence this are matters of importance in design. It must first be noted that the analysis has been based on a quasi-static approach - that is each instant is considered a snapshot in time with conditions taken to be steady and no estimation of transient effects. Thus the nominal point of contact between the cam and follower traverses across the follower face whilst the cam rotates and the follower reciprocates in the plane of the camshaft. Two important parameters characterising the generation of film thickness under hydrodynamic or elastohydrodynamic conditions are the mean velocity with which the lubricant is entrained into the contact and the effective radius of curvature at contact. For the situation under consideration these are shown in Figure 6. Because the nominal contact point traverses across the face of the follower it will be seen that there is a change in direction of the lubricant entrainment. Thus there are two positions during the cycle where the entrainment velocity is zero and with the quasi-static analysis the film thickness is predicted to be zero at these points. The nature of the contact behaviour at these positions will be critical. It must be noted that, in reality, although the generation of pressure within the lubricant film due to what is known as the hydrodynamic wedge mechanism will be non-existent at these two points, the protective mechanism of the equally powerful squeeze film effect (which has not been considered) will be brought into play.

Over the flank regions both the entraining velocity and the radius of curvature of the nominal contact are substantial, guaranteeing a good lubricant film thickness.

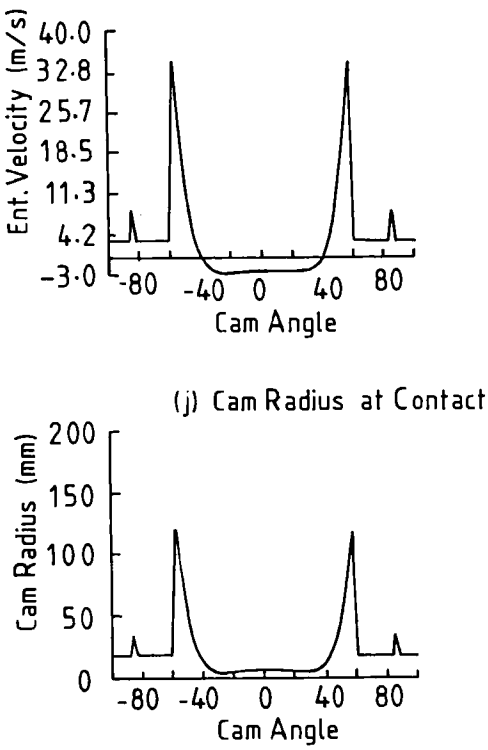


Figure 6 Variation of Mean Entraining Velocity of Lubricant and Contact Radius of Curvature as a Function of Cam and Angle for a Direct Acting Cam and Follower.

At the cam nose the radius of the cam is necessarily small, this being counter-productive to lubricant film generation. In addition the mean entraining velocity is modest. The relative constancy of these two quantities over the region of the nose explains why the film thickness varies little in this locality. It is instructive to develop the consideration of film thickness generation a little further by presenting the results of some straightforward calculations. The case considered is detailed in Appendix 1 with an operating dynamic viscosity of the lubricant of $0.01 Ns/m^2$. Two camshaft rotational frequencies will be taken - 1200rpm (20Hz) and 3000rpm (50Hz). The cam nose radius will be taken as 6mm and the film thickness also calculated on the flanks where the radius is at its maximum value of about 120mm. According to the line contact elastohydrodynamic lubrication theory the minimum lubricant film thickness predictions are shown in Table 1. At the camshaft frequency of 1200rpm the load transmitted between cam and follower was 540N at both the nose and flank locations whilst at 3000rpm the respective values were 275N and 2000N. A sample calculation is presented in Appendix 2.

Camshaft Rotational Frequency	Minimum Lubricant Film Thickness	
	Nose	Flank
1200 rpm (20 Hz)	0.095 μm	2.16 μm
3000 rpm (50 Hz)	0.20 μm	3.46 μm

Table 1. Predicted minimum film thickness at two locations at two camshaft rotational frequencies.

These calculated minimum lubricant film thicknesses may be considered in the light of typical surface roughness values of the components, namely 0.2μm (R_a) for both cam and follower. The film thickness parameter (or lambda ratio (λ)) - the ratio of the film thickness calculated on the basis of smooth surfaces to the composite surface roughness (R_a) - has become an important qualitative and quantitative parameter for a judgemental decision on the lubrication regime which is operative and the likelihood of surface distress in due course. On the flanks the (λ) value ranges between about 4 and 6.5 for the position and speed range examined. The surfaces are essentially separated and elastohydrodynamic lubrication would be taken to be operative. At the nose, however, the (λ) values for the two cases are about 0.2 and 0.4 suggesting boundary lubrication. Alternatively, one might think of the mode of lubrication at the nose as severe mixed lubrication (ie boundary and elastohydrodynamic effects). This is particularly true when it is borne in mind that, due to elastic distortion of the asperities, the surface roughness within the contact when under load may be substantially less than that measured in the metrology laboratory.

A simple calculation will now illustrate the important feature identified by Müller (1966) and noted by Naylor (1967). If it is imagined that the designer decided to increase the radius of curvature at the cam nose from 6mm to 13.7mm (not dissimilar to the situation identified in Figure 1) but retained the same overall value of base circle radius plus maximum lift, then the importance of the consideration of not only stressing but also lubrication can be clearly illustrated. For the same load condition the maximum Hertzian stress (being inversely proportional to the square root of the radius of curvature) would fall by 34%. The calculation in Appendix 2 indicates, however, that the mean entraining velocity of the lubricant into the contact would fall to zero. On the basis of a quasi-static analysis, therefore, the predicted film thickness would be zero. The prognostication would be that there

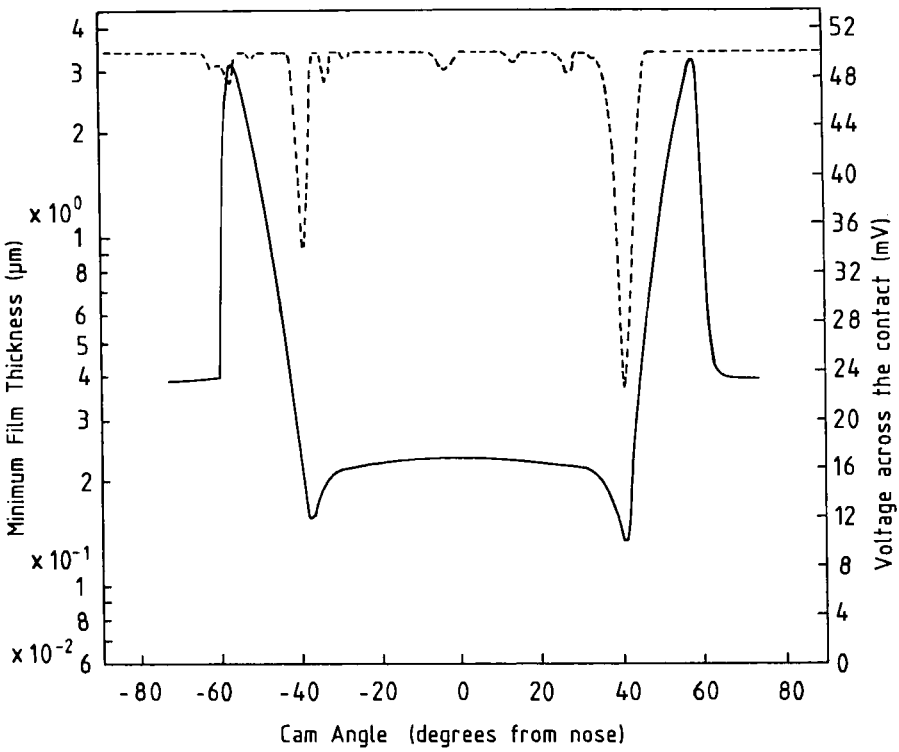


Figure 7 Predicted Transient Variation of Minimum Film Thickness Between a Cam and Direct Acting Follower (Solid Line) and Voltage Drop Measured by Electrical Resistivity Technique (dashed line) (Dowson et al (1991)).

REFERENCES

1 **Ball A D, Dowson D and Taylor C M** (1989), "Cam and Follower Design", of 15th Leeds-Lyon Symposium on Tribology - Tribological Design of Machine Elements, Butterworths, pp 111-130.

2 **Beard C A and Hempson J G G** (1962), "Problems in Valve Gear Design and Instrumentation", SAE National Power Plant Meeting, Philadelphia.

3 **Buuck B A,** (1982), "Elementary Design Considerations for Valve Gears", SAE 821574.

4 **Dowson D, Harrison P and Taylor C M** (1985), "The Lubrication of Automotive Cams and Followers", Mechanisms and Surface Distress, 12th Leeds-Lyon Symposium on Tribology, pp 305-322, Butterworths.

5 **Dowson D, Taylor C M and Zhu G** (1991), "A Transient Elastohydrodynamic Lubrication Analysis of a Cam and Follower", Frontiers of Tribology Conference IOP. To be published in Journal of Physics D : Applied Physics.

6 **Dowson D and Toyoda S** (1979), "A Central Film Thickness Formula for Elastohydrodynamic Line Contacts", Elastohydrodynamics and Related Topics, 5th Leeds-Lyon Symposium on Tribology, pp 60-65, MEP (IMechE).

7 **Dyson A and Naylor H** (1960), "Application of the Flash Temperature Concept to Cam and Tappet Wear Problems", IMechE, Proc Auto Div No 8, pp 225-280.

8 **Hoshi M** (1984), "Reducing Friction Losses in Automobile Engines", Tribology International, Vol 17, No 4, pp 185-189.

9 **Martin H M** (1916), "Lubrication of Gear Teeth", Engineering, London, 102, pp 199.

10 **Martin F A** (1985), "Friction in Internal Combustion Engine Bearings", in Combustion Engines - Reduction of Friction and Wear, Paper C67/85, IMechE Conference Publications, pp 1-7.

11 **Müller R** (1966), "The effect of Lubrication on Cam and Tappet Performance", Motor Tech Z, Vol 27, Pt 2, pp 58-61, MIRA Translation, No 27/66.

12 **Naylor H** (1967), "Cam and Friction Drives", Proc IMechE, Vol 182, Pt 3A, pp 237-247.

13 **Parker D A and Adams D R** (1982), "Friction Losses in the Reciprocating Internal Combustion Engine", in Tribology - Key to the Efficient Engine, Paper C5/82, IMechE Conference Publication, pp 31-39.

14 **Polak T A and Letts A** (1987), "When did you last Change your Camshaft ?", Automobile Engineer, Vol 12, No 3, June/July 1987, pp 6-8.

15 **Taylor C M** (1991), "Valve Train Lubrication Analysis", Proc of 17th Leeds-Lyon Symposium on Tribology in Vehicle Tribology, Elsevier, to be published.

APPENDIX 1

The calculations in section 6 have been carried out for the following automotive cam situation. The detailed geometry may be found in Dowson et al (1985).

Cam Half Period (Action Angle)	60°
Ramp Period	27°
Ramp height	0.5 mm
Ramp Velocity	0.02 mm/deg
Cam/Follower Clearance on Base Circle	0.4 mm
Maximum Lift	9.4 mm
Cam Width	12 mm
Base Circle Radius	18 mm
Valve Spring Stiffness	38 kN/m
Spring Displacement	6.4 mm
Spring Mass	0.045 kg
Valve Mass	0.157 kg
Young's Modulus of Cam/Follower	2.07×10^{11} N/m^2
Poisson's Ratio of Cam/Follower	0.29
Lubricant Dynamic Viscosity	0.01 Ns/m^2
Pressure-Viscosity Coefficient of Lubricant	2.2×10^{-8} m^2/N

APPENDIX 2

(a) Consider the calculation of film thickness at the nose at a camshaft rotational frequency of 3000rpm (50Hz).

$$R_c = R_b + L_{max} + \frac{a}{\Omega^2} \tag{1}$$

$$\therefore \frac{a}{\Omega^2} = 0.006 - 0.0274 = -0.0214\text{m}$$

$$u_e = \left[R_c + \frac{a}{\Omega^2}\right]\frac{\Omega}{2} \tag{2}$$

$$u_e = (0.006 - 0.0214)\left[\frac{2\pi.50}{2}\right] = (-)\ 2.42\text{m/s}$$

$$H_{min} = \frac{h_{min}}{R_e} = 2.65\ U_e^{0.7}\ g^{0.54}\ w_e^{-0.13} \tag{3}$$

$$U_e = \frac{\eta_o u_e}{E'R_e} = \frac{0.01.2.42}{(2.26.10^{11}).0.006} = 1.785.10^{-11}$$

$$G = \alpha E' = \left(2.2 \times 10^{-8}\right).\left(2.26.10^{11}\right) = 4972$$

$$W_e = \frac{P'}{E'R_e} = \frac{275/0.012}{(2.26.10^{11}).0.006} = 1.69.10^{-5}$$

$$\therefore h_{min} = 2.65\ (0.006)\left(2.99.10^{-8}\right)(99.1)\ (4.17)$$

$$h_{min} = 0.20\ \mu\text{m}$$

(b) Consider now if the radius of curvature of the cam at the nose location was increased from 6mm to 13.7mm whilst maintaining the value of the sum of the base circle radius and maximum lift above the ramp height ($R_b + L_{max}$).

From (1) $\frac{a}{\Omega^2} = 0.0137 - 0.0274 = 0.0137\text{m}$

From (2) $u_e = 0$ and hence $h_{min} = 0$

This Page Intentionally Left Blank

Engine Tribology / C.M. Taylor (Editor)

VALVE TRAINS – DESIGN STUDIES, WIDER ASPECTS AND FUTURE DEVELOPMENTS

Guangrui ZHU

Technology Department, John Crane EMA, Crossbow House, 40 Liverpool Road, Slough, SL1 4QX, United Kingdom.

1. INTRODUCTION

It has been argued that the lubricant film still plays a vital role in the design process of a cam and follower. The interactions of design parameters upon each other are so complex that they are often not fully appreciated by engineers. It was therefore apparent that a design aid was needed which would reveal how the various parameters affect the performance of the cam and follower.

A computer program has been developed to act as a design aid. The program is capable of analysing the kinematic and tribological performance of all types of cam and follower combinations in common use in automobile engines.

A parametric study of a cam acting directly against a flat faced follower will be presented and the results summarised in tabular form. The parametric studies were carried out using the program developed. It will be shown, using the example of a cam and flat faced follower, how a parametric study can reveal that small changes in a design may reproduce significant improvement on the tribological performance of the cam and follower system.

Modern thinking of the lubrication mechanism of the cam and follower has led to the view that some element of elastohydrodynamic lubrication must be occurring and that the mode of lubrication would be better described as 'mixed' and not 'boundary'. Theoretical studies have been carried out to predict the lubricant film thickness and friction and the cam and follower with consideration of transient effects and mixed lubrication. In addition, the theoretical predictions have been validated and correlated with experimental measurements.

Finally, a brief discussion of the failure mechanisms of the cam and follower will be presented. Some consideration of the future development of the valve train of the internal combustion engine is also given to conclude the paper.

In developing the concept of tribological design of the cam and follower the continuing theme of the direct acting follower will be addressed. However, many of the principles identified apply equally to alternative forms of the mechanism and comparable studying

of end and centre pivoted cams and followers can be found in Ball (1988) and Ball et al (1989).

1.1 Notation

e cam eccentricity
e_f cam offset
V_c velocity of contact point relative to cam
U_f velocity of contact point relative to follower
U_e resultant entraining velocity
Subscript
c cam
f follower

2. DESIGN SOFTWARE

Software has been developed within the Department of Mechanical Engineering at the University of Leeds which allows the tribological studies to be carried out upon any of the cam and follower designs in common use in today's internal combustion engines. The program is written in such a manner that it can easily be used by any competent professional engineer. It is believed that the program would be highly useful in the design of new valve trains.

3. PARAMETRIC STUDIES

In this paper a parametric study of a cam and flat faced follower mechanism is represented. In order to demonstrate how useful a tribological parametric study would be at the design stage, a suggested improved design is also presented.

3.1 A Parametric Study of a Cam and Flat Faced Follower System

The cam and follower pair used in this study is identical to that investigated by Harrison (1985). Data relating to the system is listed below and is taken as the datum condition:

Cam base circle radius = 18.00 mm
Cam width = 12 mm
Spring stiffness = 38254 N/m
Initial spring displacement = 6.40 mm
Equivalent mass at valve = 0.172 kg
Young's modulus of cam and follower = 207 GN/m^2
Poisson's ratio of cam and follower = 0.29
Lubricant viscosity = 0.050 Ns/m^2

Lubricant pressure-viscosity coefficient= 22 x $10^{-9} m^2/N$

The valve lift (l) is described by a four-power polynomial:

$$l = Y_r + l_{cmax} + C_p\left(\frac{\phi}{\phi_T}\right)^p + C_q\left(\frac{\phi}{\phi_T}\right)^q + C_r\left(\frac{\phi}{\phi_T}\right)^r + C_s\left(\frac{\phi}{\phi_T}\right)^s \quad (1)$$

where:

p = 2; q = 9; r = 78; s = 80 and
C_p = -11.6918407 mm, C_q = 2.8498942 mm,
C_r = -0.5894568 mm, C_s = 0.5314033 mm
Y_r = 0.50 mm, l_{cmax} = 9.40 mm,
ϕ_T = 60^o.

Each design parameter was changed from its datum, whilst the others remained constant, and the effect upon the predicted frictional power loss, lubricant film thickness at the cam nose and Hertzian stress at the cam nose was studied. The parameters changed were:

(a) Cam base circle radius
(b) Cam width
(c) Equivalent mass at the valve
(d) Camshaft rotational frequency
(e) Valve spring rate, and
(f) Lubricant viscosity

Figure (1) - in two parts shows the cam operating characteristics at the datum condition for a camshaft rotational frequency of 50 Hz (3000 rpm). Parametric studies were carried out at two reference rotational frequencies (16.67 Hz (1000 rpm) and 50 Hz (3000 rpm)). Figures (2), (3) and (4) identify clearly the behaviour of the dependent variables as the individual design parameters are changed. Each design parameter was changed from its datum value by −70% through to 300%, (0% change being the datum value, −50% change being half the datum value; +50% being one and a half times the datum value etc.). The important features are summarised in Table 1 with a rationalization of the variations. Only a few specific aspects will be identified briefly below.

3.1.1 Changes in cam shaft rotational frequency

Cam rotational frequency is of course not a design variable as such. Its influence is felt not only through the variation of the entraining velocity of the lubricant into the contact and the modification of sliding speed at the contact, but also through the effect on cam loading due to inertia forces.

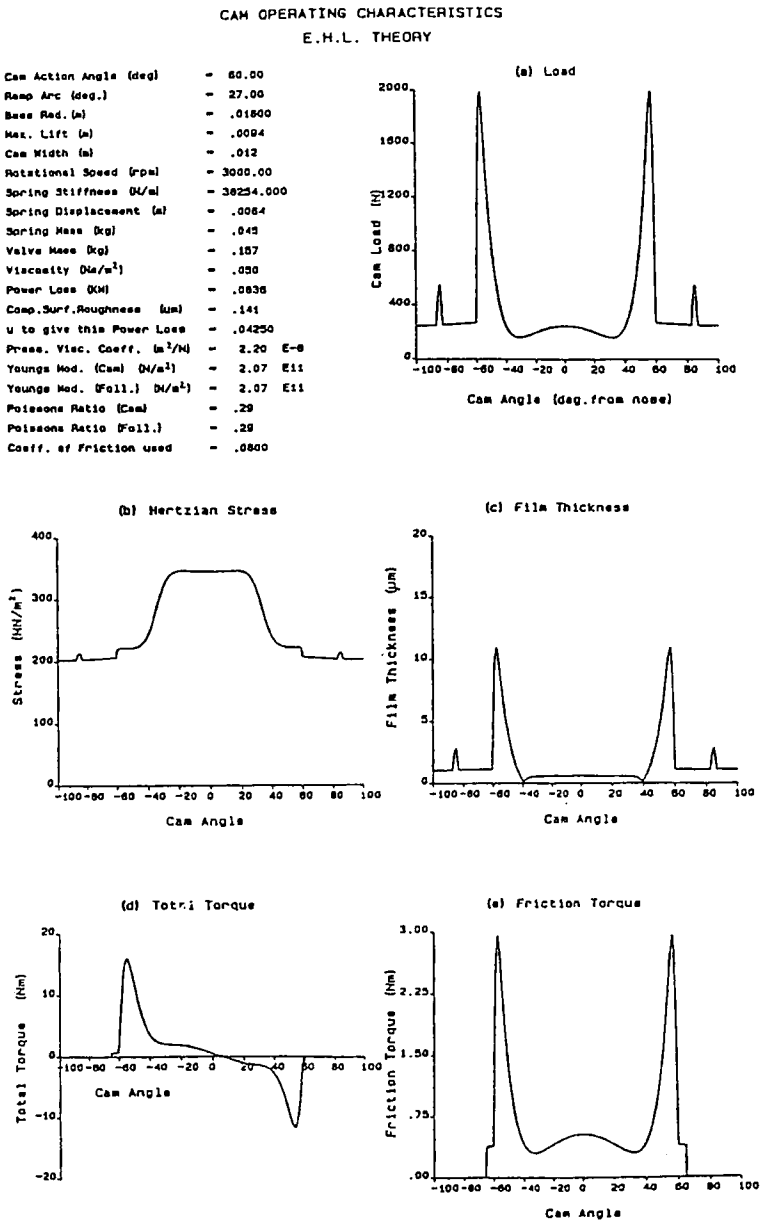


Fig. 1(a) Graphical output of computer program - operational characteristics

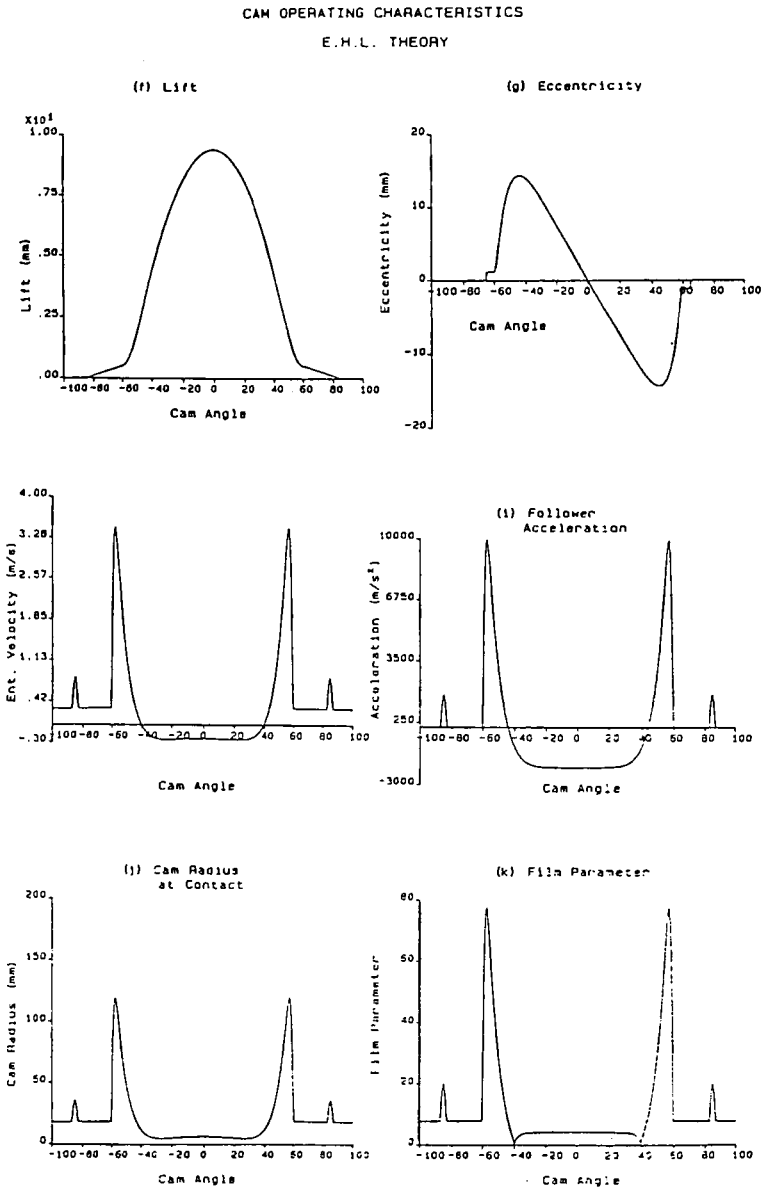


Fig. 1(b) Graphical output of computer program - interpretive data

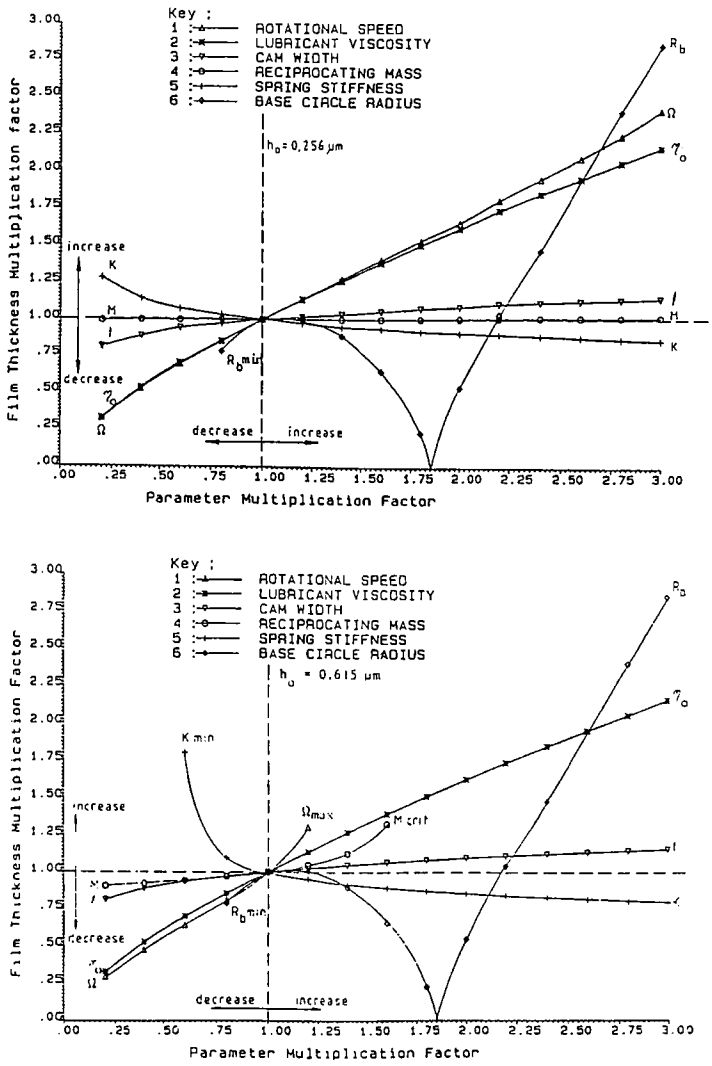


Fig. 2 (a,b) Effect of changes of parameters on minimum lubricant film thickness at the nose at a rotational frequency of 16.67 Hz (1000 rpm) (Figure 2(a) and 50Hz (3000 rpm) (Figure 2(b))

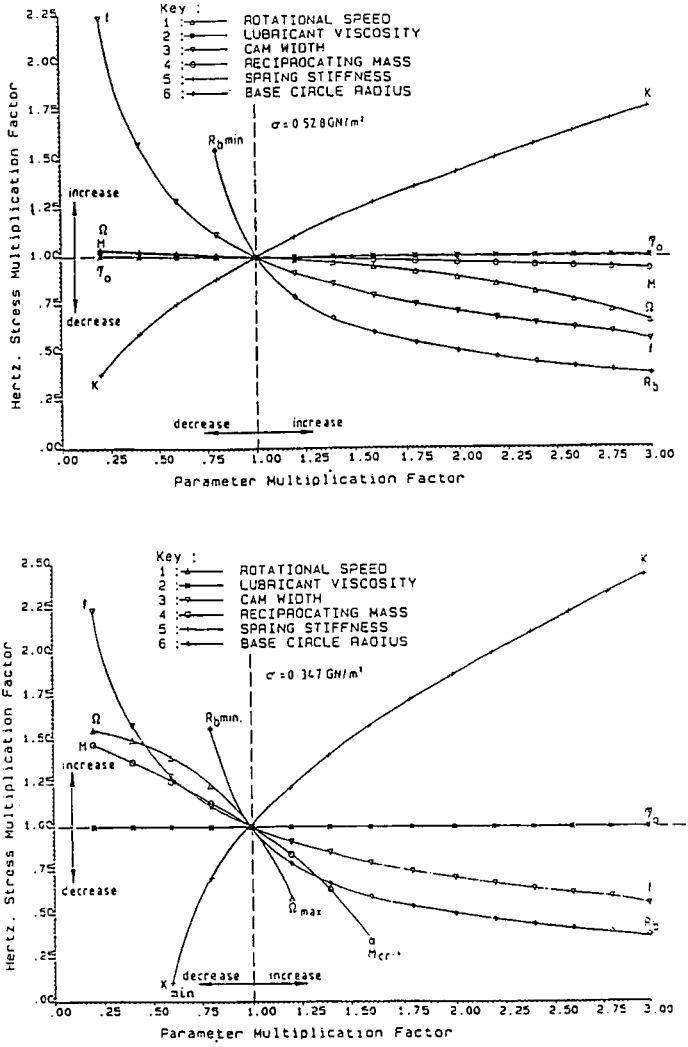


Figure 3(a,b) Effect of changes of parameters on maximum Hertzian stress at the nose and a rotational frequency of 16.67 Hz (1000 rpm) (Fig 3(a)) and 50 Hz (3000 rpm) (Figure 3(b))

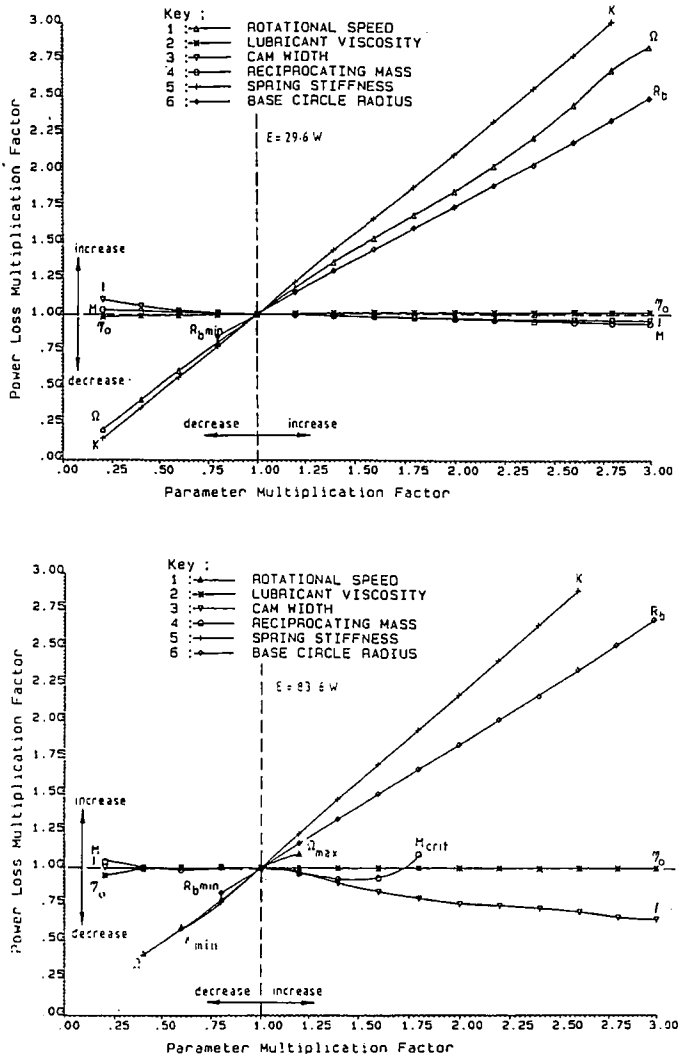


Figure 4(a,b) Effect of changes of parameters on power loss at a rotational frequency of 16.67 Hz (1000 rpm) (Figure 4(a)) and 50 Hz (3000 rpm) (Figure 4(b))

TABLE 1 Cam and Flat Faced Follower - Summarised results of Parametric Study

	MINIMUM FILM THICKNESS AT MAXIMUM LIFT POSITION	MAXIMUM HERTZIAN STRESS AT MAXIMUM LIFT POSITION	POWER LOSS	OTHER COMMENTS
BASE CIRCLE RADIUS	A decrease from the design value causes a fall in film thickness as the equivalent radius of curvature falls. An increase causes the film thickness to fall dramatically, before rising, due to the mean entraining velocity across the nose passing through a zero value.	The Hertzian stress is inversely proportional to the square root of the equivalent radius of curvature. Increasing the base circle radius therefore reduces the Hertzian stress at the cam nose.	The limiting coefficient of friction is applied over most of the cycle, hence the friction is a function of load rather than film thickness. The increase in power loss with base circle radius for a given speed is due to the increase in instantaneous radius of curvature (the distance at which the frictional force acts). If the base circle radius is increased sufficiently the entrainment velocity is enhanced adequately such that the limiting coefficient of friction is no longer applied over any of the cycle.	Increasing the base circle radius will increase the overall engine height. If on assembly, the camshaft needs to be fed through the camshaft bearings, then larger diameter bearings must also be used. A larger cam base circle radius allows a larger diameter camshaft to be used which will increase the rigidity of the valve train system; however at the expense of extra weight. Minimum value of base circle radius is dominated by the lift curve characteristics and the necessity for the cam to become concave along the flanks.
CAM WIDTH	As the cam width increases the load per unit width at the contact decreases. This leads to a small increase in the elastohydrodynamic film thickness.	As the cam width increases the load per unit width at the contact decreases, and so, therefore, does the Hertzian stress.	At low rotational frequency the limiting coefficient of friction is applied over most of the cam cycle. Hence, an increase in cam width only marginally reduces power loss. At higher rotational frequencies the proportion of the cycle governed by boundary lubrication falls with increasing cam width and there is a corresponding fall in power loss.	The cam lobe width is limited by the number of cam lobes to be fitted into a given space, i.e. the cam lobe width is governed by the valve spacing. A larger cam width necessitates a larger diameter follower which leads to a larger reciprocating mass.
RECIPRO-CATING MASS	An increase in reciprocating mass causes an increase in the film thickness due to the inertia force reducing the cam load at the nose.	Increasing the reciprocating mass reduces the cam load and hence lowers the Hertzian stress at the cam nose.	Increasing the reciprocating mass generally causes a fall in the power loss as the loading at the cam nose is reduced.	Increasing reciprocating mass increases the load on the cam flanks. decreasing mass may cause strength problems. increasing mass inertia problems.
VALVE SPRING STIFFNESS	Increasing valve spring stiffness reduces the film thickness as the cam load is increased.	A reduction in spring stiffness reduces the cam load and hence reduces the Hertzian stress.	An increase in spring stiffness causes the power loss to increase as the cam load increases.	The valve spring must provide sufficient load to restrain valve bounce and to control the valve along the cam flanks.
LUBRICANT VISCOSITY	Increasing the lubricant viscosity of the lubricant increases the film thickness.	No influence.	Insignificant influence.	Lubricant viscosity may be dictated by other engine component lubrication requirements.
CAM ROTATIONAL SPEED	Increasing speed increases entraining velocity and reduces cam load at nose. Film thickness therefore increases with increasing speed. At high speeds the cam load falls rapidly and therefore the increase in film thickness is rapid.	As the cam rotational speed increases the load decreases at the cam nose and therefore the Hertzian stress decreases.	Power loss increases with cam speed. At high speeds the rate of increase falls as valve bounce is approached.	Camshaft speed is limited by the point at which valve bounce occurs and the cam no longer controls the valve. This is due to the inertia of the reciprocating parts being greater than the valve spring load.

3.1.2 Changes in lubricant viscosity

The influence of lubricant viscosity on minimum film thickness at the nose is substantial, as would be expected, however its variation cannot be considered independently of other features.

3.1.3 Changes in reciprocating mass

Variations of reciprocating mass and valve spring stiffness have implications as regards valve train vibrational characteristics which would have to be considered in addition to the effects highlighted.

The beneficial effect of increasing reciprocating mass and cam width on reducing the maximum Hertzian stress at the nose is apparent. Also, reducing valve spring stiffness is particularly valuable as a means of reducing both the maximum Hertzian stress at the nose and the power loss.

The base circle radius is a particularly influential parameter as regards the effect of its variation on the three factors highlighted. The variation of base circle radius although it affected the physical dimensions of the cam, did not influence the lift or acceleration characteristics of the follower.

For the situation examined, as the base circle radius increased, the predicted minimum lubricant film-thickness at the nose fell to a zero value before increasing. The beneficial increase of the effective radius of curvature as regards film thickness generation is coupled with the influence upon the mean entrainment velocity of the lubricant into the contact. The latter passes through a zero value explaining the observed effect. An increase of base circle radius is also significant in reducing the maximum Hertzian stress at the nose and increasing power loss.

3.1.4 Changes in spring rate

As the spring rate increases the Hertzian stress at the cam nose increases due to the load increasing. This increase in load also causes the film thickness to decrease but the changes are not large. The frictional power loss increases as both central film thickness and maximum Hertzian pressure increase. If the limiting coefficient of friction is applied throughout the majority of the cycle, then the change in frictional power loss would be expected to be almost linear with change in spring stiffness.

3.1.5 Changes in cam lobe width

As would be expected the Hertzian stress at the cam nose will decrease as the cam lobe width is increased as the load per unit width of cam lobe decreases. The film thickness increases and the frictional power loss decreases as the cam lobe width is increased also as a direct result of this. The designer is limited in choice of cam width by the spacing between valves.

3.2 Enhancement of a Cam and Flat Faced Follower Design

Some feel for the potential overall benefits on cam performance which may accrue from the more judicious choice of influential design parameters may be obtained from the study presented in Table 2. The values of minimum lubricant film thickness and maximum Hertzian stress at the nose of the cam, together with the power loss, are detailed for rotational frequencies of 16.67 Hz (1000 r.p.m.) and 50 Hz (3000 r.p.m.). The cumulative influence upon these quantities of successive changes in base circle radius, cam width, valve spring stiffness and reciprocating mass are then presented. Finally, the percentage overall changes from the original values are listed.

The results of Table 2 demonstrate that straightforward, realistic modifications in the design of a cam and flat faced follower can effect an increase in minimum lubricant film thickness at the nose of about 10% and reductions in maximum Hertzian stress at the nose and power loss of the order of 30% and 20% respectively. The influence of modifications to the base circle radius and valve spring stiffness are particularly evident. The potential benefit both to the life and efficiency of the cam mechanism of such changes are significant and worthy of consideration.

Table 2. The effect of modification of specific parameters upon some important operating characteristics of a cam and flat faced follower

% Change of parameter from original design value	ROTATIONAL FREQUENCY 16.67 Hz (1000 rpm)			ROTATIONA FREQUENCY 50.0 Hz (3000 rpm)		
	Film Thickness (μm)	Stress (GPa)	Power Loss (W)	Film Thickness (μm)	Stress (GPa)	Power Loss (W)
0.0, As Designed	0.256	0.528	29.6	0.615	0.347	83.6
+11.1, R_b = 20 mm	0.262	0.458	32.0	0.635	0.301	91.5
+16.7, l = 14 mm	0.267	0.424	31.8	0.643	0.279	90.2
−21.6, K = 30KN/m	0.276	0.372	24.8	0.710	0.190	66.0
−4.50, M = 0.15 Kg	0.276	0.374	24.5	0.697	0.225	64.1
% Change from orininal Calculated values with all design modifications	+7.80	−29.2	−17.2	+10.4	−35.2	−22.3

Similar parametric studies of a cam acting against a centrally pivoted follower and a cam acting against an end pivoted follower were also carried out. Details can be found in Ball (1988) and Dowson et al (1988).

4. MORE SOPHISTICATED TRIBOLOGICAL ANALYSES OF THE VALVE TRAIN

As discussed in the previous sections, the squeezing action of the lubricant film and the surface roughness play an important role in the tribological performance of the valve train. More comprehensive tribological analysis has been undertaken to investigate the influence of the squeezing action and surface roughness on the lubricant film and the friction of the cam and follower.

4.1 A Transient EHL Analysis of a Cam and Follower

Quasi-static elastohydrodynamic lubrication analysis of a cam and flat faced follower reveals the typical variation of film thickness during a cycle.

A particularly important feature of cam and follower lubrication is evident in Figure (1b). At two stages during cycle, where the entraining velocity is zero, the quasi-static lubrication analysis predicts no film thickness. This entraining velocity is consequent not only upon the absolute velocities of the surface components but also the movement of the contact point itself (Dowson et al (1986)). The entraining velocity changes direction twice during a cycle resulting in the two critical points. Such behaviour is not only characteristic of the cam and flat faced follower (or cam and bucket follower) but also all varieties of the mechanism.

Whilst the quasi-static elastohydrodynamic lubrication analysis provides valuable design information and may be effected simply, a detailed study of the two critical points during the cycle is important if more sensitive assessments of potential failure and parametric influences are to be provided. Such assessments require the much more complex transient form of analysis which will enable the influence of the squeeze film lubrication mechanism to be exposed. A complex transient elastohydrodynamic lubrication analysis of the contact between a cam and flat faced follower has been undertaken (Dowson et al (1991)). The geometry considered reflected on actual automotive cam operating under realistic conditions. Four analytical approaches were developed. The results of the analysis demonstrated:

(a) Whereas previous quasi-static analysis predicted the occurrence of zero film thickness at the point of zero lubricant entraining velocity on the opening and closing flanks, the transient considerations revealed the protective and damping action of squeeze film lubrication at these positions. Indeed the minimum separation at these critical points was shown to be about half the minimum film thickness at the cam nose location. This is a significant finding.

(b) It has been demonstrated that the minimum separation occurs on the closing flank. Experimental evidence in support of this conclusion has been forthcoming but the physical explanation of why this should be so is not immediately obvious.

(c) The analytical approach developed offers the prospect for the first time of optimizing the operational conditions to promote good lubrication at the critical locations during cycle where there is a propensity for surface distress.

(d) The complete transient elastohydrodynamic analysis reported has given a benchmark against which realistic simplifications, to effect more easily a wider range of solutions whilst maintaining confidence in results, can be investigated.

4.2 Mixed Lubrication Analyses of a Cam and Flat Faced Follower

The conditions at the cam and follower contact are very severe. The quasi-static analysis of the lubricant film thickness has shown that the contact around the cam flanks may enjoy elastohydrodynamic lubrication, while the contact on the cam nose would operate in the mixed or boundary lubrication regime. On the cam nose, therefore, the surface roughness of the cam and follower has important effects on the tribological

performance of the valve train. A mixed lubricating analysis of a cam and flat faced follower has been undertaken (Dowson et al (1986)) with the assumption of rigid bulk component. This analysis has been further extended to cover the elastohydrodynamic lubrication of the cam and follower with consideration of surface roughness (Zhu (1988)). The analysis of a four-power polynomial cam and a bucket follower has yielded following conclusions:

(1) The cyclic variation of the nominal minimum film thickness is greatly enhanced by including the effects of surface elastic deformation, surface roughness and the variation of lubricant viscosity with pressure. The general features of the cyclic variations of the nominal minimum film thickness between the cam and follower are similar those predicted by the rigid-isoviscous solution for rough surfaces and the full numerical transient elastohydrodynamic solution for smooth surfaces.

(2) Surface roughness exerts a significant influence upon the lubrication performance of the cam and follower. The increasing of the composite surface roughness height causes an increase of the nominal minimum film thickness but a decrease in the film parameter.

(3) The effect of the distribution of surface roughness between the cam and follower upon the nominal minimum film thickness varies in different parts of the cam lift portion. Comparing the results with those for the case of the same roughness on cam and follower, in the cam nose region a smooth follower and rough cam generates a higher predicted film thickness, whilst a smooth cam and rough follower produces a lower film thickness. By contrast, on the cam ramps the effects are reversed.

(4) Under certain circumstances, such as the reference case adopted by Zhu (1988), the cam and follower works under mixed and elastohydrodynamic lubrication conditions for the greater part of its lift portion. This implies that the nature of the contact between the cam and follower is not as severe as it is generally envisaged to be.

5. EXPERIMENTAL EVIDENCE

It is important to validate experimentally the predictions of theoretical analysis and in relation to lubricant film thickness between a cam and bucket follower and the friction of the valve train. Three sources of empirical data will be reviewed (Zhu (1988), Williamson et al (1989) and Dowson et al (1990)).

5.1 Cyclic Film Thickness

Williamson et al (1989) used a capacitance technique to infer the oil film thickness between a cam and follower in a motored head driven at 1600 rpm (26.67 Hz). Seven fully formulated oils, without the inclusion of a viscosity index improver, were tested and the oil film thickness variation for the authors lubricant A is reproduced in Figure (5). This particular oil was an SAE Grade 20W having a dynamic viscosity of 4.97 MPa.s at $100°C$. The general similarity between the curve of Figure (5) and the film thickness variation of Figure 1(c) is evident. Indeed, the magnitudes are not dissimilar to the

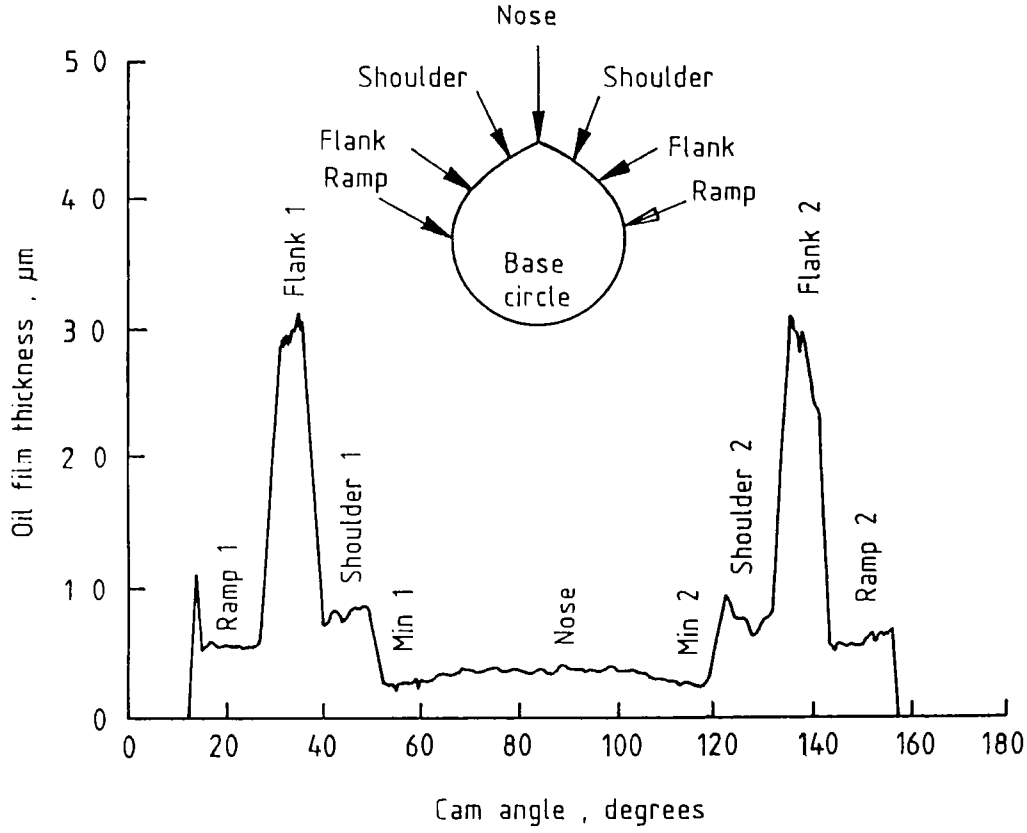


Figure 5 Sketch of voltage drop between cam follower to aid interpretation of Figure 1(a)

example calculation presented in Section 3. Williamson et al (1989) noted that for their cam follower system, "the magnitude of the variation (of oil film thickness) with viscosity is strongly dependent upon the position within the cam cycle. The response is greatest in the flank regions and least in those regions of minimum oil film thickness and on the cam nose". This observation provides a valuable correlation with the results of analysis. Firstly, if lubricant viscosity is influential in the development of film thickness then the lubrication mode being encountered is not that normally understood as boundary lubrication. Secondly, as noted from the analytical predictions presented in Section 3, on the flanks the correlation of oil film thickness with dynamic viscosity is best (elastohydrodynamic lubrication with minimal asperity contact) whilst this is less so over the nose region (mixed lubrication with significant surface interaction).

One of the conclusions of Williamson et al is worthy of recording, namely that "for all the oils tested there was always a lubricant film present between the cam and follower surfaces". The oils had a range of dynamic viscosity variation from 5 to 23MPa.s at $100°C$. Whilst the theoretical predictions of oil film thickness on the flanks did underestimate the measured values, particularly for the thinner lubricants, the importance of the elastohydrodynamic generation of oil film thickness was clearly established.

Dowson, Harrison, Taylor and Zhu (1990) used an electrical resistivity technique to measure what they describe as film state. This qualitative assessment of film thickness involved the determination of voltage drop between an appropriately insulated cam and bucket follower. With no contact a 50 millivolt potential drop existed between the components, whilst if they come into contact this would drop to zero. Five oscilloscope traces are shown in Figure (6) for a camshaft rotational frequency of 1200 rpm (20 Hz). The traces were taken at start up and then after 15 minutes, 1 hour, 25 hours and at the end of a test, 100 hours. The level of the trace is a measure of the drop from a no contact situation to one of intimate contact between the surfaces.

The sketch of Figure (7) will serve to illustrate the interpretation of Figure (6a), the start up trace. The 'no-contact' level represents essentially a potential difference of 50 millivolt between the surfaces corresponding to a high resistance. The 'contact' level implies zero potential difference between cam and follower, that is they are in contact. Thus Figure (7) can be interpreted as follows (symmetric cam).

A-B : cam and follower separated by 0.4 mm over base circle.

B-C : constant velocity ramp of cam contacts follower.

C-D : rising flank with good separation.

D-E : nose region with a very thin (partial) oil film.

E-F : falling flank with good separation.

F-G : constant velocity ramp with parting of components.

G-H : base circle region.

The initial trace of Figure (6a) once again presents a solid experimental confirmation of the trends of oil film thickness generation between a cam and follower presented in

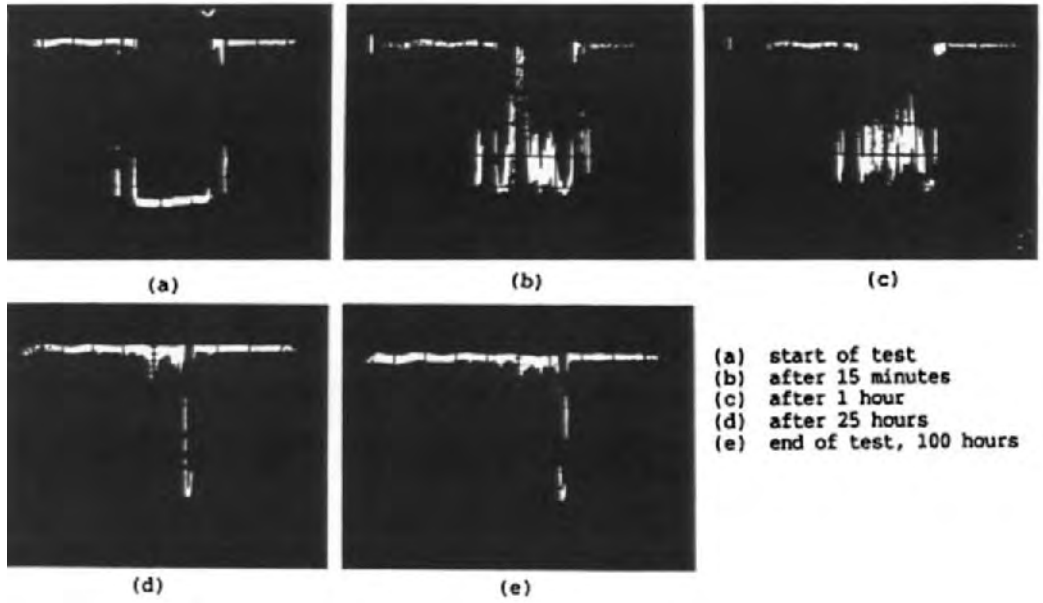


Figure 6 Electrical resistivity (voltage drop) traces during 100 hours of running at a camshaft rotational frequency 20 Hz (1200 rpm) and bulk temperature of about 50°C

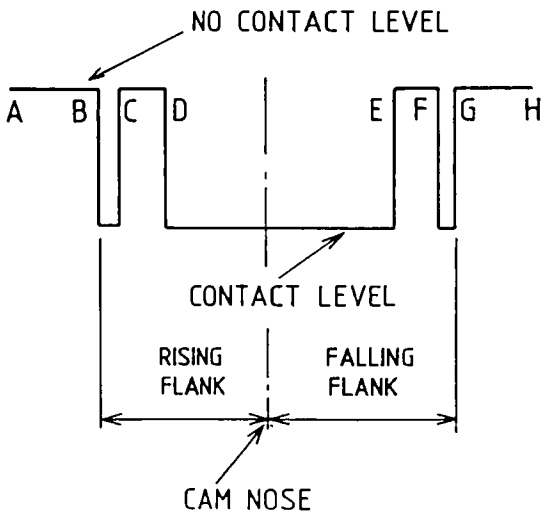


Figure 7 Sketch of voltage drop between cam follower to aid interpretation of Figure 6(a)

Figure (1). At the end of the tests recorded in Figure (6) it will be seen that the voltage trace remains at the 'no contact' level for virtually the whole cycle which might be interpreted as there being a substantial film separating the cam and follower for the whole action period. However, the presence of thin reaction films on the surface having electrical insulation properties has confused the understanding of the traces by this time. This point was also noted by Williamson et al (1989).

The electrical resistivity technique of film state assessment also provided a remarkable correlation of film thickness variation at the locations of zero entraining velocity. The details are presented in Figure (8) after Zhu (1988), these results being noted by Dowson et al (1990). Whilst a quasi-static analysis of cam and follower lurbriction will predict zero film thickness at the positions of zero entraining velocity, for such situations a consideration of squeeze film effects will reveal the protective action of this mechanism and result in a finite film thickness and indication. Zhu (1988) undertook such a transient, elastohydrodynamic analysis and the results of one aspect of his study are given in Figure (8). Here on the opening flank as the cam and follower are approaching each other (cam angle about $-40°C$), the predicted lubricant film thickness is marginally higher than that at the zero entraining velocity position on the closing flank. This sensitive difference was validated by the electrical potential measurements where as can be seen the drop from the 'no contact' position was greater for the thinner film. The difference in theoretical predictions for these two points relates to time history and possibly cavitation influences.

Further evidence which supports the theoretical predictions comes from inspection of the follower surface. Zhu (1988) inspected the wear track of the follower surface after testing with a Talysurf machine. The lubricant film thickness predicted by a transient elastohydrodynamic lubrication analysis with the consideration of surface roughness (Zhu (1988)), is compared with the Talysurf trace of the wear track. Figure (9) shows the cyclic variation of the lubricant film thickness along the follower surface. The deep groove at the right hand end of the trace corresponded to the zero entraining velocity region near the cam closing flank, whilst the shallower groove located at the left hand end of the trace related to the other zero entraining velocity region near the cam opening flank. Interestingly, the variation of the lubricant film thickness exhibited the same tendency as the wear profile on the follower surface with not only the position but also the depth of the wear profile being in harmony with the predicted lubricant film thickness. The two grooves observed in the middle of the tract reflect the two points cut which the cam initiated and completed contact with the follower due to a 0.4mm clearance between the cam base circle and the follower.

5.2 Friction and Wear

Whilst the correlation between theoretical and experimental assessment of lubricant film thickness generation between an automotive cam and follower lends the strongest

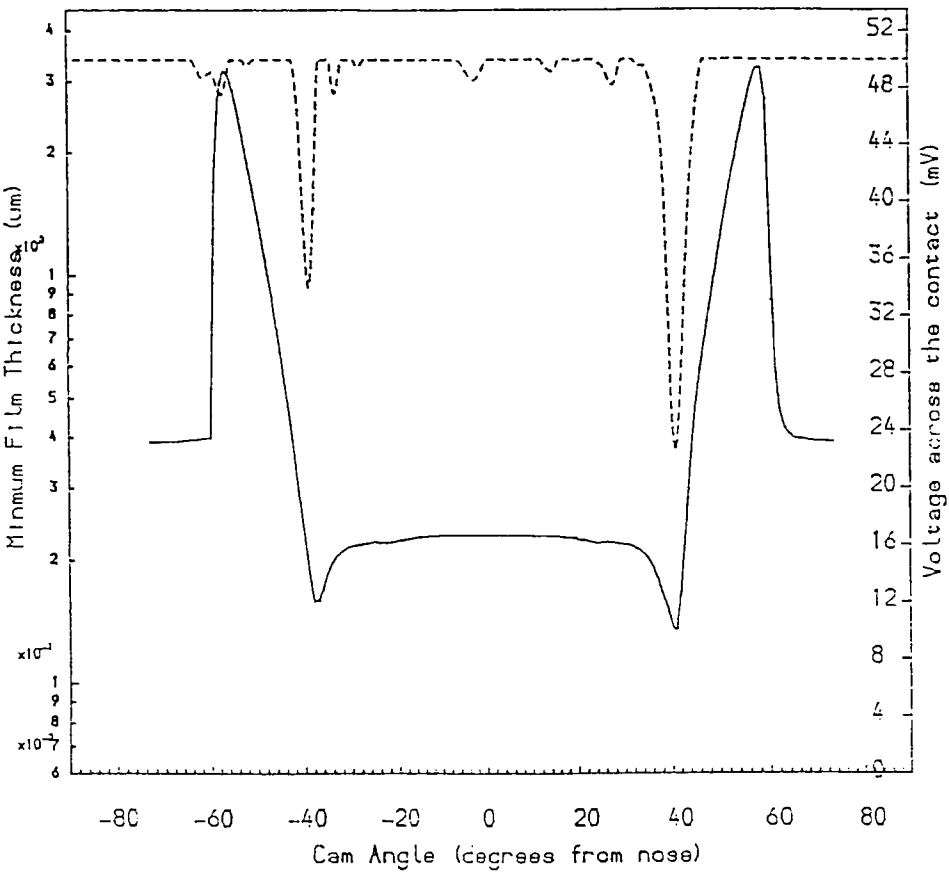


Figure 8 Predicted oil film thickness and measured voltage drop between cam and follower. (Bulk temperature 50°C rotational frequency 20.7 Hz (1240 rpm)

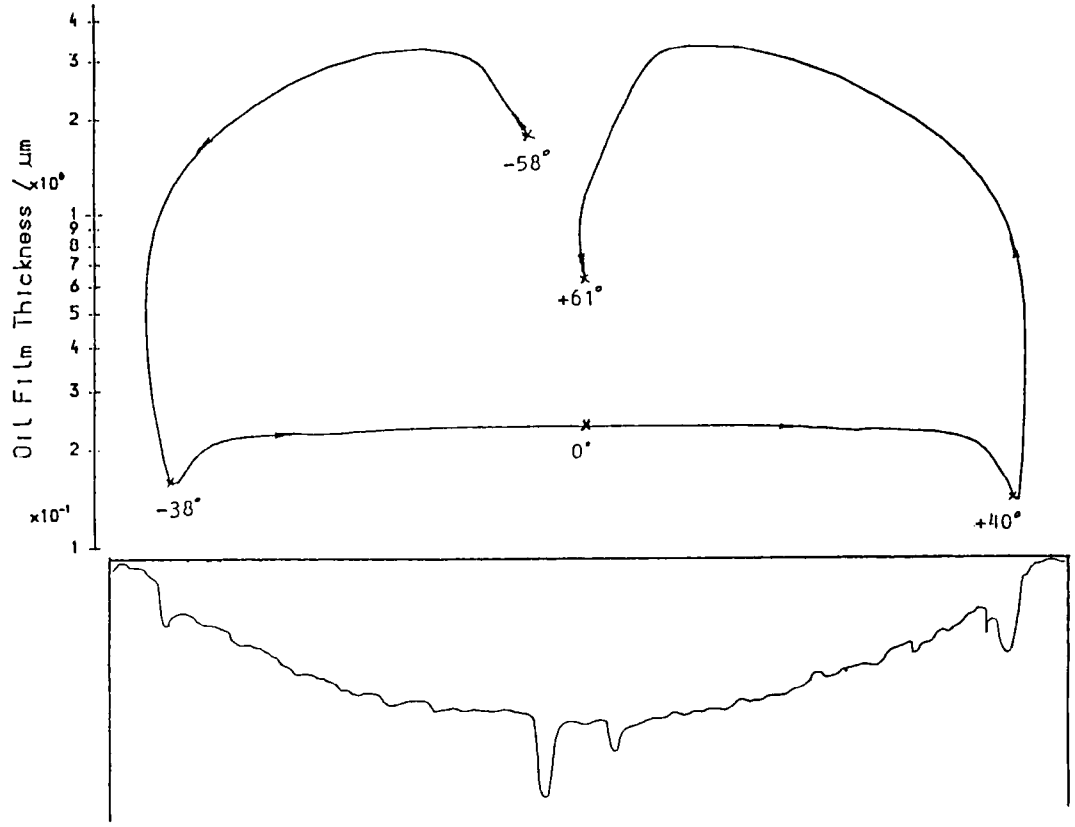


Figure 9 Comparison of the talysurf trace of wear tracks on the follower with the oil film thickness predicted by the lubrication analysis

support to the noting that thin film hydrodynamics has a role to play in addition to boundary lubrication influences, other evidence is available.

Dowson et al (1989) undertook an experimental study of the friction and power loss in an automotive four power polynomial cam and flat faced bucket follower. The influence of camshaft speed and bulk temperature upon friction was examined and encouraging agreement with theoretical prediction was obtained. The theoretical model developed in association with the study enabled the transient, elastohydrodynamic analysis of a line contact cam and follower to be undertaken. The simplified analysis had been based upon a more rigorous study and enabled a consideration of mixed lubrication. The roughness of the surfaces was incorporated into an averaged Reynolds equation after the method of Patir and Cheng (1978). In the prediction of friction the approach of Patir and Cheng was also followed, but when the shear stress exceeded the shear strength of the boundary film, the latter was taken for computing friction force. The boundary shear stress due to asperity interaction was taken to be a linear function of the local pressure.

The agreement between predictions of the lubrication based model for estimating friction and experimental measurements was good. Figure (10) is a plot of friction coefficient averaged over a cycle against chamshaft speed for a bulk temperature of $105°C$. Apart from the lower camshaft speeds, where severe mixed lubrication occurred and the theoretical model adopted would be inappropriate, agreement of experiment and theory to within 5% for friction over the speed range 650 rpm (10.88 Hz) to 2000 rpm (33.3 Hz) was obtained. Similar agreement was observed for a constant speed test with the bulk temperature of the cam and follower contact varying from $50°C$ to $120°C$.

Not only was the theoretical/experimental agreement of frictional torque averaged over a cycle satisfactory but even the instantaneous frictional torque during a cycle bore acceptable comparison. Experimental friction torque for two camshaft rotational frequencies at bulk temperature of $105°C$ are shown in Figure (11) and comparison with theoretical predictions made. Setting aside the friction associated with the ramp regions where the cam and follower meet and part, the overall trend of experimental variation is well predicted by the theoretical model. This is particularly interesting as the instantaneous mode of lubrication will vary through a cycle.

It is also worthy of note that wear modelling in cams and followers has begun to investigate the influence of elastohydrodynamic films. Bell et al (1990) have developed a wear model incorporating a wear coefficient dependent upon the lubricant film thickness, or to be more precise the film thickness ratio (λ) referred to previously. For full film lubrication the value of this wear coefficient was taken as zero, for boundary lubrication at a fixed value appropriate to the circumstances and between these two extremes a linear variation was adopted in the mixed lubrication regime. For the finger follower system investigated boundary lubrication was found to be the predominant wear controlling factor. However, the contribution of hydrodynamic effects to wear control on the flanks was noted.

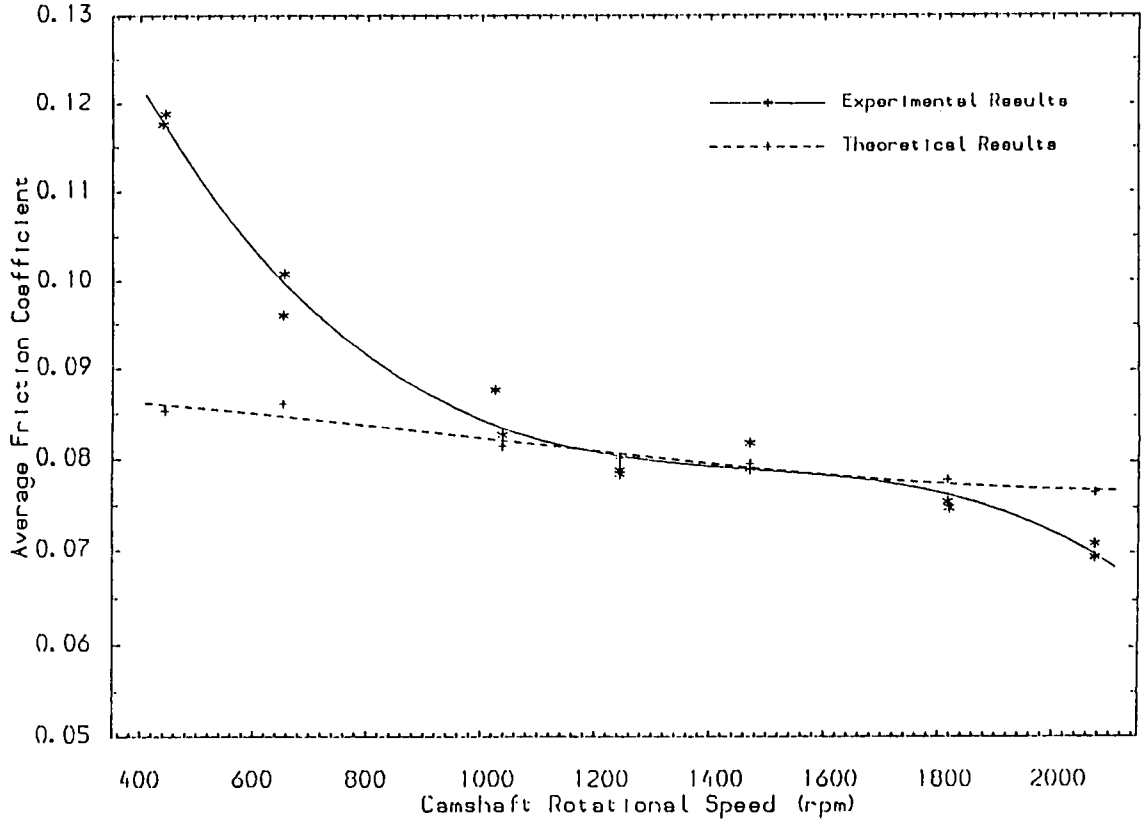
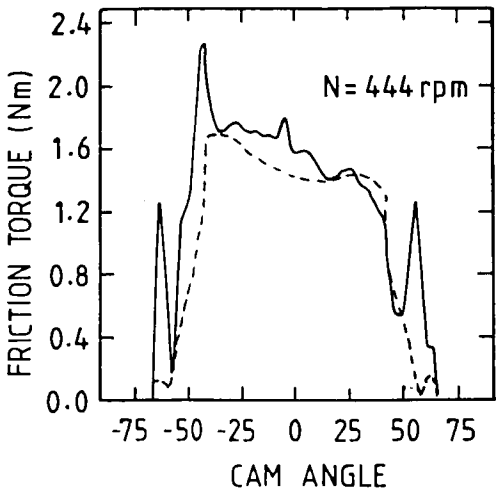


Figure 10 Average friction coefficient as a function of camshaft frequency at a bulk temperature of 105°C



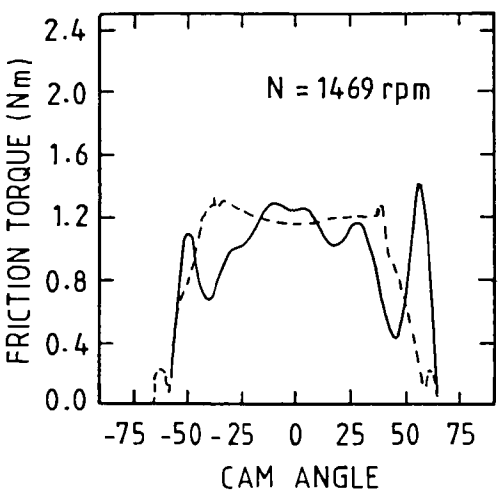


Figure 11 Comparison of theoretical (dotted) and experimental (solid) data for instantaneous camshaft friction torque at rotational frequencies of 444 and 1469 r/min (bulk temperature 105^{o}C)

6. THE FAILURE OF CAMS AND FOLLOWERS

A detailed discussion on the failure mechanism is beyond the scope of this paper. Some basic conclusions, therefore in association with the failure of cams and followers are outlined as follows.

Generally speaking, three main forms occur. These are pitting, polish wear and scuffing. Failure may occur on either the cam or follower, often in differing degree on both. Neale (1973) defines the three forms of wear as follows:

Pitting is the failure of a surface, manifested initially by the breaking-out of small roughly triangular portions of the material surface. This failure is primarily due to high stresses causing fatigue failure to start at a point below the surface where the highest combined stresses occur. Heavily loaded surfaces will continue to pit with increasing severity with time.

Polish wear is the general attrition of the contacting surfaces. When conditions are right this could be small, but occasionally very rapid wear can occur, particularly with chilled and hardened cast iron flat-faced followers. Polish wear appears to be an intermediate case between pitting and scuffing assisted by some form of chemical action involving the lubricant.

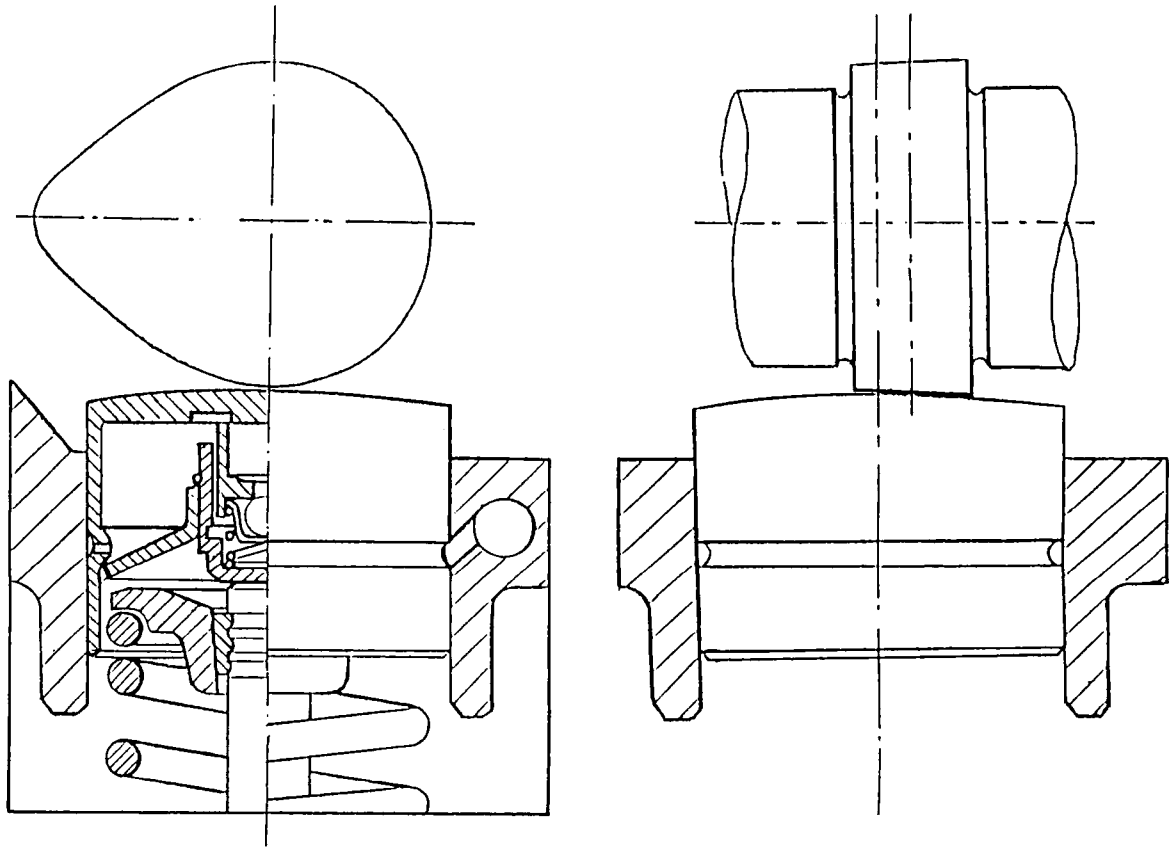
Scuffing is believed to be caused by the local welding of two heavily loaded surfaces, particularly when a high degree of relative sliding occurs under poor lubrication conditions, followed by the tearing apart of the welded material.

Dyson and Naylor (1960-1961) addressed the application of the flash temperature concept to wear problems and this paper was the precursor of many contributions (e.g. Naylor (1967-1968), Colgan and Bell (1990)). They believed that the scuffing and polishing wear are controlled by the temperature in the contact zone.

7. THE FUTURE

Many developments are still required in relation to the understanding of the lubrication of cams and followers in automobile valve trains. At the present time there is a strong trend to the use of roller followers and the cam and bucket follower. Whilst much of the information presented in this paper has related to cam and flat faced followers with nominal line contact, there remain many challenging problems in more practical arrangements. Figure (12) is a sketch of a domed tappet, with an angled offset cam, the rotating follower being incorporated with hydraulic lash adjuster.

An analysis to understand the tribological implications of the arrangement presented in Figure (11) is complex and challenging. Quite apart from the point (or elliptical) contact situation, the requirement to model the lash adjuster and the need for a satisfactory analysis for the follower guide, there are some particularly difficult engineering science problems. Thus because the follower is permitted to rotate, the entrainment of the lubricant into the contact is at a varying angle to the principal axis of the contact ellipse



DOMED TAPPET INCORPORATING HYDRAULIC LASH ADJUSTER.

Figure 12

(Figure (13)). Data is available to determine lubricant film thickness for this situation (Dowson (1985)) with elastohydrodynamic lubrication, however, the precise determination of the entraining velocities will depend upon the calculation of frictions which are not in themselves simple to obtain precisely. In addition, it has been observed that the cam offset, coupled with its taper, can lead to a trucated rather than an elliptical contact (see Figure (14)) and with poor designs edge loading may occur. Such circumstances will not be easy to analyse.

8. CONCLUSIONS

A computer programme to enable the assessment with ease of the influence of the variation of a range of design parameters has been developed.

Parametric studies have been presented for a cam acting directly against a flat faced follower. The results of the parametric study has been summarised.

It has been shown that design approaches for cams and followers, which have traditionally been based upon considerations of materials, lubricant additive package and Hertzian stress values, may be strengthened by parametric studies of performance based on lubrication analysis. The importance of the elastohydrodynamic or mixed lubrication modes in the characterization of the performance of the automobile cam and follower has been identified. Conventional thinking in relation to this has mainly dealt with concepts associated with boundary lubrication.

The strongest evidence of the significance of hydrodynamic effects is presented through consideration of experimental and theoretical considerations of cyclic film thickness variation between the cam and follower. A remarkable correlation has been obtained. In addition examination of frictional torque and wear behaviour lends support to the view the film lubrication may have an important role.

There remain many challenging tribological problems of a lubrication nature in relation to valve trains which must be addressed if efficacy and durability are to be reliably improved.

9. ACKNOWLEDGEMENT

Permission to reproduce material from the following papers published by Elsevier is gratefully acknowledged: Ball et al (1989); Dowson et al (1988) and Taylor (1991).

10. REFERENCES

Ball, A.D., (1988), "A Tribological Study of the Design and Performance of Automotive Cams", Ph.D. Thesis, Department of Mechanical Engineering, University of Leeds.

Ball, A.D., (1989), "A Computer Program for Valve Train Analysis" Internal Report, Department of Mechanical Engineer, University of Leeds.

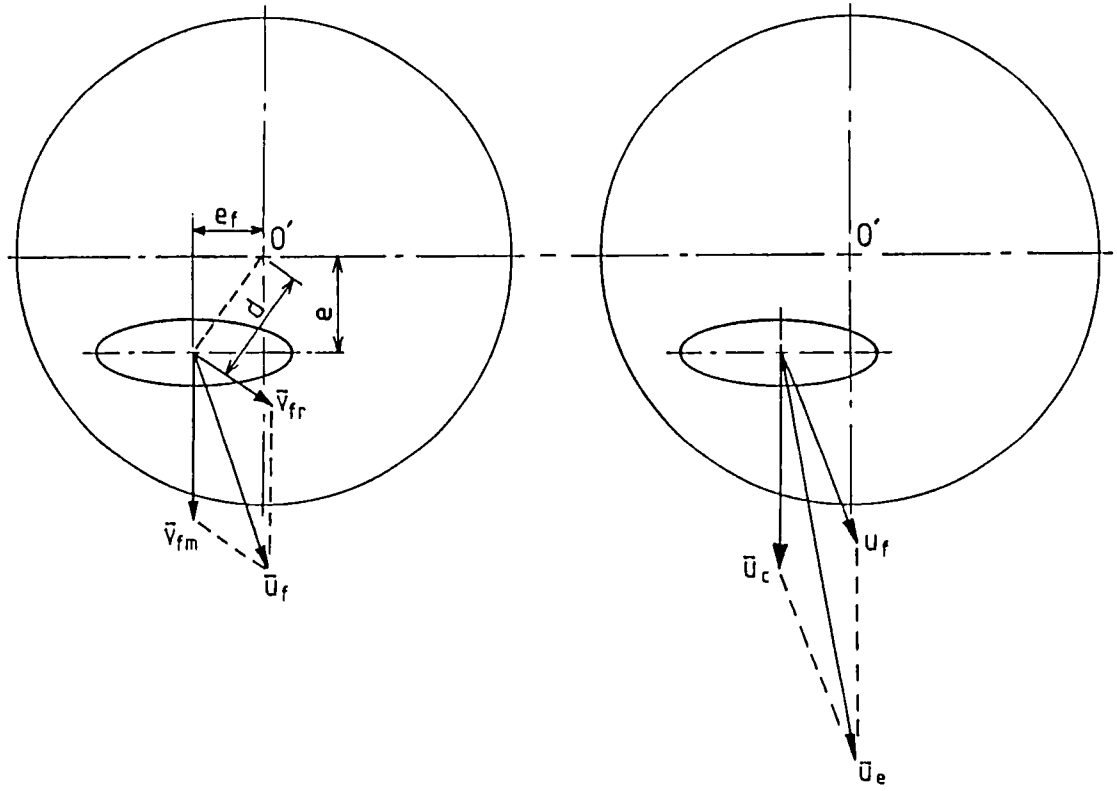


Figure 13 Entraining velocities into contact

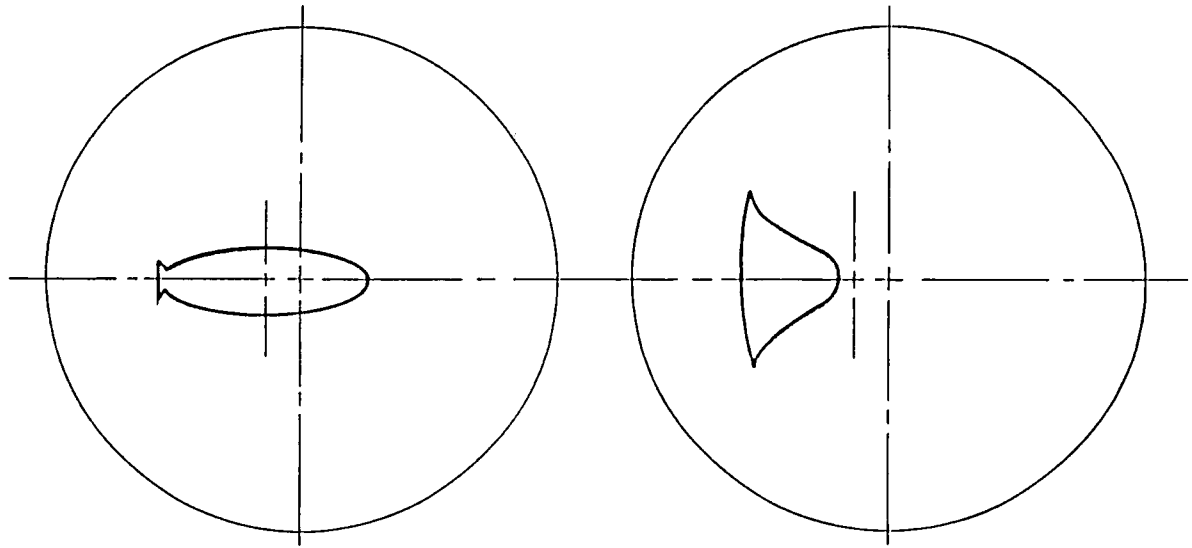


Figure 14 Contact footprint for truncation and edge loading

Ball, A.D., Dowson, D. and Taylor, C.M. (1989), "Cam and Follower Design", 15th Leeds-Lyon Symposium on Tribology, pp 111-130, Butterworths.

Chittenden, R.J., Dowson, D., Dunn, J.F. and Taylor, C.M. (1985), "Elastohydrodynamic Lubrication of Concentrated Contacts (Short Title) Part 1", Proc. Roy. Soc., A.397, pp 245-2609, Ibid, (1985), A397, pp 271-294.

Colgan, T. and Bell, J.C. (1990), "Modelling Valve Train Wear For Cycling Conditions", proc. Int. Trib., Conf., Nagoya, Japan, Vol. II., pp 1231-1236.

Dowson, D., Harrison, P. and Taylor, C.M. (1985), "The Lubrication of Automotive Cams and Followers", Mechanisms and Surface Distress, 12th Leeds-Lyon Symposium on Tribology, pp305-322, Butterworths.

Dowson, D., Harrison, P., Taylor, C.M. and Zhu, G. (1990), "Experimental Observations of Lubricant Film Sate between a Cam and Bucket Follower Using the electrical Resistivity Technique", Proc. Int. Trib. Conf., Nagoya, Japan, Vol. 1, pp 119-124.

Dowson, D. and Higginson, G.R. (1977), "Elastohydrodynamic Lubrication", SI Edition, Pergamon Press.

Dowson, D., Taylor, C.M. and Zhu, G. (1986), "Mixed Lubrication of a Cam and Flat Faced Follower", 13th Leeds-Lyon symposium on Tribology, pp 599-609, Butterworths.

Dowson, D., Taylor, C.M. and Zhu, G. (1989), "An Experimental Study of the Tribology of a Cam and Flat Faced Follower", Proc. I.Mech.E., 2nd Int. Conf. on Combustion Engines - Reduction of Friction and Wear, Paper C375/025, pp 97-108.

Dyson, A. and Naylor, H., (1960), "Application of the Flash Temperature Concept to Cam and Tappet Wear Problems", I.Mech.E., Proc. Auto. Div., No. 8, pp 225-280.

Naylor, H. (1967-1968), "Cam and Friction Drives", Proc. I.Mech.E., Vol. 182, Pt. 3A, pp 237-247.

Patir, N. and Cheng, H.S. (1978), "An Average Flow Model for Determining Effects of Three Dimensional Roughness on Partial Elastohydrodynamic Lubrication", Trans ASME, Jrl. Lub., Tech., Vol. 100, No. 1, pp 12-17.

Taylor, C.M. (1991), "Valve Train Lubrication Analysis", Engine Tribology - Proc., 17th Leeds-Lyon Symposium in Tribology, Elsevier, Amsterdam, pp 119-132.

Williamson, B.P., Galliard, I.R. and Benwell, S. (1989), "Measurement of Oil Film Thickness in the Elasto-Hydrodynamic Contact between a Cam and Bucket Follower in a Motored Cylinder Head. Part 1 : Newtonian Oils. Int. Fuels and Lubricants Meeting and Exposition, Baltimore, Maryland, U.S.A., SAE Preprint 892150, pgs. 11.

Zhu, G. (1988), "A Theoretical and Experimental Study of the Tribology of a Cam and follower", Ph.D. Thesis, Department of Mechanical Engineering, University of Leeds.

This Page Intentionally Left Blank

Engine Tribology / C.M. Taylor (Editor)

Piston Assemblies; Background and Lubrication Analysis

D Dowson

Institute of Tribology, Department of Mechanical Engineering, The University of Leeds, Leeds LS2 9JT, England

Abstract

The history of piston ring development is reviewed briefly. The basic mechanics of the simple crank mechanism and a single piston ring in its groove are then considered. An analysis of the lubrication of a single ring is then developed in relation to the cyclic variation of film thickness, friction and lubricant transport. A standard procedure for the calculation of inter-ring pressures in a multi-ring pack is presented and brief consideration is given to the lubrication of a ring pack consisting of compression and oil control rings. Brief reference is made to the influence of distortion of rings, grooves and cylinder bores upon ring pack performance.

1. INTRODUCTION

The piston and piston rings in modern reciprocating engines account for the majority of the friction and power loss associated with the mechanical system. The exact proportion depends upon the type of engine and the operating conditions, but typical figures lie in the range 40%-60%. The tribological performance of the piston ring pack is thus of supreme importance in terms of engine reliability and efficiency.

The major role of piston ring assemblies is to form a dynamic seal between the working fluid in the combustion chamber and the crankcase. If the sealing action is poor, pressure in the combustion chamber will be reduced during the compression, explosion and expansion stages of the engine cycle, thus leading to a reduction in power output and efficiency and a rapid deterioration of the lubricant in the crankcase. The ring pack must also restrict the passage of lubricant from the crankcase to the combustion chamber and the exhaust, while ensuring adequate control of friction and wear at the ring face - cylinder wall interface. This role directly affects oil consumption in engines and the reliability and life of ring packs. Finally, the piston rings provide a heat path for the effective cooling of the piston. Heat enters the piston from the hot gases in the combustion chamber and is thence distributed by conduction along the connecting rod, by convection from the piston

skirt and by conduction through the piston rings, the oil film and cylinder wall to the engine coolant.

The mechanism of engine bearing lubrication is well understood and design procedures are well advanced and reliable. A full appreciation of piston ring lubrication has emerged only in the past twenty years or so and many aspects of ring pack performance remain to be explored. However, a major development has been the recognition that piston rings operate in the fluid-film regime of lubrication throughout much of each cycle and that hydrodynamic analysis can reveal valuable indications of the overall tribological performance of the piston seal. The analysis can also be extended to take account of mixed lubrication and estimates can be made of the relative contributions of fluid-film and boundary lubrication to power loss predictions.

In this chapter the essential features of piston ring pack lubrication analysis are considered.

1.1 Notation

a	radial width of piston ring
b	axial height of piston ring
f	overhang of piston ring beyond effective pivot point (approximately equal to the distance from the pivot to the cylinder wall)
h	enthalpy of gas
"	oil film thickness
k	exposed length of starved piston ring at top end of ring face
ℓ	exposed length of starved piston ring at bottom end of ring face
"	connecting rod length
m,m-1,...	designation of inter-ring volumes
p	pressure
p_1	pressure above piston ring
p_2	pressure below piston ring
s	piston displacement from T.D.C.
t	time
u	internal energy of gas
"	entraining velocity = $(U_1+U_2)/2$
v	gas velocity
x,z	axial and radial coordinates
$\bar{x},\bar{z}$	Location of piston ring centre of mass
$x_{C.P.}$	axial location of resultant hydrodynamic normal force from mid-plane of ring
C_d	discharge coefficient of orifice
E,F,G	constants
F_h	hydrodynamic shear force/unit circumferential length (radial)
F_p	ring pivot friction force/unit circumferential length (radial)

I	moment of inertia of piston ring
K	torsional stiffness
M	mass of piston ring/unit circumferential length
M	mass flow rate of gas through orifice
P_s	spring (sealing) force/unit circumferential length (radial)
P_z	hydrodynamic normal force/unit circumferential length (radial)
R	gas constant
"	crank radius
R_p	piston/groove reaction force/unit length (axial)
T	temperature
U	piston velocity
V	volume rate of flow of gas through orifice
α	angle of twist of piston ring
"	angle between connecting rod and cylinder axis
γ	ratio of specific heats at constant pressure and constant volume
θ	crank angle (from T.D.C.)
η	lubricant viscosity
Λ	ratio of calculated film thickness/composite surface roughness
Ω	angular velocity of crank

2. HISTORICAL BACKGROUND

The present forms of piston ring assemblies owe much to the development of this form of seal in the era of steam rather than internal combustion engines. The Newcomen atmospheric beam engine of 1712 erected at Dudley Castle for the purpose of draining a coal mine had a thermal efficiency of about 0.5%. The bore of the cast iron cylinder was 19 in. (0.48m) and the stroke 72 in. (1.83m) and one of the major problems was the leakage of steam past the piston. It was not uncommon for the bores of large early 18th century steam engines to vary by about 0.2 in. (5mm), but John Smeaton improved the sealing by means of rope packing to achieve thermal efficiencies of about 1.4% by 1774.

James Watt improved the efficiency of steam engines substantially when he introduced separate condensers in 1769. The firm of Boulton and Watt dominated the development and manufacture of steam engines in the latter half of the 18th century. Cylinders were bored by the new technique introduced by John Wilkinson, whereby the cylinder was clamped securely on a stationary work support while the boring bar was passed through it. This produced a much improved cylinder bore, such that Watt was able to promise his customers that;

> .... *"a 72 in. (1.83m) cylinder would not be further from absolute truth than the thickness of a thin sixpence in the worse part".*

By 1800 the thermal efficiency of 'expansive' steam engines was close to 5%. The common sealing material was rope or hemp, but as steam pressures increased and pistons became heavier, the axial length of packing in an 18in. (0.46m) piston extended to about 2.5in (63.5mm). The piston had to be fabricated to enable the packing material to be compressed and the use of bolts, nuts, cotters and pins provided plenty of scope for loosening and the release of damaging metallic components within the cylinder.

Mr John Ramsbottom (1854) introduced metallic piston rings which revolutionised the form of the piston seal. He machined three grooves 0.25 in. (6.35 mm) wide, 0.25 in. apart and 3/16 in. (4.76 mm) deep into the circumference of a cast iron piston to receive rings of brass, iron or steel. The rings were drawn to a suitable section to fit the grooves and rolled into a circular form with a diameter about 10% greater than the cylinder bore. The simple principle of operation was that the elasticity of the ring would provide an effective seal when the piston was inserted into the cylinder. The piston assembly had the merit of lightness, simplicity, structural security, reduced friction and extended life. A set of three rings, which cost 2s 6d (12.5p), ran for 3,000-4,000 miles and returned a 12% improvement in fuel consumption (5.7 lb. of coal per mile recorded over an aggregate distance of 269,800 miles). In an interesting response to a question by Mr Archibald Slate, President of the Institution of Mechanical Engineers, Ramsbottom stated that he had intended to rivet a stud into each groove to prevent ring rotation, but this had not been found to be necessary, since the leakage of

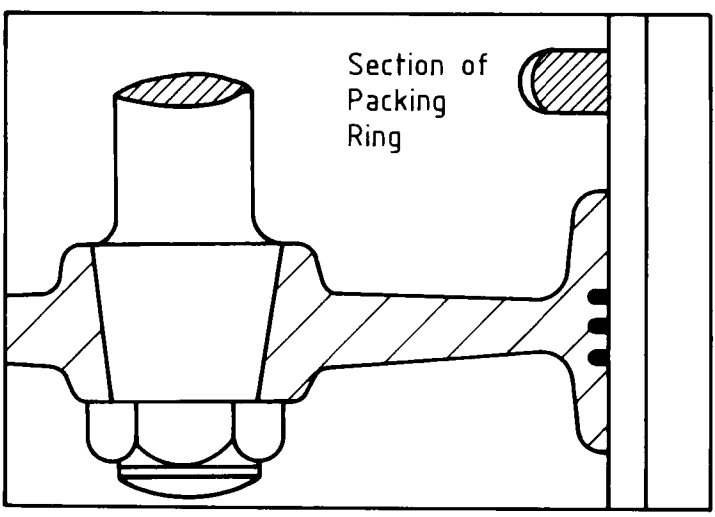


Figure 1 The Ramsbottom (1854) Steam Engine Piston Assembly

steam was not significant and the shifts in position yielded much improved surface finishes. An illustration of the Ramsbottom piston assembly is shown in Figure 1.

Some of the Ramsbottom rings broke under the action of the steady, elastic, bending stresses and the superimposed dynamic stresses associated with ring movement and this prompted Mr David Joy (1855) to introduce a spiral ring into the grooves. However, it was Miller (1862) who first used steam pressure, rather than the inherent elasticity of the ring, to provide the closing force between each ring and the cylinder wall. He drilled small holes from the piston crown to the back of the grooves, thus providing peak closing forces when the sealing action was most needed. This concept has been retained to the present day, but the gas pressure is now exerted on the rear of the rings by leakage into the grooves from the piston-liner clearance spaces, rather than through holes in the piston.

3. KINEMATICS OF PISTON-CONNECTING ROD-CRANK MECHANISM

An illustration of the crank mechanism is shown in Figure 2.

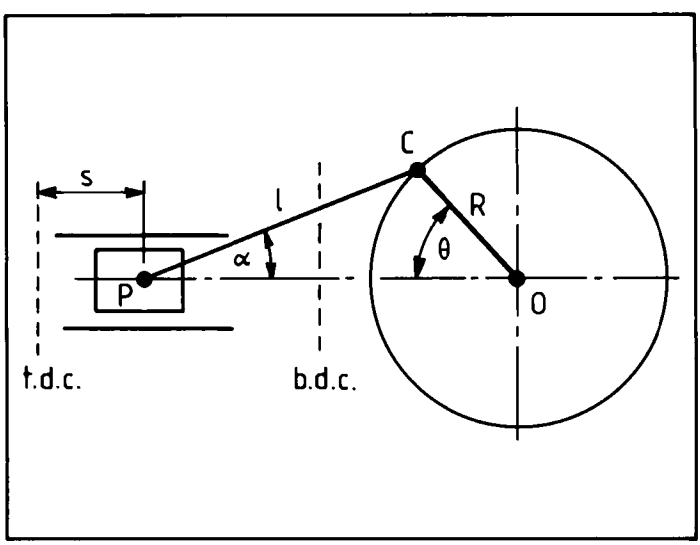


Figure 2 Crank Mechanism

In an analysis of the lubrication of piston rings it is necessary to determine the velocity of the piston (ring) as a function of crank angle. Furthermore, the ring

acceleration influences the phenomenon of ring lift. Piston displacement from top dead centre (t.d.c.) is

$$s = R(1-\cos\theta) + \ell\ (1-\cos\alpha) \tag{1}$$

$$\sin\alpha = \left(\frac{R}{\ell}\right)\sin\theta$$

Now,

$$\cos\alpha = \sqrt{1-\sin^2\alpha} = \sqrt{1-\left(\frac{R}{\ell}\right)^2 \sin^2\theta} \approx 1-\left(\frac{R}{\ell}\right)^2 \sin^2\theta.....$$

Hence,

$$s = R\left[1-\cos\theta+\frac{1}{2}\left(\frac{R}{\ell}\right)\sin^2\theta+\right] \tag{2}$$

Piston Velocity $\quad U = \frac{ds}{dt} = \frac{ds}{d\theta}\cdot\frac{d\theta}{dt} = \Omega\frac{ds}{d\theta}$

or,

$$U = R\Omega\left[\sin\theta+\frac{1}{2}\left(\frac{R}{\ell}\right)\sin2\theta+.....\right] \tag{3}$$

Piston Acceleration $\quad f = \frac{dU}{dt} = \frac{d^2s}{dt^2} = \frac{dU}{d\theta}\cdot\frac{d\theta}{dt} = \Omega\frac{dU}{d\theta}$

or,

$$f = R\Omega^2\left[\cos\theta+\left(\frac{R}{\ell}\right)\cos2\theta+.....\right] \tag{4}$$

4. PISTON RING GEOMETRY AND MECHANICS

4.1 Ring Face Geometry

Piston rings generally have a cross section which is approximately rectangular in form. The ring face adjacent to the cylinder liner may initially be flat or profiled, but the ring twists during each cycle and wear causes a curved profile to form on the face. The hydrodynamic performance of bearings is influenced more by the

change in film thickness along the bearing length than by the form of the change and hence a good estimate of piston ring lubrication performance can be achieved by representing the ring face by a curved profile of parabolic form. The influence of initial, machined, profiles or the effects of wear can be assessed by allowing the parabolic profile to be symmetrical about the mid-plane or offset, as shown in Figure 3.

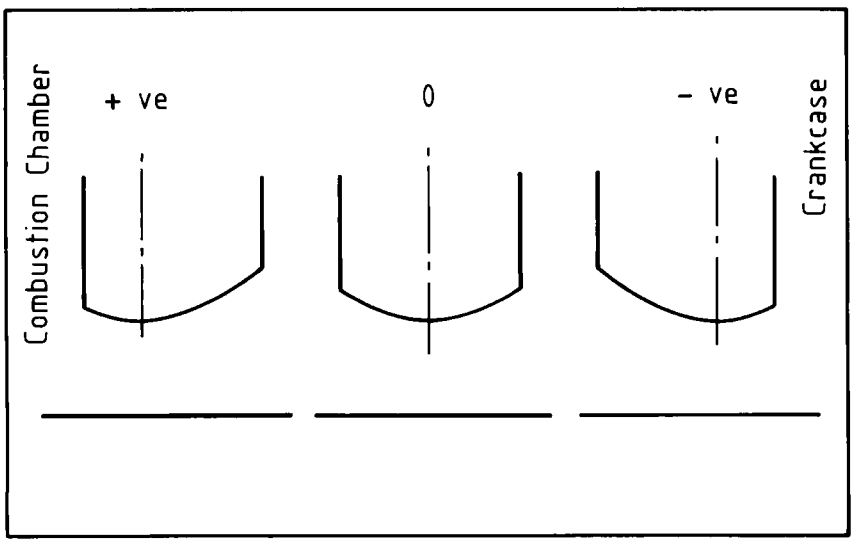


Figure 3 Parabolic Ring Face Profiles and Offsett

For the downward scraping second rings, which are used increasingly to control oil consumption and emissions, it is more realistic to represent the ring face as a plane inclined surface.

4.2 Forces and Moments

A cross section of a representative piston ring is shown in Figure 4.

The gas pressures above, below (axial) and behind (radial) the ring exert normal forces on the ring section, while both normal (radial) and tangential (axial) forces are applied to the ring face from the lubricating film or the contact between the ring and the liner. If parts of the ring face are exposed to the gas pressures above and below the ring, the associated normal (radial) forces relieve the sealing force. The loading associated with the elastic spring force on rings is represented as an equivalent pressure acting on the rear of the ring. It is typically 10^4 Pa-10^5 Pa, but this represents only about one percent of the peak gas pressure.

The ring nominally sits on the piston groove flanks, but ring twist, flank wear and piston distortions due to thermal and elastic effects in a firing engine will cause the ring to twist about an effective pivot line. It is difficult to specify the locations of such pivot points, or lines, unless full analyses of distortions are carried out, but

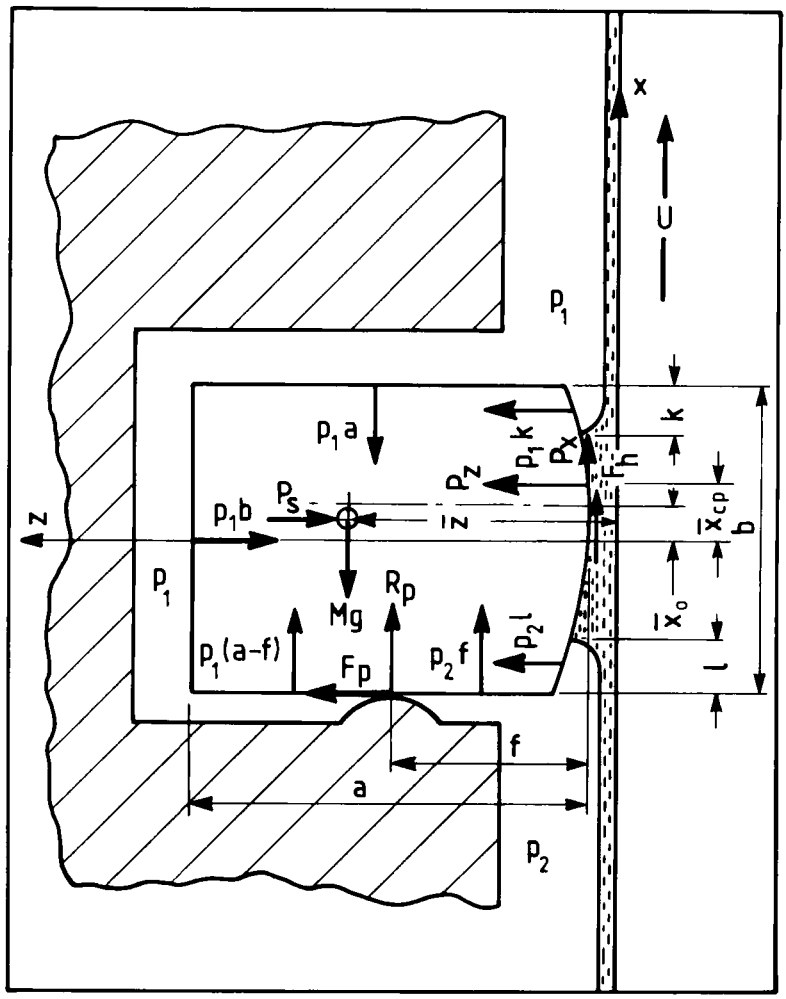


Figure 4 Forces Acting on a Piston Ring

if they can be located there will be normal (axial) and tangential (radial) components of the pivot reaction force acting on the ring. Finally, the inertia of the

ring and the axial accelerations (equation (4)) have to be considered when the seating of the ring on the ring groove flank is considered. It is also necessary to determine the centre of area (mass) for ring twist calculations.

The radial and axial force balance equations and the requirement for equilibrium of moments about the centre of mass of the ring section can be written as;

Radial

$$p_1 (b-k) - p_2(\ell) - F_p - P_z + P_s = M \frac{d^2 z}{dt^2} \approx 0 \tag{5}$$

Axial

$$R_p + F_h + P_x - (p_1 - p_2) f - Mg = M \frac{d^2 x}{dt^2} \tag{6}$$

Moments

$$(p_1 - p_2)(f)\left(\bar{z} - \frac{f}{2}\right) + p_2(\ell)\left(\frac{2\bar{x} + b - \ell}{2}\right) - p_1(k)\left(\frac{-2\bar{x} + b - k}{2}\right)$$

$$+ F_p\left(\frac{2\bar{x} + b}{2}\right) - F_h(\bar{z}) - P_z\left(\bar{x}_{C.P.} - \bar{x}\right) - P_x(\bar{z}) - R_p(\bar{z} - f) - K\alpha = I \frac{d^2 \alpha}{dt^2} \tag{7}$$

If the sum of the external axial forces shown on the left hand side of equation (6) is smaller than the inertia force on the right hand side calculated by introducing the piston acceleration from equation (4), the ring can be assumed to lift (flutter) from the flank of the groove.

5. LUBRICATION ANALYSIS OF A SINGLE RING

5.1 Mode of Lubrication

It should be recalled that an understanding of the mechanism of lubrication of piston rings has emerged only during the past two decades or so. One procedure adopted to determine the mode of lubrication is to assume that the bearing surfaces are perfectly smooth, to calculate the cyclic variation of film thickness and to relate the findings to the composite surface roughness of the ring face and cylinder liner. The ratio of calculated fluid-film thickness to composite roughness is known as the Lambda (Λ) ratio.

$$\Lambda = \frac{\text{calculated fluid film thickness}}{\text{composite surface roughness}} \tag{8}$$

If the Lambda ratio is greater than about three, effective fluid-film lubrication can be expected, whereas if it is less than unity conditions are akin to boundary lubrication. In the intermediate zone ($1 < \Lambda < 3$) a mixed lubrication regime will apply.

5.2 Gas Pressures

It has already been noted in Section 4.2 that the gas pressure behind a ring provides the major contribution to the sealing force. In the case of the top compression ring the gas in the combustion chamber passes down the clearance space between the piston crown land and the cylinder liner and thence into the top ring groove to load the rear face of the ring. There will be some attenuation of this pressure compared with the combustion chamber pressure, but it is usual in lubrication analyses of piston ring packs to assume that the pressure in the top ring groove is at all times equal to the combustion chamber pressure.

Flow of gas into the volume formed between the piston land between the first and second rings is, of course, strictly limited by the top ring. If it is assumed that lubricant occupies any clearance between the ring and the cylinder liner and that no gas passes between the ring and the flank of the ring groove, gas leakage will be restricted to an orifice formed by the ring gap and the radial clearance between the ring and the liner. The similarity between the mechanism of gas flow through a ring pack and that through a labyrinth seal prompted Eweis (1935) to propose an **'orifice and volume'** method for the calculation of inter-ring pressures. A more recent and complete account of this approach has been presented by Ting and Mayer (1974).

The general model for flow through these orifices between the inter-ring volumes is shown in Figure 5.

Adiabatic flow is assumed and an equation for the mass rate of flow through each orifice (e.g. between the volumes (m-1) and (m)) can be obtained directly from the energy equation for a compressible fluid. If potential energy changes are deemed to be negligible,

$$\frac{P_{m-1}}{\rho_{m-1}} + \frac{v^2_{m-1}}{2} + u_{m-1} = \text{constant} = \frac{p_m}{\rho_m} + \frac{v^2_m}{2} + u_m$$

Furthermore, $\frac{p}{\rho} + u = \text{enthalpy (h)} = C_p T$,

and if v_m represents the mean velocity of the jet of gas entering volume (m) through the orifice and the kinetic energy of the fluid in the reservoir (m-1) is deemed to be negligible,

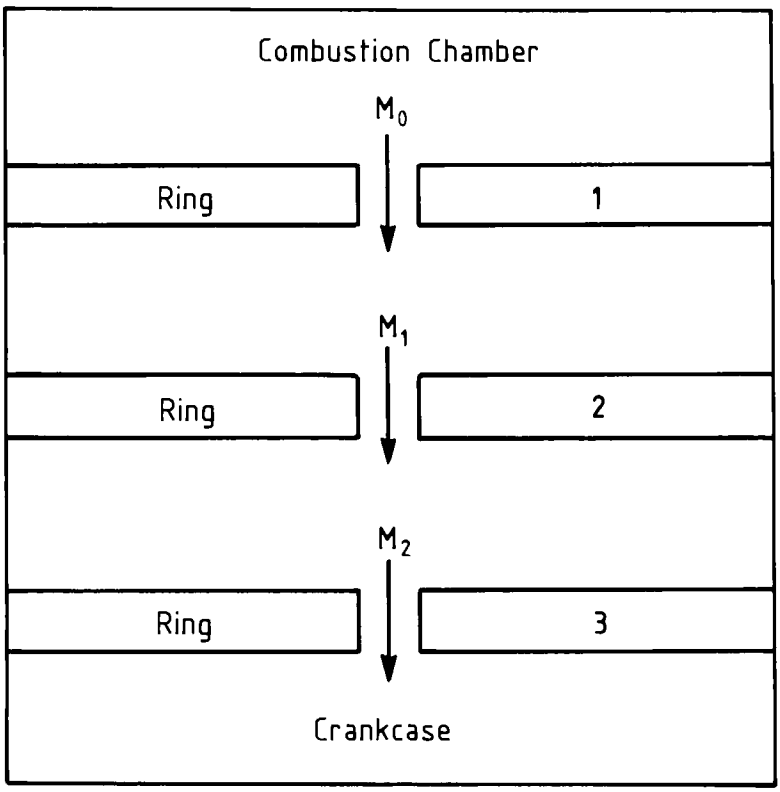


Figure 5 Model for Gas Flow Between Inter-Ring Volumes

$$v_m = \left[\left(\frac{2\gamma}{\gamma-1}\right) R\, T_{m-1} \left(1 - \frac{T_m}{T_{m-1}}\right)\right]^{\frac{1}{2}}$$

For an adiabatic process,

$$\frac{T_m}{T_{m-1}} = \left(\frac{P_m}{P_{m-1}}\right)^{\frac{\gamma-1}{\gamma}}$$

and hence,

$$v_m = \left[\left(\frac{2\gamma}{\gamma-1}\right) R\, T_{m-1} \left\{1 - \left(\frac{p_m}{p_{m-1}}\right)^{\frac{\gamma-1}{\gamma}}\right\}\right]^{\frac{1}{2}}$$

The mass rate of flow (dM/dt) through an orifice of cross sectional area (A) having a discharge coefficient (C_d) is given by;

$$\frac{dM}{dt} = C_d \rho A v = \Omega \frac{dM}{d\theta}$$

and hence the mass rate of flow from the reservoir (m-1) into the volume (m) due to the pressure difference $[(p_{m-1}) - (p_m)]$ is given by;

$$\left(\frac{dM_{m-1}}{d\theta}\right) = \frac{C_d A_m}{\Omega}\left[\frac{2\gamma}{R(\gamma-1)\,T_{m-1}}\right]^{\frac{1}{2}}(p_{m-1})\left(\frac{P_m}{P_{m-1}}\right)^{\frac{1}{\gamma}}\left[1-\left(\frac{P_m}{p_{m-1}}\right)^{\frac{\gamma-1}{\gamma}}\right]^{\frac{1}{2}} \qquad (9)$$

If (p_m) exceeds (p_{m-1}) reverse flow takes place and the subscripts (m) and (m-1) should be reversed in equation (9) in order to determine the mass rate of flow out of the volume (m).

It is, of course, the nett rate of change of mass associated with the flow through two orifices that determines the mass changes within a ring pack. If the initial mass in volume (m) is (M_m) and the nett mass gain from the two orifices is $[(M_m)_{in} - (M_m)_{out}]$, the inter-ring pressure at a new crank angle is given simply by the perfect gas equation as;

$$p_m = \frac{RT_m}{V_m}\left[M_m + \left\{(M_m)_{in} - (M_m)_{out}\right\}\right] \qquad (10)$$

The equations can be solved iteratively until changes in pressures in the inter-ring volumes fall below some specified tolerance at each crank angle.

The flow through the orifice increases as the pressure ratio (p_m/p_{m-1}) increases, but when the velocity becomes sonic, the mass flow rate reaches a maximum as the nozzle chokes. This condition can be identified by differentiating equation (9) with respect to (p_m/p_{m-1}) and equating the result to zero, to yield;

$$\frac{p_m}{p_{m-1}} = \left(\frac{2}{\gamma+1}\right)^{\frac{\gamma}{\gamma-1}} \qquad (11)$$

A representative value of (γ) for internal combustion engine gases is (1.3) and hence, for choked conditions, (p_m/p_{m-1}) = 0.546 and equation (9) reduces to

$$\frac{dM_{m-1}}{d\theta} = 0.227\,\frac{C_d A_m}{\Omega}\left[\frac{2\gamma}{R(\gamma-1)\,T_{m-1}}\right]^{\frac{1}{2}}(p_{m-1}) \qquad (12)$$

About a decade ago Ruddy et al (1981(b)) enhanced the analytical approach to inter-ring pressure calculation to take account of the cyclic variation of temperature in the combustion chamber, the volume of gas within the ring groove, the variation in ring gap in a running engine due to thermal and elastic distortion and wear, and the lack of axial alignment of the ring gaps in a ring pack.

It is unusual to have experimental values of inter-ring pressures available in studies of piston ring lubrication and an essential first step is to obtain the cyclic variation of combustion chamber pressure from a (P-V) diagram and then to calculate the cyclic variation of all the inter-ring pressures by the above procedures. It is usual to assume that the crank case is held at atmospheric pressure in such analyses. A comparison between the theoretical predictions of inter-ring pressures and the measured values in the Abingdon B-1 single cylinder, two stroke test engine is shown in Figure 6.

There will also be some leakage of gas through the free volumes between the contacting ring and the groove flank, since both surfaces are rough and the axial load pushing the solids together is carried by a large number of discrete contact points. Such problems have been analysed extensively in relation to gas seals and relationships have been established for the flow rates at given pressure drops as functions of the surface topography and the applied load.

The parallel leakage path can readily be included in the analysis of inter-ring pressures in piston ring packs, but there is generally less certainty about the metal-to-metal contact conditions than there is about the orifice flow conditions between the ends of the ring. For this reason the orifice flow model is often used to provide

an initial estimate of inter-ring pressures in piston ring packs. It should be remembered, however, that although the height of the effective slot between the two contacting surfaces is small (typically 10^{-6}m), the circumferential width of the slot is large (typically 10^{-1}m), giving effective areas of slot cross section of 10^{-7} m^2 - 10^{-6} m^2. Orifice sizes in ring gaps vary enormously from one engine size to another and they increase with time due to wear. However, they are generally of order 10^{-6} m^2. The ring gap thus usually presents the larger leakage cross-sectional area, even though the effective length of the slot between contacting rings and groove flanks is quite large. It is therefore a good approximation to consider only the leakage through the ring gaps.

5.3 The Reynolds Equation

The conjunction between a piston ring and a cylinder liner represents a supreme example of a dynamically loaded bearing. The load varies rapidly and substantially in each cycle (say 30-50 times per second), the entraining velocity varies in a roughly sinusoidal manner (see equation (3)) and reduces to zero at the top and bottom dead centre positions, while the temperature and hence the

viscosity of the lubricant changes along the length of the stroke. Furthermore, if ring twist and the effects of thermal and elastic distortions are considered, the effective shape of the bearing changes throughout each cycle. On a longer time

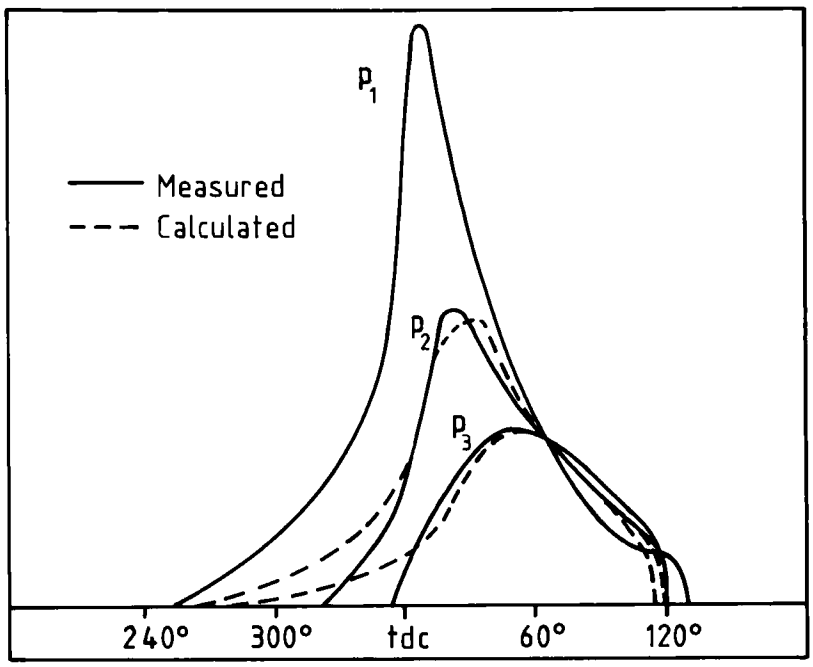


Figure 6 Measured and Calculated Inter-Ring Gas Pressures in Single Cylinder, 2-Stroke Test Engine (Abingdon B-1)

scale the progressive wear of both the piston rings and the cylinder liner will affect the bearing configuration.

The appropriate form of the Reynolds equation for an incompressible lubricant, in which circumferential flow of the lubricant is neglected, is;

$$\frac{\partial}{\partial x}\left[\frac{h^3}{\eta}\left(\frac{\partial p}{\partial x}\right)\right] = 12u\left[\frac{\partial h}{\partial x}\right] + 12\,\frac{\partial h}{\partial t} \qquad (13)$$

- where the first term on the right hand side represents entraining action ($u = (U_1+U_2)/2$) and the second term the influence of normal velocities (squeeze-film

action). If (θ) represents crank angle and the angular velocity of the engine is (Ω), the latter term can be written more conveniently as $12\Omega(\partial h/\partial\theta)$.

Since (h) is a function of (x) alone at a given instant, equation (13) can be integrated to yield,

$$\frac{\partial p}{\partial x} = 12\eta u\left[\frac{1}{h^2} + \frac{\Omega}{u}\left(\frac{\partial h}{\partial\theta}\right)\frac{x}{h^3} + \frac{C_1}{h^3}\right] \tag{14}$$

Further integration reveals the following general expression for pressure distribution within the lubricating film.

$$p = 12\eta u\left[E + \frac{1}{R}\left(\frac{R\Omega}{u}\right)\frac{\partial h}{\partial\theta}\cdot F + (C_1)\,G + (C_2)\right] \tag{15}$$

- where $E = \int\frac{dx}{h^2}$; $F = \int\frac{x}{h^2}\,dx$; and $G = \int\frac{dx}{h^3}$

For rings of parabolic profile the integrals E, F and G have analytical solutions and hence the form of the pressure distribution can readily be obtained for a given entraining velocity (u) and squeeze-to-entraining velocity ratio. It will be noted that the pressure distribution expression involves two constants of integration and the full determination of pressure thus calls for two boundary conditions, as discussed in the next section.

Once the pressure distribution has been determined the load supporting capacity for a given minimum film thickness is obtained by numerical integration. If the film thickness at a given location has been estimated, the initial evaluation of load supporting capacity will invariably fail to match the external sealing force applied to the ring. It is therefore necessary to adjust the minimum film thickness successively until a satisfactory force balance is achieved.

As a solution for film thickness emerges at a given crank angle ($\theta+\delta\theta$), the current value of ($\partial h/\partial\theta$) can be re-assessed by means of numerical differentiation of current and previous values of the film thickness and constantly updated in the iterative procedure outlined for film thickness determination. It is convenient to commence the cyclic film thickness calculation at a location where the squeeze-film velocity is likely to be small e.g. at mid-stroke. The process converges rapidly and it is usual to achieve convergence during the second cycle of crank rotation.

5.4 Boundary Conditions

A representative form of the hydrodynamic pressure distribution in the partially filled clearance space between a piston ring face and cylinder liner is shown in

Figure 7. The gas pressures above and below the ring, determined according to the procedure outlined in the previous section, are (p_1) and (p_2) and the oil film pressure must therefore be equal to these values at locations (x_1) and (x_4) respectively.

The main region for the generation of hydrodynamic pressures which yield a fluid-film force capable of balancing the instantaneous sealing force is located between (x_1) and (x_2). At (x_2) the pressure falls to the saturation pressure for the

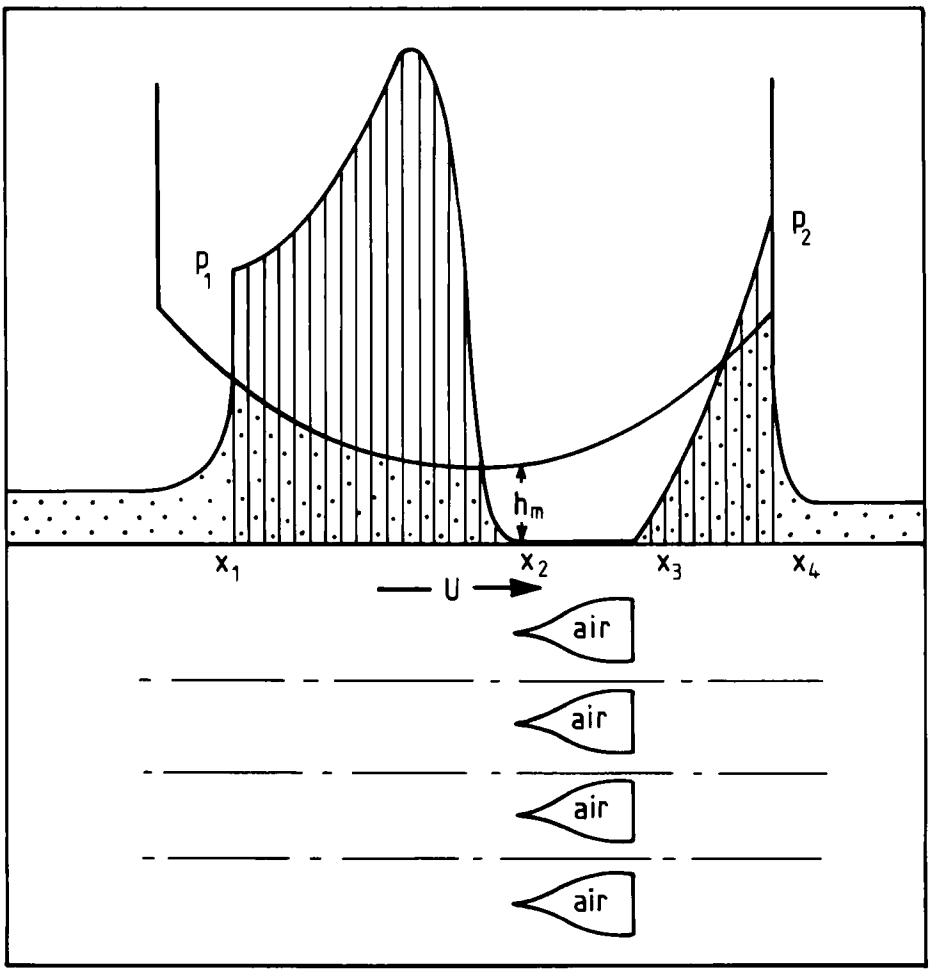


Figure 7 Hydrodynamic Pressure Distribution in a Starved Piston Ring/ Cylinder Liner Conjunction

lubricant, normally atmospheric, such that gas is released from the liquid to create a ruptured or cavitated region between (x_2) and (x_3). Continuity of flow considerations in the full film and ruptured film regions on either side of (x_2) call for the imposition of the Reynolds cavitation boundary condition ($p = dp/dx = 0$) at (x_2). The location of (x_2) varies cyclically under the dynamic conditions encountered in piston ring lubrication, but if the rate of movement is not too great, the Reynolds boundary condition is still valid, as noted by Olsson (1965). The two integration constants (C_1) and (C_2) in the integrated Reynolds equation are determined by these boundary conditions at (x_1) and (x_2). The double condition represented by the Reynolds boundary condition at (x_2), on the pressure and its first derivative, also locates (x_2). It should be noted that the film thickness (h_m) at (x_2) is equal to the film thickness at the line of maximum pressure, where ($\partial p/\partial x$) is also zero. Furthermore, (h_m) is related to the volume rate of flow of lubricant per unit circumferential width by the expression ($Q_x = Uh_m/2$).

There is little experimental evidence to indicate the nature of the cavitated region between piston rings and cylinder liners, but until such time that this position is clarified, a general solution can be obtained in which the pressure (p_2) in the cavitated region is not necessarily equal to the downstream pressure in the inter-ring space (p_4).

If it is ascertained experimentally that the cavitation pressure (p_2) is equal to the downstream pressure (p_4), the appropriate solution can readily be found by assuming again that the Reynolds boundary condition applies at (x_2), while the pressure in the cavitated zone is equated to (p_2).

If the downstream pressure (p_2) exceeds the saturation pressure (atmosphere), a full lubricating film will reform to fill the clearance space at (x_3) to enable the hydrodynamic pressure to build up to (p_2) at the exit location (x_4). The pressure at (x_4) is thus equated to (p_2) and the Reynolds equation, with the previously determined value of (C_1), is integrated until the pressure falls to atmospheric at (x_3). The locations of (x_1, x_2, x_3 and x_4) are thus determined and the magnitude of the hydrodynamic pressures established in the regions (x_1 to x_2) and (x_3 to x_4). In the cavitated region (x_2 to x_3) the pressure is constant and equal to the saturation pressure, but the clearance space is only partially filled with lubricant. There is little direct evidence of the physical disposition of the gas bubbles and the liquid lubricant in the cavitated zone in piston ring lubrication, but it seems reasonable to assume that thin fingers of lubricant extending across the full height of the clearance between the ring face and cylinder liner will be separated in the circumferential direction by discrete air bubbles, as observed in journal bearings and other convergent-divergent bearing configurations. This is important when viscous friction and power loss are evaluated in the cavitated region. It can readily be shown from flow continuity considerations that the proportion of the circumferential width filled with oil at any location ($x_2 < x < x_3$), where the clearance height is (h), is simply (h_m/h),

The full set of boundary conditions on pressure for the solution of the Reynolds equation in piston ring lubrication, can thus be stated in relation to Figure 7 as shown below. It is convenient to deal with gauge pressures and thus to record the atmospheric (cavitation) pressure as zero.

Boundary Conditions

$p = p_1$	at $x = x_1$	Inlet
$p = 0 \left(\text{or } p_2\right)$ and $\frac{\partial p}{\partial x} = 0$	at $x = x_2$	Cavitation
$p = 0 \left(\text{or } p_2\right)$	at $x = x_3$	Reformation
$p = p_2$	at $x = x_4$	Outlet

5.5 Lubricant Viscosity

It is customary to assume that the lubricant viscosity is constant throughout the film at any crank angle. The temperature of the ring face may not be the same as that of the liner at a particular crank angle, but for the purpose of initial analysis it is adequate to identify the viscosity of the lubricant with the surface temperature of the cylinder liner.

The liner temperature varies along the length of the cylinder and it is therefore necessary to determine experimentally, or by analysis, the axial distribution of liner temperature in a firing engine. In the absence of firm data it is usually possible to make realistic estimates of liner temperature distributions based upon experience with other engines. In any event it is necessary to determine the relationship between the effective lubricant viscosity and crank angle.

$$\eta = f(\theta) \tag{16}$$

5.6 Numerical Analysis and Computation

5.6.1. Cyclic Variation of Film Thickness. The basic requirement is to determine the cyclic variation of film thickness, since other overall derived quantities such as friction and power loss and the nett transport of lubricant between the ring and cylinder liner can then be derived with ease. The entraining velocity (see equations (3) and (13)) and the inter-ring pressure equations (9,10) are solved to yield (u), (p_1), (p_2) at each crank angle.

An initial estimate is made of the minimum film thickness at some specified crank angle. It does not matter where the calculation is started in the engine cycle, but convergence is hastened if a location is selected where the variation of film thickness with crank angle is small (e.g. at mid-stroke). The integrated Reynolds equation (15) is then solved according to the boundary conditions recorded in Section 5.4 for a local mean viscosity (η) determined by the liner temperature at the selected crank angle. The normal or squeeze-film velocity required to generate hydrodynamic pressures capable of balancing the sealing force is then established for the assumed, initial value of minimum film thickness at crank angle (θ). This initial squeeze-film velocity enables an estimate to be made of the film thickness at the next increment ($\theta+\delta\theta$) of crank angle. The integrated Reynolds equation (15) is then solved at ($\theta+\delta\theta$) until the force balance equation (5) is once again satisfied and a new estimate of the squeeze film velocity established. The trapezium rule is then applied to re-calculate the minimum film thickness at ($\theta+\delta\theta$) based upon the film thickness at (θ) and the squeeze-film velocities at (θ) and ($\theta+\delta\theta$);

$$h_{(\theta+\delta\theta)} = h_{\theta} + 0.5\left[\left(\frac{\partial h}{\partial \theta}\right)_{\theta} + \left(\frac{\partial h}{\partial \theta}\right)_{\theta+\delta\theta}\right]\delta\theta \qquad (17)$$

This cycle of calculations is repeated as necessary until the film thickness changes by a quantity smaller than some specified tolerance.

The calculation then advances to the next crank angle and in this way it is possible to march out the complete engine cycle to yield the cyclic variation in film thickness. The calculation procedure converges rapidly and it is rarely necessary to pursue the calculation beyond 1½ to 2 complete cycles.

5.6.2. Friction, Power Loss and Lubricant Transport. At each crank angle the viscous shearing force can be calculated. Integration of these quantities enables the power loss in each cycle and the nett lubricant transport to be established. It should be noted, however, that the predicted film thickness generally falls below the composite surface roughness of the ring and liner near one or both of the dead centre positions. This implies that boundary lubrication conditions then apply and in these circumstances a coefficient of friction (say 0.08) is assumed for the purpose of friction and power loss calculation.

5.7 Typical Results

An indication of the predicted and measured inter-ring pressures in a two-stroke diesel engine has already been shown in Figure 6. An important feature of this figure is the relative magnitudes of the inter-ring pressures. It is also interesting to note that, under some circumstances, the pressure beneath a ring can exceed that above it; a result of some importance in relation to emissions.

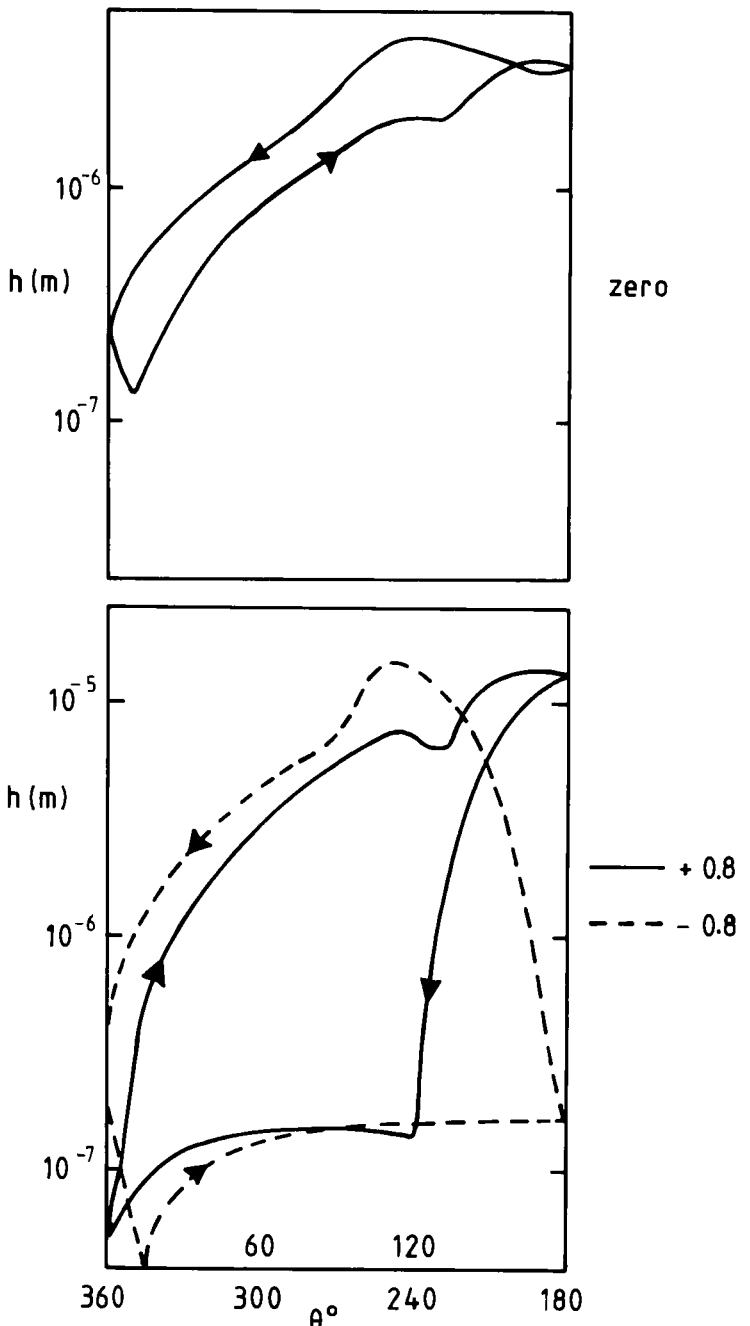


Figure 8 Typical Piston Ring Film Thickness Profiles - Fully Flooded - Effect of Offset

The most basic result of piston-ring lubrication analysis is the prediction of cyclic variation of film thickness. As typical set of results for a 2-stroke diesel engine with a fully flooded ring exhibiting symmetry or considerable positive or negative offset is shown in Figure 8.

A number of important features are evident in Figure 8.

- For a symmetrical ring face profile the film thicknesses on the relatively lightly loaded upstroke exceed those on the downstroke.
- Offset has a pronounced effect upon film thickness predictions.
- The minimum film thickness generally occurs after T.D.C.
- The ring experiences fluid-film lubrication throughout much of the cycle.

The minimum film thickness experienced during each cycle is also influenced by the ring face radius of curvature and the axial height of the ring. The former tends to increase as the radius increases for realistic values, but clearly reaches a maximum and then decays to zero as the radius approaches infinity, while the latter can be optimised for a given radius of curvature.

The viscous shear force developed between smooth rings and liners can be determined at each crank angle, once the film thickness is known. This enables the viscous power loss to be determined, but such predictions tend to underestimate the losses experienced in real engines. A major reason for this discrepancy is that although a fully flooded single ring may enjoy fluid-film lubrication throughout much of the cycle, the effective film thickness is small compared with the composite roughness of ring and liner near to T.D.C. The influence of the lambda Λ factor in determining the mode of lubrication then becomes important as outlined in Section 5.1. If the predicted film thickness falls below the composite roughness, transition to boundary lubrication can be assumed and a constant coefficient of friction (≈ 0.08) adopted. More sophisticated analyses allow the friction coefficient to vary from the fluid film to the boundary lubrication values as a function of (Λ) in the range ($1 < \Lambda < 3$). A typical friction force trace is shown in Figure 9.

Lubricant transport beneath the ring during each full cycle can also be computed once the cyclic variation of film thickness has been ascertained. The nett flow can be ascertained by subtracting the volumes involved in successive up and down strokes and the sensitivity of the calculation to offset is shown in Figure 10. The concept of the "zero nett leakage" seal is apparent for a ring with a small positive offset.

The lubricant transported within the oil film beneath each ring is but one contribution to the total flow, since oil may pass through the ring gap. If the ring lifts

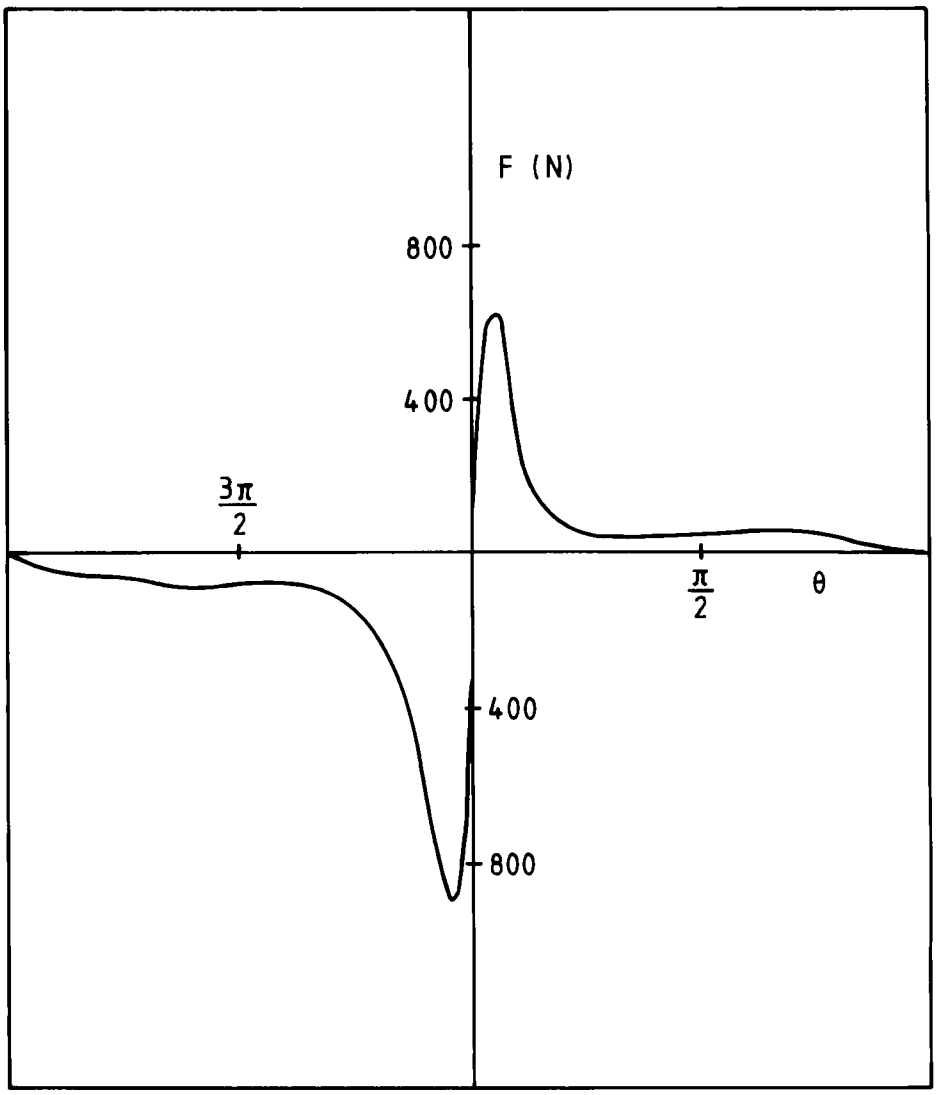


Figure 9 Typical Friction Force Trace for a Single Piston Ring
(Fully flooded. $\mu = 0.1$ when $h < 0.25$ μm

from its seat on the groove flank a relatively large volume of oil may also pass the ring from the inter-ring volumes. A simple squeeze-film calculation can be used to estimate the cyclic surge in lubricant flow towards the combustion chamber associated with the phenomenon of ring lift or flutter.

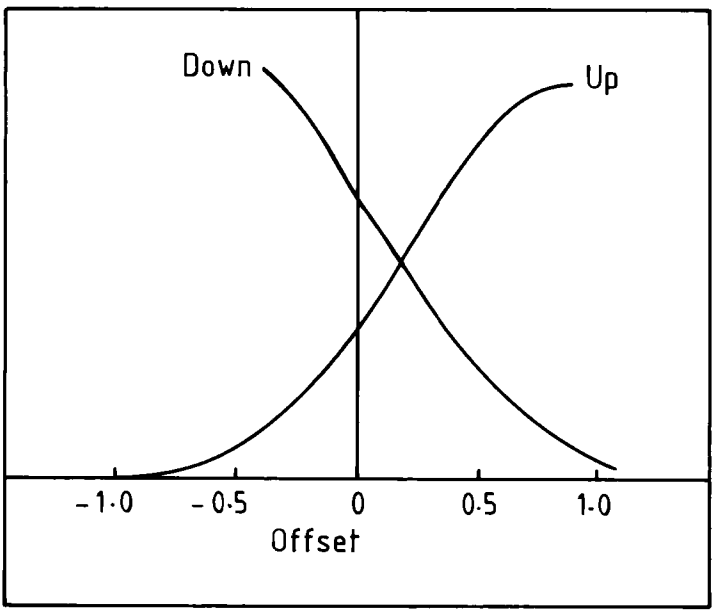


Figure 10. Lubricant Transport - Fully Flooded Ring

6. MULTIPLE RINGS

6.1 Compression Rings

The analysis presented in Section 5 relates to a single piston ring, but ring packs contain at least two compression rings. The analysis of a multiple compression ring configuration is straightforward, providing that consideration is given to the requirements of continuity of flow as shown in Figure 11. This is satisfied by assuming that the lubricant available to the (nth) ring is equal to that passing beneath the leading ring (n+1) or (n-1) in any stroke. It is necessary to apply the concept of lubricant starvation to the analysis of each ring.

In many modern engines the top compression ring presents a parabolic profile, while the second scraper ring is initially of straight taper form. Such scraper rings cannot promote fluid film lubrication on the downstroke and hence they operate in the boundary lubrication regime. On the upstroke they are capable of excellent hydrodynamic action, providing that there is an adequate layer of oil deposited on the cylinder liner. Scraper rings readily wear as a consequence of the severe lubrication conditions on the downstroke. This action changes the scraper ring profile and in due course it is often possible to represent the profile by a parabola of large negative offset.

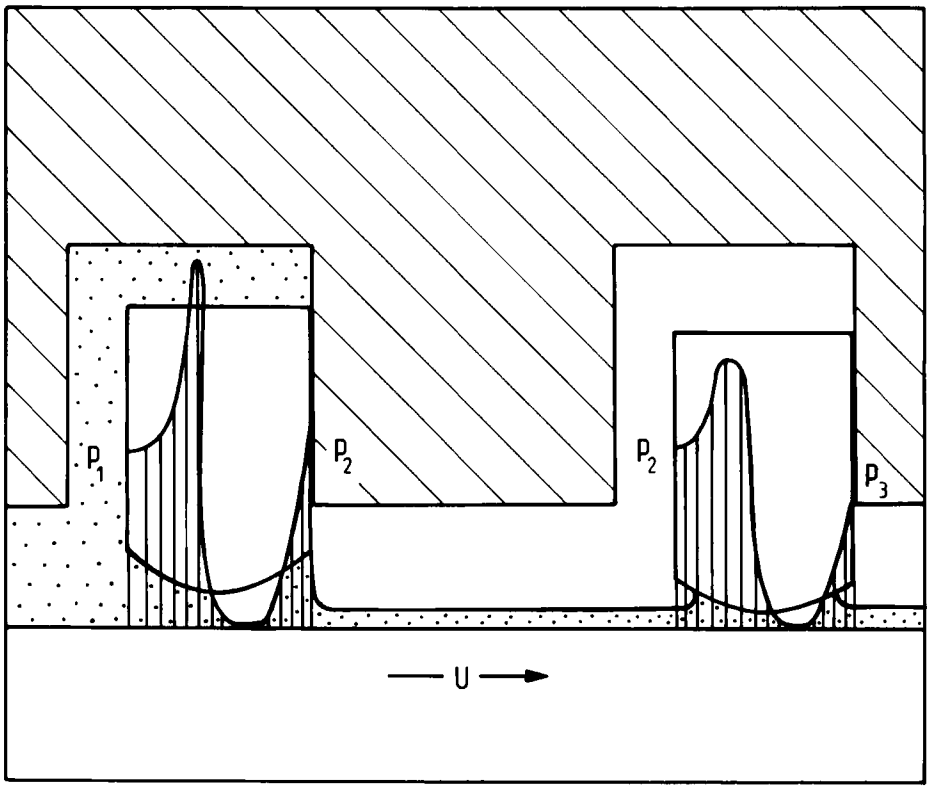


Figure 11 Multiple Ring Pack

6.2 Oil Control Rings

Many oil control rings present two connected or independent lands in contact with the cylinder wall as shown in Figure 12. The lubrication and friction of twin

land oil control rings has been analysed by Ruddy et al (1981(a)). The ring readily twists, thus promoting wear of the ring faces, and it was found that predicted performance was sensitive to the details of the separate and relative geometries of the two lands and the composite surface roughness of the ring faces and the cylinder liner.

It is not uncommon to find that the steady spring load on the twin lands of the oil control ring causes it to operate in the mixed or boundary lubrication regime. The friction force and hence the power loss can thus be estimated by assuming a constant coefficient of friction. In the mixed lubrication theory advanced by Ruddy

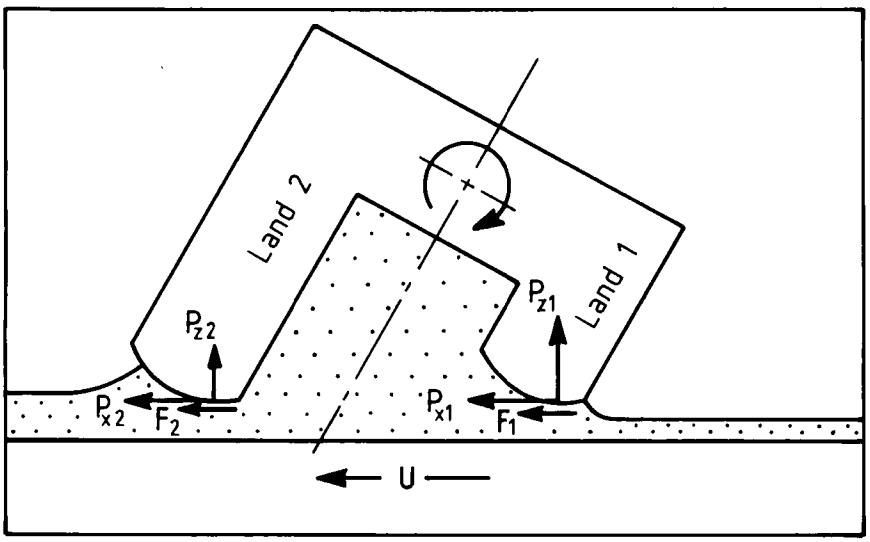


Figure 12 Twin Land Oil Control Ring

et al (1981(a)) a sharing of the applied load between asperity interaction and partial hydrodynamic films was considered, but this greatly complicates the analysis.

6.3 Complete Ring Pack

The analysis of a complete ring pack, consisting of compression rings and an oil control ring, calls for a careful consideration of the essential link between all the rings through the requirements for flow continuity. This is relatively straight forward in relation to the volume rates of flow beneath each ring, but there are still uncertainties about the role of oil accumulated in the inter-ring volumes and the ring grooves. A ring pack analysis has been presented by Dowson et al (1973). In most analyses it is assumed that each ring encounters a layer of oil on the liner of thickness equal to that deposited by the previous ring. It is assumed that the oil control ring is fully flooded on the downstroke and that the top compression ring encounters the layer which it deposited on the downstroke as it moves towards T.D.C. The situation is further complicated if ring lift occurs at any stage in the cycle.

6.4 Ring Lift

Once the lubrication analysis of a complete ring pack has been completed and the hydrodynamic force components calculated, it is possible to assess the possibility of ring lift by considering the force balance represented by equation (6).

Ring lift has been analysed by Ruddy et al (1979) (See also Ruddy, B L 1979). Ring lift is most likely to occur near to T.D.C. at high speeds (large ring inertia) and light loads (small restraining gas pressures) shortly after T.D.C. Once the ring has lifted off the groove face the gas pressures communicate readily with the inter ring volumes. The action may also be associated with the release of volumes of oil accumulated in the inter-ring spaces and ring grooves.

7. DISTORTED RINGS, GROOVES and CYLINDER BORES

The analysis presented earlier has assumed symmetry between the piston and cylinder bore and a steady ring face geometry. In reality the forces and associated moments acting on the ring (see Figure 4 and equations (5,6,7)) will cause the ring section to twist by an amount which varies throughout the cycle. The influence of ring twist upon the hydrodynamic performance of piston rings has been considered by Ruddy et al (1979). Since a small angular twist readily changes the offset and the latter has a pronounced effect upon film thickness, friction and lubricant transport, it is clear that ring twist is an important feature of ring pack performance.

The effective geometry of the ring face presented to the cylinder liner may also be affected by thermal and elastic distortions of the piston and liner and also by wear of the ring groove flanks and the liner. The influence of such distortions upon piston ring lubrication has been considered by Ruddy et al (1980).

Cylinder bores are not strictly circular in section, as assumed in this simple analysis. There are errors of shape, ovality produced at manufacture to anticipate shape changes in the running engine, assembly deformations and changes in elastic/ thermal distortions throughout each cycle. The elastic distortions created by bolting on the cylinder head generally have a pronounced effect upon cylinder bore profiles. The bores often present multiple lobes such that it is necessary to determine whether or not the rings conform and follow the liner profile, albeit with varying contact pressures in the circumferential direction.

In distorted bores the film thickness varies circumferentially, but the circumferential pressure gradients are still small compared with the axial gradients, such that the simple, no side-leakage, form of the Reynolds equation (13) can be retained. The lubricant transport is greatly affected by deviations from the assumed circular profile, but overall power loss is much less seriously affected.

8. SUMMARY

The analysis of piston ring lubrication is less advanced than that of engine bearings, but considerable progress has been recorded in the past fifteen years or so. Many piston rings experience fluid-film lubrication for much of each operating

cycle and hence their operating characteristics in these periods can be analysed by conventional hydrodynamic theory. However, like many dynamically loaded reciprocating components, they also encounter mixed and boundary lubrication during the severe conditions of operation near T.D.C. Indeed, Dowson et al (1983) have suggested that elasto-hydrodynamic lubrication may be experienced near to T.D.C. Piston rings are therefore subjected to wear, with a continuous process of evolution of surface finish, ring face geometry and hence of tribological characteristics throughout the engine life.

Interest in reducing engine power loss and oil consumption in recent years has stimulated further advances in the analysis and design of satisfactory piston ring packs. Both power loss and oil consumption have been reduced markedly through improved understanding of piston ring tribology. An excellent review of tribological aspects of piston assembly design has been presented by Ruddy and Hillyard (1991).

Further progress is likely to come from improvements in analysis, design, materials and manufacture. In the realm of analysis and design it is evident that the life-cycle performance of a ring pack must be considered and that simplified design procedures must be developed.

References

1 **Dowson, D, Economou, P N, Ruddy, B L, Strachan, P J and Baker, A J S. (1973),** 'Piston Ring Lubrication. Part II - Theoretical Analysis of a Single Ring and a Complete Ring Pack', in 'Energy Conversion through Fluid-Film Lubrication Technology', A.S.M.E., pp 23-52.

2 **Dowson, D, Ruddy, B L and Economou, P N (1983),** 'The Elastohydrodynamic Lubrication of Piston Rings, Proc. R. Soc. London, A386, pp 409-430.

3 **Ewiss, M. (1935).** 'Reibungs-und Undichtigkeitsverluste an Kolbenringen', Forschungshefte des Vereins Deutscher Ingenieure, No. 371.

4 **Joy, D. (1855).** 'Description of a Spiral Coil Piston Packing', Proc. I.Mech.E., pp 171-176.

5 **Miller, G M (1862).** 'On a Packing for Pistons of Steam Engines and Pumps', Proc. I.Mech.E.

6 **Olsson, K O. (1965).** 'Cavitation in Dynamically Loaded Bearings', Chalmers Tek. Högskolas Handl, No. 308, pp 1-60.

7 **Ramsbottom J. (1854).** 'On an Improved Piston for Steam Engines', Proc. I.Mech.E., pp 70-74.

8 **Ruddy, B L. (1979).** 'The Lubrication and Dynamics of Piston Rings and the Theoretical Predictions of Ring Pack Gas Flow', Ph.D. Thesis, University of Leeds.

9 **Ruddy, B L, Dowson, D, Economou, P N and Baker, A J S. (1979),** 'Piston Ring Lubrication Part III - The Influence of Ring Dynamics and Ring Twist', Energy Conversion Through Fluid Film Lubrication Technology, A.S.M.E. pp 191-215.

10 **Ruddy, B L, Dowson, D. and Economou, P N. (1981(a)).** 'A TheoreticaAnalysis of the Twin-Land Type of Oil-Control Piston Ring ', I.Mech.E. Journal of Mechanical Engineering Science, Vol. 23, No. 2, pp 51-62.

11 **Ruddy, B L, Dowson, D and Economou, P N. (1981(b))**, 'The Prediction of Gas Pressures Within the Ring Packs of Large Bore Diesel Engines', I.Mech.E., Journal of Mechanical Engineering Science, Vol. 23, No. 6, pp 295-304.

12 **Ruddy, B L and Hildyard, M L (1991).** 'A Review of Tribological Aspects of Piston Assembly Design', Proc. 17th Leeds-Lyon Symposium on Tribology, Elsevier, pp 93-102.

13 **Ruddy, B L, Parsons, B, Dowson, D. and Economou, P N. (1980).** 'The Influence of Thermal Distortion and Wear of Piston Ring Grooves upon the Lubrication of Piston Rings in Diesel Engines', Proc. 6th Leeds-Lyon Symposium on Tribology, Elsevier, pp 84-94.

14 **Ting, L L and Mayer, J.E. Jr. (1974).** 'Piston Ring Lubrication and Cylinder Bore Analysis - Part I, Theory; Part II - Theory Verification', Trans. A.S.M.E., Series F, Journal of Lubrication Technology, Vol. 96, No. 2, pp 258-266, No. 3, pp 305-314.

Engine Tribology / C.M. Taylor (Editor)

ANALYSIS OF THE PISTON ASSEMBLY, BORE DISTORTION AND FUTURE DEVELOPMENTS

R.J. Chittenden* and M. Priest*

*The Industrial Unit of Tribology, Institue of Tribology, Department of Mechanical Engineering, The University of Leeds, Leeds LS2 9JT, United Kingdom

1. THE PISTON

1.1 Introduction

When considering the tribology of the internal combustion engine the part played by the piston is often noted but its importance may not be fully recognised. The historical focus of attention has been the ring pack which has been found to have an important influence over may aspects of engine operation. Consideration of frictional dissipation, however, shows that the typical losses associated with the piston skirt might be up to one hundred and fifty percent of those associated with the ring pack. In addition the design of the piston can have a significant part to play in the understanding of engine noise, where the phenomenon of "piston slap" has been the subject of a number of investigations, for example Kaiser, Schmillen and Spessert (1988).

In operation the piston is not only driven axially up and down within the cylinder liner (the primary motion) but can also translate in the transverse direction and rotate around the gudgeon pin axis. This secondary motion is influenced by the following factors:-

- gas and inertia forces
- piston clearance
- offset of the gudgeon pin centre from the centre line of the piston
- hydrodynamic friction forces acting around the gudgeon pin and along the piston skirt
- elastic deformation of the piston skirt during impact with the wall of the cylinder liner

The experimental investigation of piston assembly friction may be traced back to the years of the Second World War. These studies concentrated upon the experimental evaluation of different parameters affecting the design of a piston. The development of a theoretical analysis of the piston skirt was started in the early 1950's though clearly a full solution, considering all the influences listed above, could not be expected from these first calculations. For example, Meir (1952) who published an analysis of piston lateral motion

in which the hydrodynamic action was neglected. (Experimental evidence as to the availability of oil to form a hydrodynamic film was subsequently provided by Greene (1969), who used a glass cylinder to allow the lubricant supply to be monitored photographically during the engine cycle.)

The advent of the digital computer allowed the piston to be modelled in much greater detail (Bishop and Leavitt (1975)), and when coupled with the measurements obtained from the parallel evolution of transducers, (Munro and Parker (1975)), a link with the design of pistons could begin to be established. Such work, however, was still limited to the dynamic action of the piston skirt and did not consider the influence of any lubricant film on the cylinder liner. In the early 1980's this restriction was removed by the work published by several research groups, Knoll and Peeken (1982), and Li, Rhode and Ezzat (1982). The latter work also led to a further refinement in the computations where modelling of the skirt by finite elements allowed any elastic deflections of the piston assembly to be accounted for within the analysis, Oh, Li and Goenka (1987).

The analysis of the piston skirt may be seen to have progressed greatly over the past forty or so years. Calculations can be made which show the influence of the various important features listed above. This is therefore of interest to the piston designer who can use these predictions as a guide before developing trial piston configurations which can then be evaluated on real engines. The experimental testing of pistons to show the influence of lightweight design, Azevedo and Filho (1988) or investigate their structural integrity Frisch (1988) should eventually be much reduced though not yet entirely eliminated.

1.1.1 Notation

c	piston radial clearance within the cylinder liner (m)
e_b, e_t	eccentricity at the top and bottom of the piston skirt respectively (m)
F_{cg}	Axial inertia force due to piston mass (N)
F_G	Axial force due to gas pressure acting on the piston crown (N)
F_{pin}	Axial inertia force due to gudgeon pin and connecting rod small end masses (N)
F_{rod}	Force acting along the connecting rod due to the reciprocating action of the piston (N)
G_H	Radial hydrodynamic force (N)
G_{cg}	Radial inertia force due to piston mass (N)
G_{pin}	Radial inertia force due to gudgeon pin and connecting rod small end masses (N)
h	lubricant film thickness (m)
I_{pis}	Moment of inertia of the piston about its centre of mass (kg m^2)
l_c	length of the engine crank (m)
l_p	length of the piston skirt (m)
L	length of the connecting rod (m)
m_{pin}	combined gudgeon pin and connecting rod small end masses (kg)
m_{pis}	mass of the piston (kg)
M_H	Hydrodynamic moment acting about the centre of mass of the piston (Nm)
M_I	Inertia moment acting about the gudgeon pin axis (Nm)

N	engine speed (rpm)
p	hydrodynamic fluid pressure (Pa)
P_G	combustion chamber gas pressure (Pa)
r	engine half-stroke (crank radius) (m)
R	radius of the piston skirt (m)
t	time (s)
U	piston sliding velocity (note, $^+$ve towards the crankcase) (m/s)
X	axial coordinate measured from the top of the piston skirt (m)
x_{cg}	distance of piston centre of mass from top of piston skirt (m)
x_{pin}	distance of gudgeon pin centre from top of piston skirt (m)
Y	radial coordinate for the piston skirt, measured from the piston axis (m)
y_{cg}	radial displacement of piston centre of mass from the piston centre line (m)
y_{pin}	radial displacement of the gudgeon pin from the piston centre line (m)
η	lubricant dynamic viscosity (Pas)
ω	engine crankshaft rotational speed (rad/s)
μ	coefficient of friction under non- hydrodynamic conditions (-)
θ	connecting rod angle (rads)
ϕ	circumferential angle around the piston axis, measured from major thrust axis (rads)

Subscripts

t	top (towards combustion chamber) of the piston skirt
b	bottom (towards crank case) of the piston skirt

1.2 Assumptions made by the Model

In developing a numerical model of this even this complexity it is necessary to make a number of simplifying assumptions. In general these follow the restrictions imposed upon the analytical model for the piston rings, with additional aspects to account for the more complex three-dimensional nature of the problem.

- the piston skirt and cylinder liner are perfectly smooth and round
- the influence of the ring pack upon the piston motion is negligible
- any friction due to the twisting of the piston pin may be neglected as may the rotational inertia of the pin
- the lubricant viscosity is constant along the length of the skirt and during the engine cycle
- the lubricant is Newtonian (ie. shear stress is proportional to shear rate)
- the hydrodynamic pressures generated do not lead to significant change in the lubricant viscosity or deformation of the bounding solids
- connecting rod inertia may be neglected

1.3 Forces Acting upon the Piston

In developing a simple mathematical model of the piston the forces and moments acting upon it must first be considered. They may be divide into four main groups:

- the force acting in the axial direction and arising from the reciprocating action of the piston and the gas pressure on the piston crown [F_{pin}, F_{cg}, F_G and also a component of M_I]
- hydrodynamic forces and moment acting against the transverse motion of the piston [G_H, M_H and G_{pin}]
- inertia forces acting in the transverse direction due to the transverse accelerations of the piston [G_{cg} and a component of M_I]
- forces acting along the connecting rod, which may be attributed to the reciprocating mass of the piston assembly [F_{rod}]

These forces and moments are shown in Figure 1.1 along with the required dimensions for the calculation of moments.

1.4 The Equations of Motion

The equations of motion may easily be written for both the axial and transverse directions. Considering the equilibrium of forces and moments about the gudgeon pin the following expressions may be derived.

For the piston axis (the X direction):-

$$F_G + F_{cg} + F_{pin} + F_{rod} \cos\Phi = 0 \tag{1.1}$$

Similarly for the piston transverse direction (the Y direction):-

$$G_H + G_{pin} + G_{cg} - F_{rod} \sin\Phi = 0 \tag{1.2}$$

Looking now at moments about the gudgeon pin:-

$$M_I + M_H - F_{cg}\, y_{cg} + F_G\, y_{pin} + G_{cg}\,(x_{pin} - x_{cg}) = 0 \tag{1.3}$$

Note, the reciprocating inertia forces for a constant crankshaft speed (ω) are given by:-

$$F_{cg} = -m_{pis}\, \ddot{X} \tag{1.4}$$

$$F_{pin} = -m_{pin}\, \ddot{X} \tag{1.5}$$

where x is given by the standard expression for piston axial acceleration derived from the engine geometry.

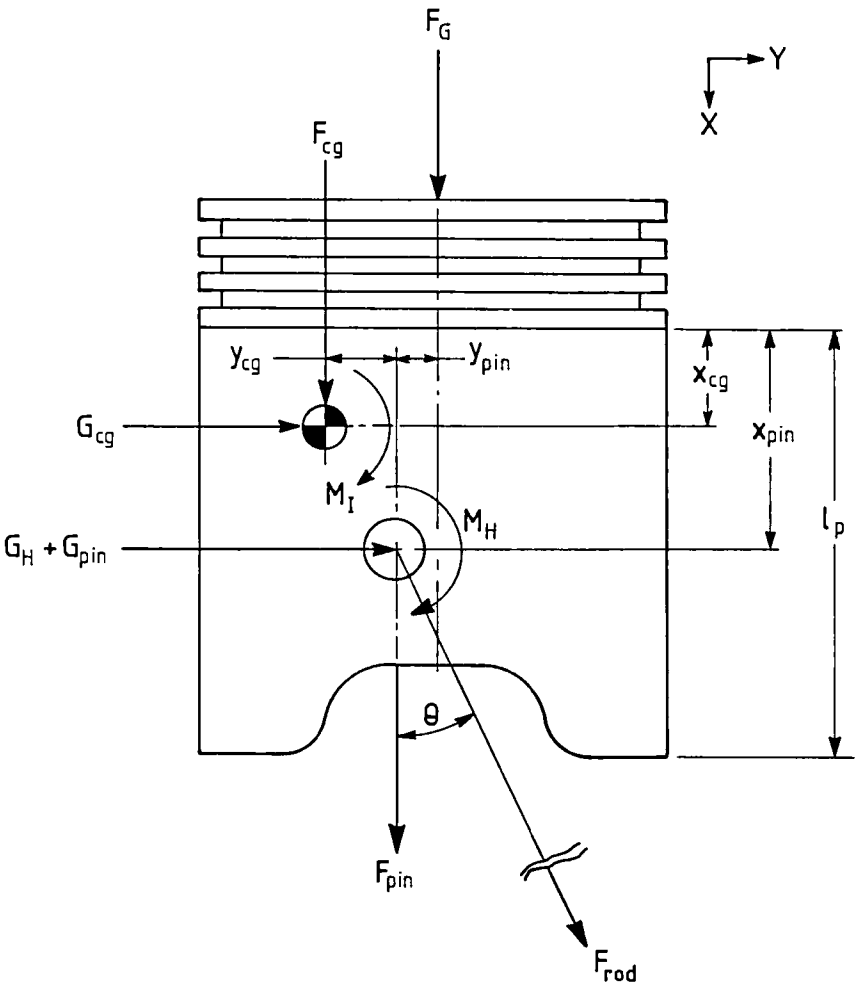


Figure 1.1 Forces and Moments Acting on the Piston

1.5 Lubrication Analysis for a Reciprocating Piston

The Reynolds' equation for a fixed set of axes, and conditions in which lubricant flow is assumed to take place along and across the entraining direction, may be written, in polar coordinates as:-

$$\frac{\partial}{\partial\phi}\left(h^3\frac{\partial p}{\partial\phi}\right)+R^2\frac{\partial}{\partial X}\left(h^3\frac{\partial p}{\partial X}\right)=12\eta R^2\left[\frac{\partial h}{\partial t}-\frac{U}{2}\frac{\partial h}{\partial X}\right] \qquad (1.6)$$

where X is in the piston axial direction and ϕ is measured around the piston circumference from the major axis, and the film thickness, h, is obtained from the approximation:-

$$h = c + e_t\cos\phi + X/l_p(e_b - e_t)\cos\phi \qquad (1.7)$$

Note the similarity between this version of the Reynolds' equation and that, also in a polar coordinate form, used for the analysis of dynamically loaded journal bearings.

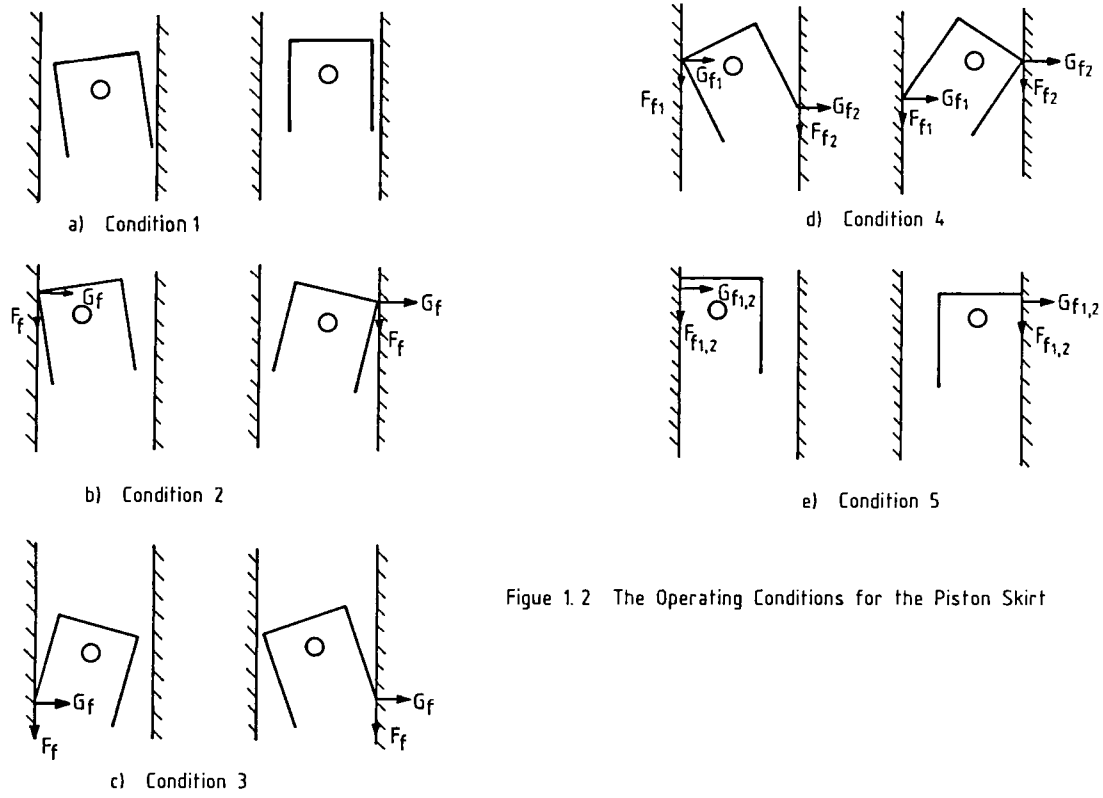
It may be seen that the left hand side of this expression involves pressure gradients in both the piston axial and circumferential directions. The right hand side requires a transverse, squeeze, velocity resulting from the motion of the piston across the piston/ liner clearance and knowledge of the changes in the shape of the lubricant film in the direction of motion.

1.6 Solution Considerations

From the analysis presented in the preceding sections it is clear that the calculation of piston transverse motion, and hence indication of piston slap or power loss, requires the simultaneous solution of the equation governing the dynamic and hydrodynamic action of the piston. This may be achieved my a number of methods which will not be covered in these notes. The following few comments, however, illustrate some of the points that must be covered by any solution algorithm.

Consider the piston moving transversely across the piston/ cylinder liner clearance, and rotating about the gudgeon pin at the same time. In the general situation detailed above it was assumed that contact was not made between the piston and the cylinder wall. This may be thought of as "condition 1" and is illustrated in Figure 1.2(a). The motion of the piston may be such that contact occurs, resulting in one of the four following possibilities:-

"condition 2" - contact at the top of the skirt only, see Figure 1.2(b)
"condition 3" - contact at the bottom of the skirt only, see Figure 1.2(c)
"condition 4" - contact at both top and bottom of the skirt but on different sides of the piston, see Figure 1.2(d)



Figue 1. 2 The Operating Conditions for the Piston Skirt

Figure 1.2 Operating Conditions for the Piston Skirt

"condition 5" - contact at both top and bottom of the skirt on the same side of the piston, see Figure 1.2(e)

In the latter conditions were contact exists the force and moment expressions developed earlier must be amended to incorporate the resulting additional friction forces. The simplest way to do this, and that used in the following examples, is by the introduction of an *equivalent coefficient of friction under non- hydrodynamic conditions* (μ). As an example, "condition 3" would require the introduction of two forces as shown in Figure 1.2(c), whilst "condition 4" and "condition 5" would result in a state of quasi- static equilibrium.

Parameter	Notation	Value
Crank Radius	l_c	0.045m
Length of the Connecting Rod	L	0.150m
Mass of the Piston	m_{pis}	0.75kg
Gudgeon Pin and Connecting Rod Small End Masses	m_{pin}	0.15kg
Offset of Gudgeon Pin Axis from Piston Centre Line	y_{pin}	-0.15×10^{-3}m
Offset of Piston C of G from Piston Centre Line	y_{cg}	0.0m
Position of Piston C of G from Top of Skirt	x_{cg}	0.02×10^{-3}m
Position of Gudgeon Pin Centre from Top of Skirt	x_{pin}	0.017m
Radius of the Piston Skirt	R	0.048m
Length of the Piston Skirt	l_p	0.060m
Effective Lubricant Dynamic Viscosity	η	0.0065Pas
Moment of Inertia of the Piston about its centre of mass	I_{pis}	6.5×10^{-4}kg m^2
Effective Piston/ Liner Clearance	c	1.5×10^{-5}m
Engine Speed	ω	2000rpm
Effective Coefficient of Friction under non- Hydrodynamic Conditions	μ	0.07

Table 1. Parameters Employed in the Example Calculations

1.7 Example Calculations of Piston Dynamics

As an example of the influence the various parameters have upon the operation of the piston skirt consider the engine operating conditions for a typical medium bore automotive spark ignition engine shown in Table 1. The corresponding combustion chamber pressures are illustrated in Figure 1.3

The movement of the skirt within the piston/ liner clearance is illustrated for a single engine cycle in the following figures. Note that the position of a point on the piston axis at the top and bottom of the skirt is presented in terms of an eccentricity ratio (e_t/c or e_b/c). A value of unity would therefore imply that movement had taken place from the concentric position such that the entire radial clearance had been taken up and the piston would therefore be in contact with the cylinder liner. Positive values indicate movement towards the minor thrust side and negative values towards the major thrust side.

The parameters used to show the changes in the operation of the piston skirt are set out in Table 2.

x_{pin} (m)	x_{cg} (m)	y_{pin} (m)	y_{cg} (m)	c (m)	η (Pas)	Case ref
0.017	0.00002	-0.00015	0.0	0.000015	0.0063	1
0.017	0.00002	0.00000	0.0	0.000015	0.0063	2
0.017	0.00002	0.00015	0.0	0.000015	0.0063	5
0.017	0.00002	-0.00030	0.0	0.000015	0.0063	11
0.030	0.00002	-0.00015	0.0	0.000015	0.0063	3
0.045	0.00002	-0.00015	0.0	0.000015	0.0063	4
0.000	0.00002	-0.00015	0.0	0.000015	0.0063	6
0.017	0.00002	-0.00015	0.0	0.0000075	0.0063	7
0.017	0.00002	-0.00015	0.0	0.0000225	0.0063	8
0.017	0.00002	-0.00015	0.0	0.000015	0.0239	9
0.017	0.00002	-0.00015	0.0	0.000015	0.0027	10

Table 2 Parameters used in the analysis of the piston skirt

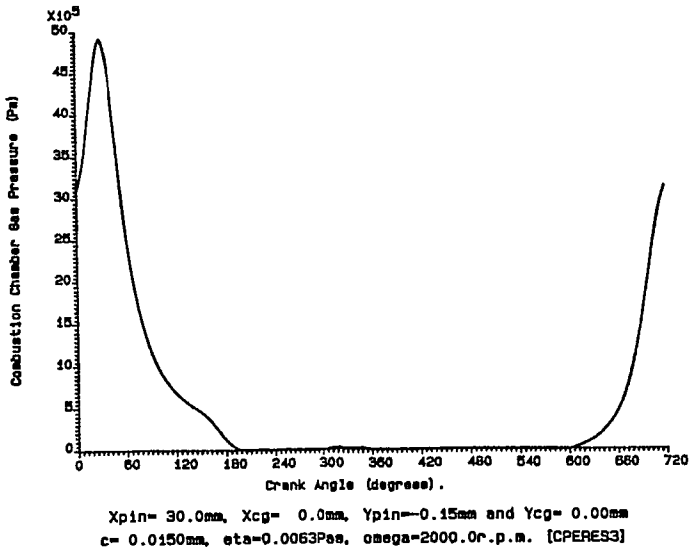


Figure 1.3 Combustion Chamber Pressures

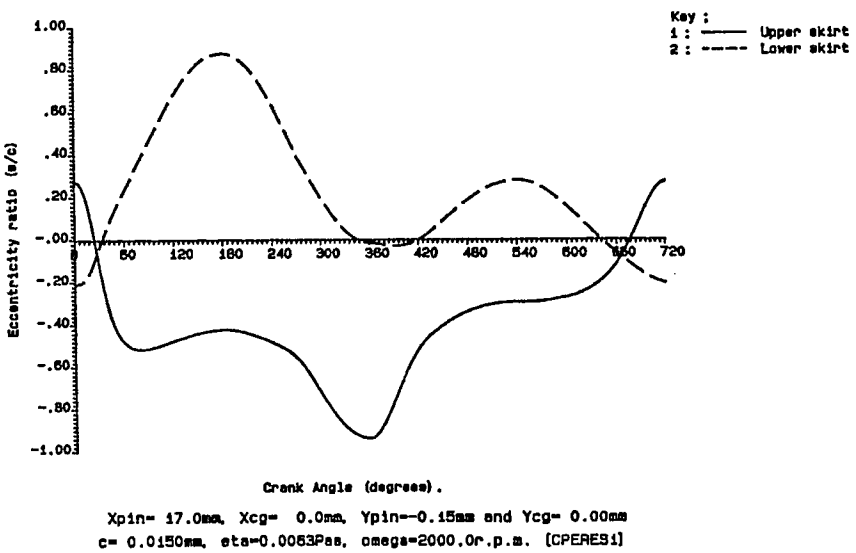


Figure 1.4 Predicted Piston Motion for Case 1

Figure 1.4 shows the predicted curves when the parameters take the values listed in Tables 1 and 2 (Case 1). It may be seen that as the gas pressures build up to top dead centre firing (0 or 720 degrees of crack angle) the effect is to twist the piston rather than to drive it towards one side of the cylinder liner or the other. Once 180 degrees of crank angle are reached, however, the translational effect becomes dominant. The piston first moves towards the major thrust side and then after 360 degrees towards the minor thrust side before the gas pressure starts to increase again and the cycle repeats itself.

Figure 1.5, case 2, shows a very similar situation, but here the position of the gudgeon pin is actually on the piston centre line rather than slightly displaced to the major thrust side. The only significant difference may be seen between 120 and 180 degrees of crank angle where the sharp change in motion of the lower skirt indicates contact between the piston and cylinder liner.

If the position of the gudgeon pin centre is positioned an identical distance from the piston centre line as for Figure 1.4 (Case 5), but this time towards the minor thrust side the result is illustrated in Figure 1.6. Here the contact situation has worsened compared with Figure 1.5 and hence the friction losses will have increased and the sharp gradient around 30 degrees of crank angle suggest that "piston slap" may be significant.

An increase in the displacement of the gudgeon pin centre, compared with case 1, from the piston centre line towards the major thrust side, however, shows that the performance of the piston skirt may be improved, Figure 1.7, case 11. The change is not large and the full range of engine operating conditions would have to be examined before it could be decided if such a position was actually of most benefit.

The axial position of the gudgeon pin is also important. This is illustrated in the next few figures starting with Figure 1.8, case 3, where the axial distance has been set to 0.030m from the top of the skirt compared with the 0.017m used in the base case, Figure 1.4. The result shows a very smooth piston motion where the only significant changes occur around the area of peak gas pressure.

Increasing the displacement of the gudgeon pin from the top of the skirt, however, does not continue to have a beneficial effect upon the predicted piston motion. This is shown in Figure 1.9, case 4, where the pin is now 0.045m from the top of the skirt. The piston eccentricities will be seen to be much higher and contact with the cylinder liner is indicated around the position of top dead centre firing.

Decreasing the axial position of the gudgeon pin from the top of the piston, case 6, to a point actually at the top of the skirt shows that an optimum axial position for the gudgeon pin must exist. The illustration of this case, Figure 1.10, shows that number of degrees of crank angle over which contact is indicated has increased, but without any improvement in other parts of the engine cycle.

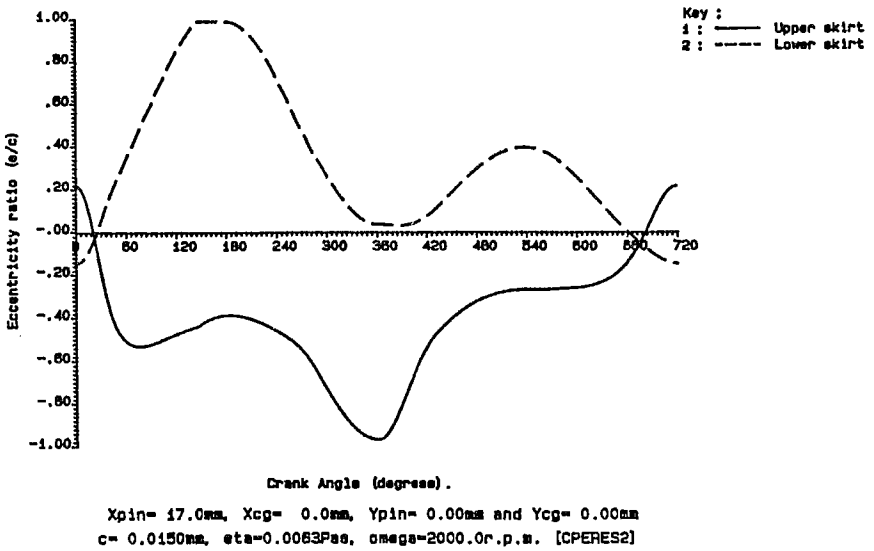


Figure 1.5 Predicted Piston Motion for Case 2

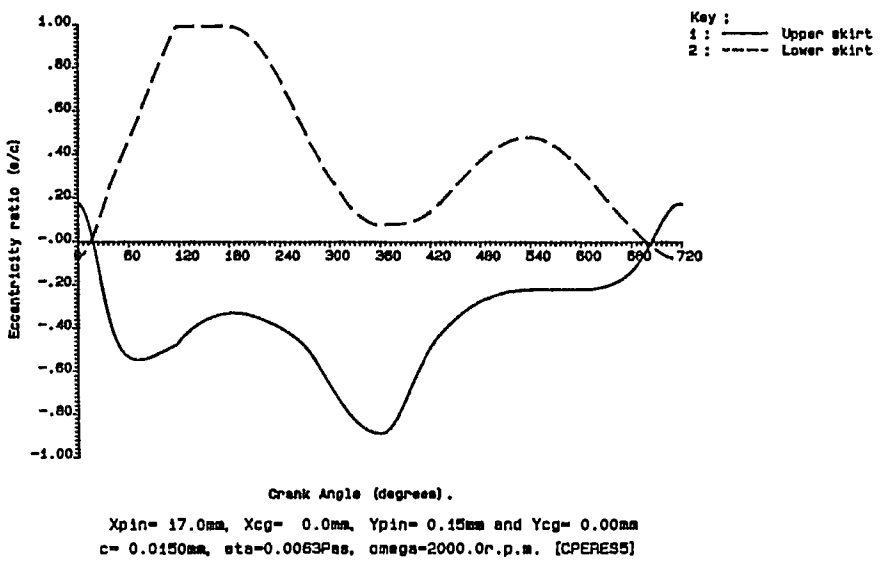


Figure 1.6 Predicted Piston Motion for Case 5

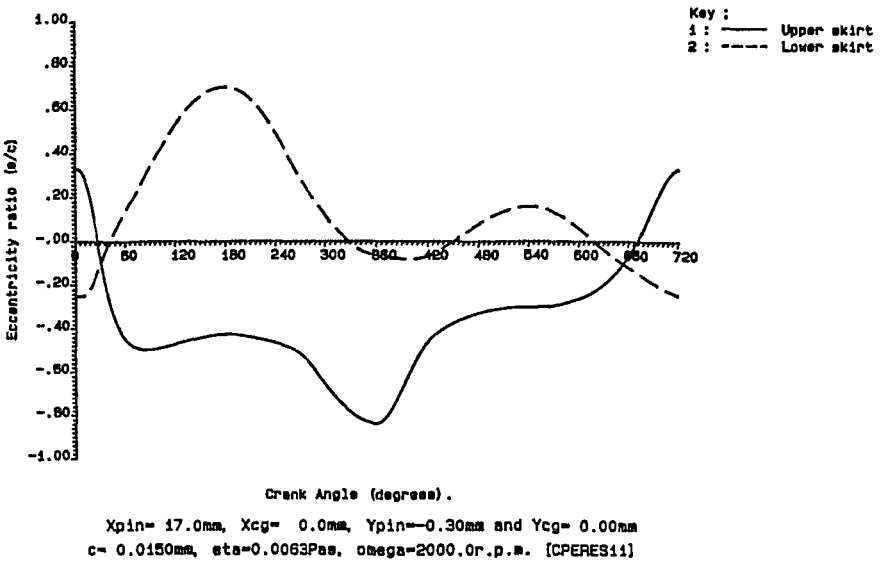


Figure 1.7 Predicted Piston Motion for Case 11

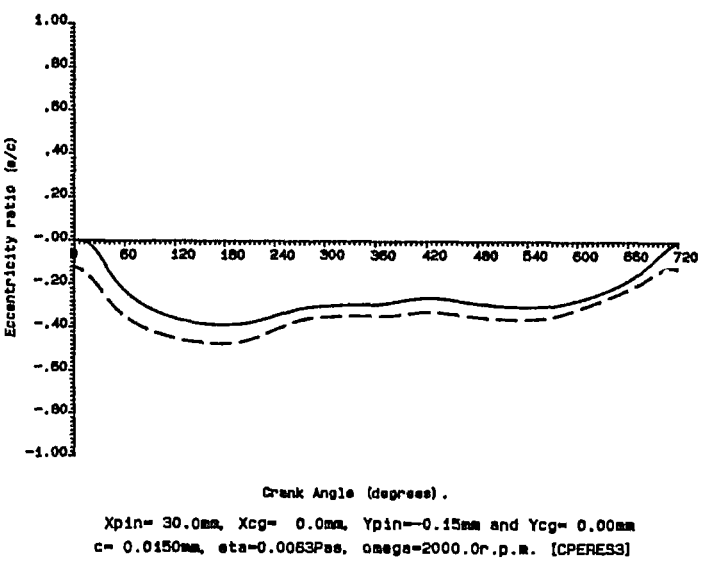


Figure 1.8 Predicted Piston Motion for Case 3

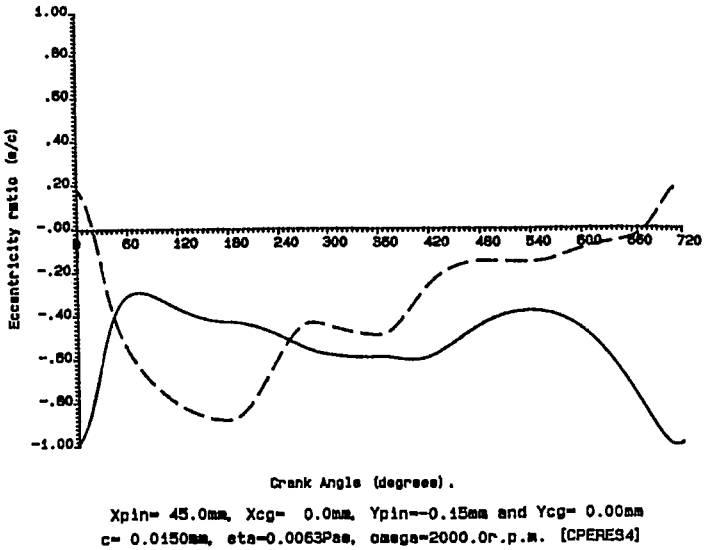


Figure 1.9 Predicted Piston Motion for Case 4

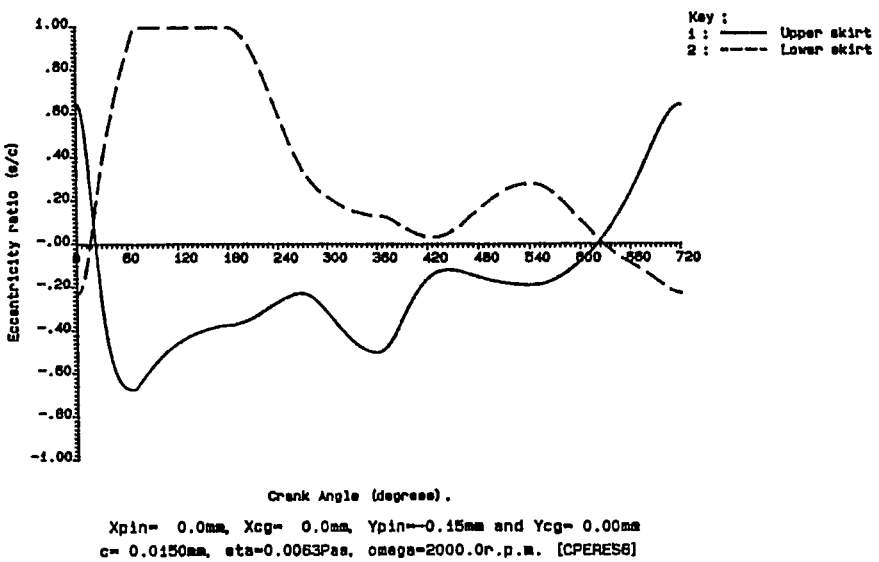


Figure 1.10 Predicted Piston Motion for Case 6

The influence of piston/ cylinder liner radial clearance is shown in Figures 1.11 and 1.12, cases 7 and 8 respectively. The former shows the effect of a reduced radial clearance. This may be seen to provide the piston with a much more controlled motion in which the piston not only experiences a smaller degree of tilt but also operates closer to the centre line of the cylinder liner. In the latter case the increased clearance, and hence a greater potential for inertia effects to influence the motion, results in the complete opposite. A significant tilt is observed around bottom dead centre after firing and both top and bottom of the skirt reach appreciable eccentricity ratios. Contact is also indicated between the lower skirt and cylinder liner at a similar position to maximum tilt.

Finally, Figures 1.13 and 1.14 (Cases 9 and 10) show how changes in the operating viscosity of the lubricant, whether caused by a change in temperature or oil grade, can feed through to the performance of the piston skirt. Comparison of Figure 1.13 with Figure 1.11 and Figure 1.14 with Figure 1.12 shows that changes in viscosity have very similar, but opposite, effects to changes in piston radial clearance. The form of Figures 1.13 and 1.14 are such that the comments made regarding the previous figures are equally applicable.

1.8 Conclusions

The examples considered in the previous section indicate that the analysis of the piston skirt has an important role in the understanding of the piston assembly as a whole.

The positioning of the gudgeon pin has been shown to have a strong influence over the piston lateral motion and tilt, with bad positioning causing piston/ cylinder liner contact to be predicted.

The influence of radial clearance and lubricant viscosity have also been examined. The results shown that if an engine is required to run on a lower viscosity lubricant then the loss in performance may be counterbalanced by a tightening of the piston radial clearance.

2 BORE DISTORTION

2.1 Introduction

In the preceding tribological models of the piston ring pack and piston skirt it was assumed throughout that the cylinder bore is perfectly circular and of nominal bore size along its entire length. In reality this is often not the case.

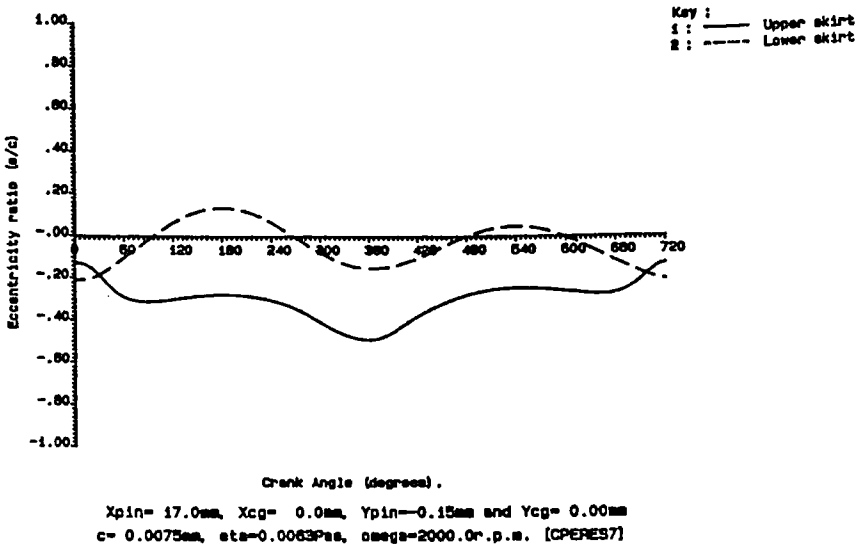


Figure 1.11 Predicted Piston Motion for Case 7

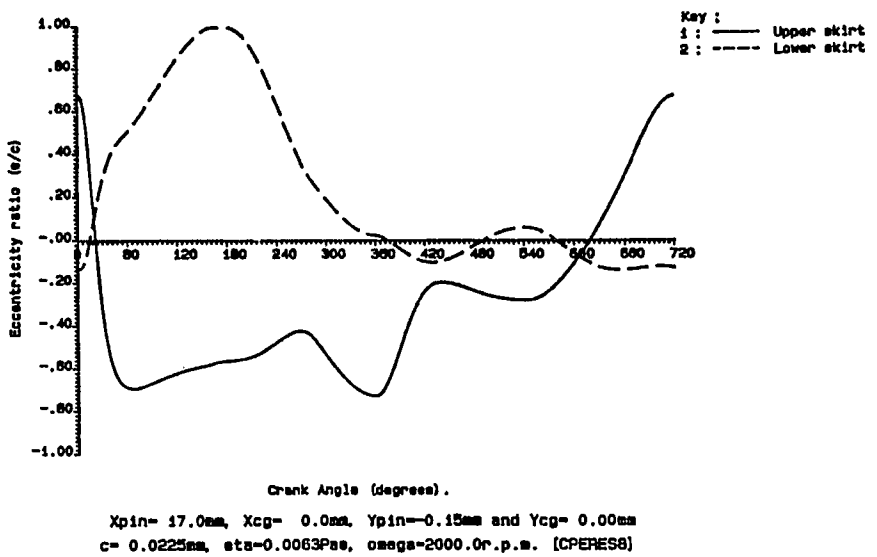


Figure 1.12 Predicted Piston Motion for Case 8

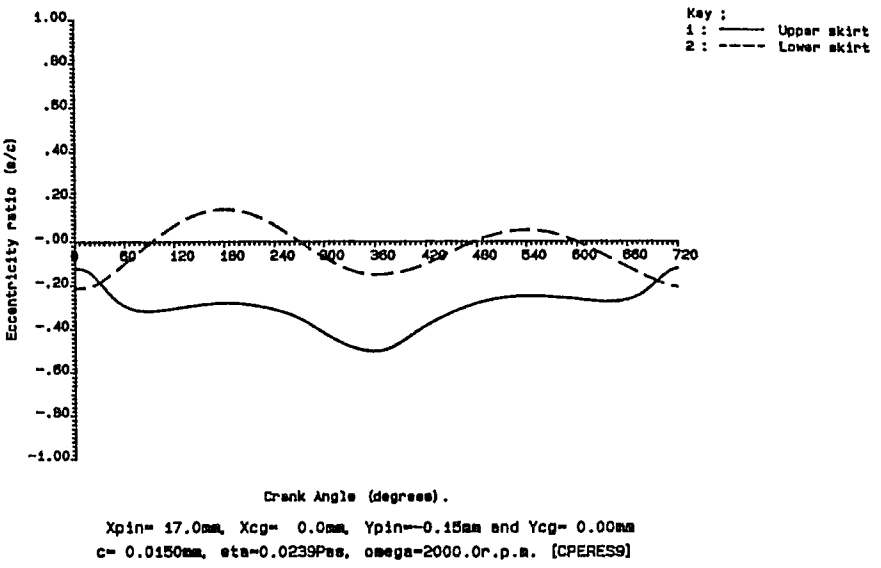


Figure 1.13 Predicted Piston Motion for Case 9

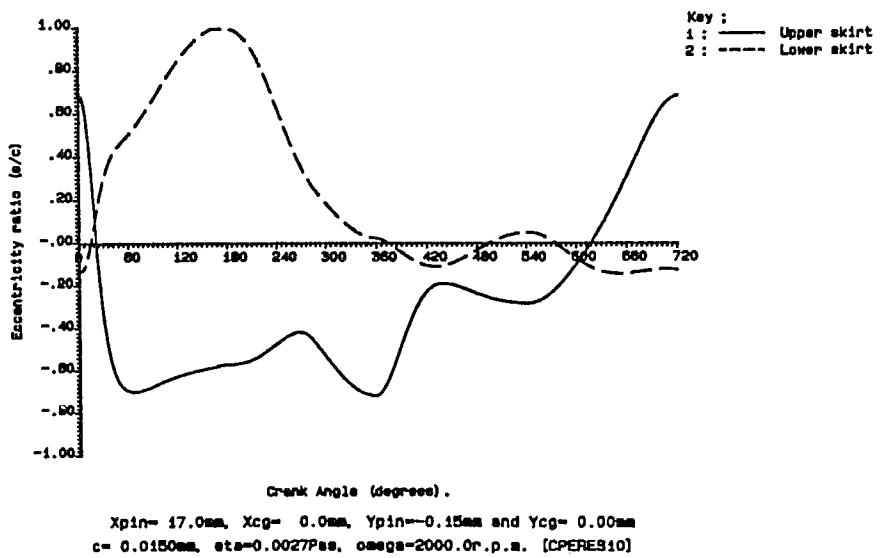


Figure 1.14 Predicted Piston Motion for Case 10

2.2 Causes of Bore Distortion

Non-circularity of cylinder bores may result from:

- Manufacture

Engineering components are machined to a level of accuracy specified in terms of tolerances, i.e. a maximum and minimum value for a given dimension. For a two litre automotive engine the difference between the maximum and minimum permissible diameter of the cylinder bore may be 30μm. In comparison, the thickness of the oil films between the piston rings and the cylinder bore and between the piston skirt and the cylinder bore are typically of the order of 1μm or less. The deviations from circularity within the manufacturing tolerances are therefore likely to have a significant effect on the performance of the piston assembly.

- Assembly

The process of assembling an engine can introduce additional stresses and deformations into the cylinder walls. An example of this is the tightening of cylinder head bolts.

- Thermal Effects

If the rate of cooling is not constant around the circumference of the cylinder and along its length then differential expansion will occur leading to distortion of the cylinder bore. This may result from inadequate cooling of a region of the cylinder as in `Siamese' bore engine designs or from over-cooling a region of the cylinder as in the case of a poorly designed coolant circulation system.

- Pressure Effects

This refers to distortion of the cylinders walls due to the pressure of the gases generated in the combustion chamber. It is only likely to be significant in highly rated diesel engines with thin-walled wet liners.

In most cases non-circularity of a cylinder bore results from a number of these effects acting together. For instance the combination of assembly and temperature can lead to badly distorted bores due to the expansion of the cylinder relative to its restraints such as the cylinder head bolts. In a firing automotive engine the difference between the maximum and minimum bore diameter may be as high as 150μm, which is five times the manufacturing tolerance.

2.3 Representation of Non-Circular Bore Geometries

The variation in geometry around the circumference of a distorted cylinder bore may be modelled by a general Fourier series

$$R(\phi) = \sum_{i=0}^{i=n} (A_i \cos\phi + B_i \sin\phi) \qquad (2.1)$$

where

$R(\phi)$ = radial coordinate
ϕ = angular coordinate
A_i, B_i = amplitude constants
i = order
n = highest order distortion to be considered

The coordinate system and various Fourier orders of bore distortion are illustrated in Figure 2.1.

Hill and Newman (1984), Fesser (1988) and Dunaevsky (1990) have all adopted this approach to describe distorted cylinder bores. It enables complex bore geometries to be broken down, using a Fast Fourier Transform (FFT) algorithm, into simpler components which can then be analysed to assess the reasons for the distortion and possible remedial action. For example if a fourth order distortion (see Figure 2.1) is a major component in a distorted cylinder bore then the tightening of the four cylinder head bolts would be questioned.

Zero and first order bore distortions (see Figure 2.1) are a function of the size and location tolerances of the cylinder bore. As these components are circular they are of little consequence to piston ring conformity. In automotive applications the predominant bore distortion components are the second, third and fourth orders. Higher orders of distortion may be of significance in particular engines.

2.4 Measurement

To obtain a true picture of the bore distortion of a particular engine it needs to be measured in a firing engine. One way of doing this would be to fit a number of proximity transducers to the piston skirt and measure the piston to bore clearance throughout the engine cycle. These clearance measurements would be a function of the transducer position on the piston skirt, the nominal size of the components, the distortion of the piston, the dynamics of the piston and bore distortion. If the other parameters can be measured, accurately predicted by computer models or neglected then a reasonable estimate of the bore geometry is achievable.

This form of experimentation is most definitely a research activity and is very expensive to undertake. It does not fulfil the need of the automotive industry for a tool to measure bore

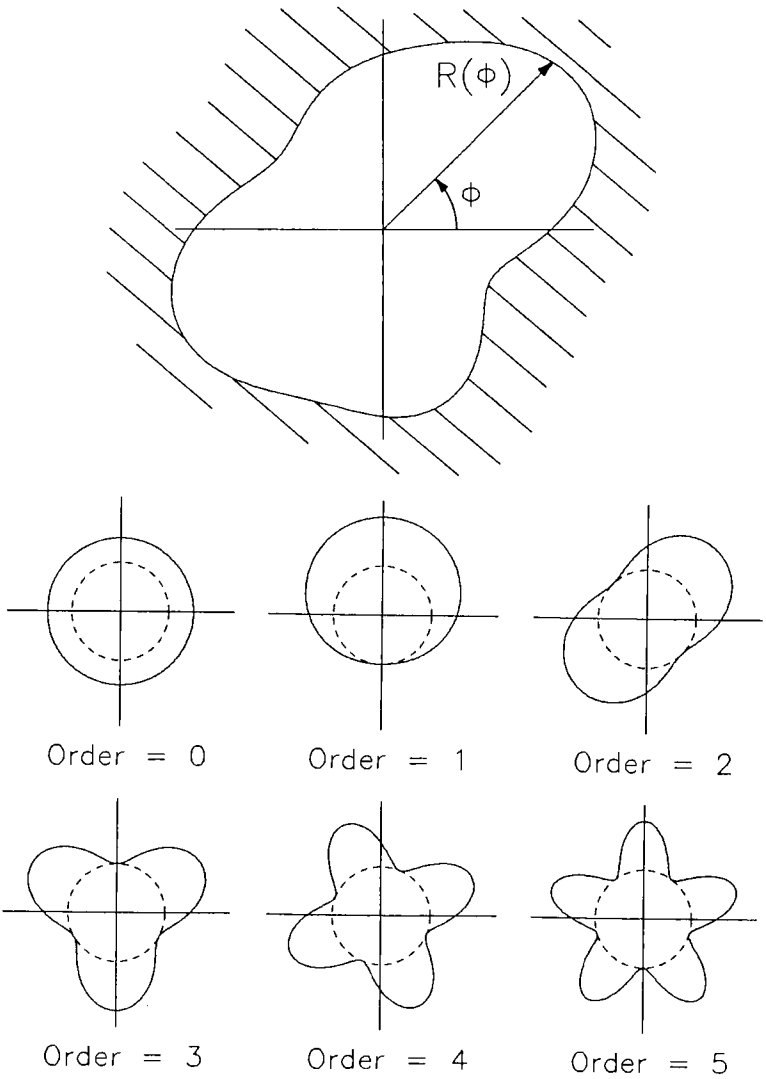


Figure 2.1 Non-Circular Bore Geometries;
Cylindrical Coordinate System and Fourier Orders

distortion in product development and quality assurance departments. Measurement devices have been developed for such applications using mechanical stylus instruments which are lowered in to the bore with the cylinder head removed. Bore distortion due to manufacture and assembly can be measured providing the cylinder head bolts are fitted. A rough indication of the effect of temperature can be obtained by heating the cylinder block. An example of this type of device is the 'Incometer' developed by Goetze AG (Loenne and Ziemba, 1988).

2.5. Piston Ring Conformity

By far the most significant consequence of bore distortion to the piston assembly is the potential loss of compliance, or conformity, between the piston rings and cylinder bore. The ability of the piston ring pack to conform to the cylinder bore surface is critical to the performance of an engine. If large gaps develop between the rings and the bore then excessive oil flow and blow-by of combustion chamber gases may occur (Daskivich, 1980). Precisely how piston rings react to bore distortion is therefore of great importance.

In evaluating piston ring to bore conformity two main options are available:

- empirical relationships derived from experimental data.
- stress analysis likely to extend to finite element analysis.

The former approach is more amenable to a preliminary study and is adopted here.

Much work has been done on establishing empirical relationships and models for the development of light-tight seals between piston rings and cylinder bores (Mierbach, Dueck and Newman, 1983; Hill and Newman, 1984; Brombolich and McCormick, 1988). In their study Hill and Newman (1984) derived the following relationships for the conformability factor K, which is a measure of the ability of a piston ring to follow the geometry of the bore, and for U_i, which is the level of deformation of the cylinder bore of the ith order able to be sealed statically by a piston ring

$$k = \frac{3F_t(d-a)^2}{Eha^3} \tag{2.2}$$

where

F_t = piston ring tangential load (N)
E = modulus of elasticity (Pa)
d = nominal cylinder bore diameter (m)
a = radial thickness of piston ring (m)
h = axial height of piston ring (m)

$$U_i = \frac{rK}{\left(i^2 - 1\right)^2} \tag{2.3}$$

where

r = nominal cylinder bore radius $\left[= \frac{d}{2}\right]$ (m)

U_i is the difference between the maximum radial coordinate and the minimum radial coordinate of the distorted bore which is twice the amplitude of any given Fourier order. In the piston ring analysis described earlier in this course it is implicit that

$$F_t = \frac{F_r(d - a)}{2} \tag{2.4}$$

where

F_r = piston ring radial load per unit circumference

Substituting equation (2.4) in (2.2)

$$K = \frac{3F_r(d - a)^3}{2Eha^3} \tag{2.5}$$

and substituting equation (2.5) in (2.3)

$$U_i = \frac{3F_r d(d - a)^3}{4Eha^3\left(i^2 - 1\right)^2} \tag{2.6}$$

If both gas pressure and inherent ring tension are included in the radial load term F_r then a useful expression is obtained to examine the ability of a piston ring to conform to a distorted cylinder bore. It should be noted however that this is an empirical expression and as such is only approximate. Dynamic effects and the influence of temperature are neglected and the equation is not valid near the ring gap.
Equation 2.6 has been applied to the following bore distortions for a typical automotive compression ring with no gas pressure acting such that the radial load is due only to the inherent tension of the ring:

Case 1 - second order bore distortion of 25μm amplitude
Case 2 - third order bore distortion of 25μm amplitude
Case 3 - fourth order bore distortion of 25μm amplitude
Case 4 - a typical engine bore distortion composed of various orders

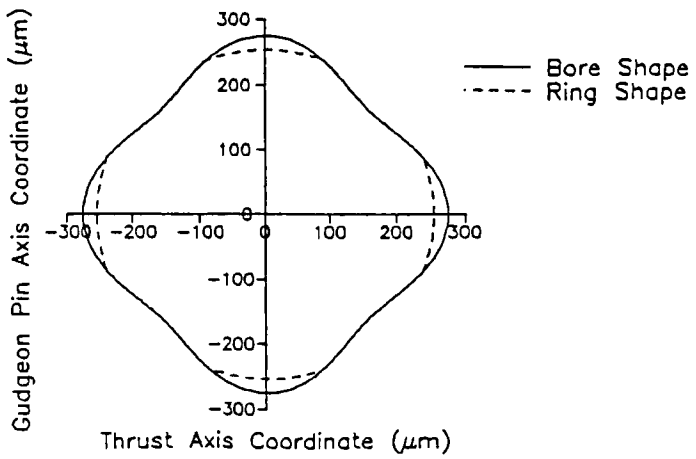
The results are shown in Figures 2.2 to 2.5 with the cylinder bore and piston ring distortions superimposed on a reference circle of 250μm radius. With the second order distortion of Case 1 shown in Figure 2.2, the piston ring conforms completely to the cylinder bore. The third order distortion of Case 2 shown in Figure 2.3, which is of the same amplitude, produces three significant regions where the piston ring cannot follow the distortion of the cylinder bore. These zones will be referred to as non-conformed regions. The maximum gap size between the piston ring and the cylinder bore in the non-conformed regions is about 10μm. Case 3 shown in Figure 2.4 exhibits four non-conformed regions with a maximum gap size of about 20μm. So it can be seen that as the order of bore distortion increases there is a consequent increase in the extent of non-conformity. A similar increase in non-conformity is obtained with increasing amplitude of bore distortion. Case 4 shown in Figure 2.5 is a combination of second, third and fourth order bore distortions and exhibits a much more complicated ring geometry with five non-conformed regions. This is a good illustration of the power of Fourier analysis in this context.

The non-conformed regions have essentially no radial load capacity as there will be very little hydrodynamic pressure in the relatively thick oil films between the piston ring and the cylinder bore. This imposes greater load on the remainder of the piston ring circumference. There is also the possibility of the flow of combustion chamber gases within the ring pack displacing the low pressure oil in the non-conformed regions. Bore distortion therefore has very serious implications for oil flow and blow-by in the piston assembly.

With gas pressure acting on the piston ring, conformity is greatly increased but not necessarily eliminated. In addition, gas pressures are only of significant magnitude for a small proportion of the engine cycle for a particular piston ring.

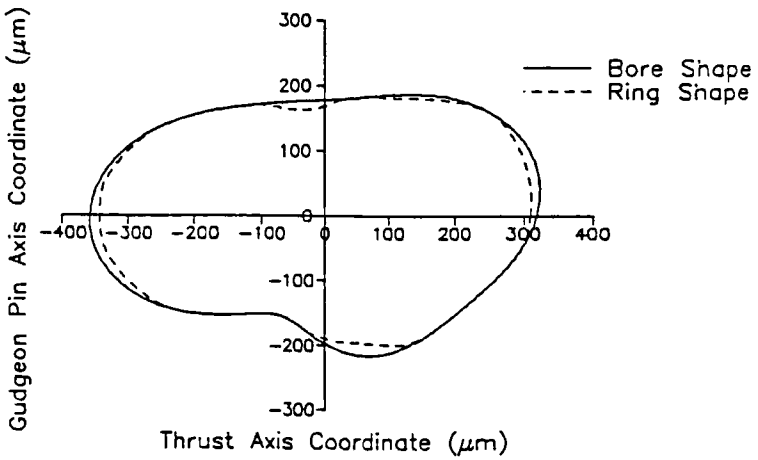
2.6 Implications for Engine Design

If the piston ring pack is unable to conform to the distorted cylinder bore there are only two options. Either change the design of the piston rings, to increase conformity, or reduce bore distortion. Equation (2.6) suggests that changing the geometry of the ring or increasing its inherent tension will improve conformity. Ring geometry is often determined by other design considerations, however, and increasing ring tension has the undesirable effect of increasing friction. Conformability can also be increased by making the structure of the ring more flexible as has been achieved in recent years with oil-control ring designs. The main possibilities for reducing bore distortion are improved assembly and minimising thermal distortion via effective cylinder cooling and controlled interaction between the cylinder and adjacent components such as the cylinder head gasket (Loenne and Ziemba, 1988).



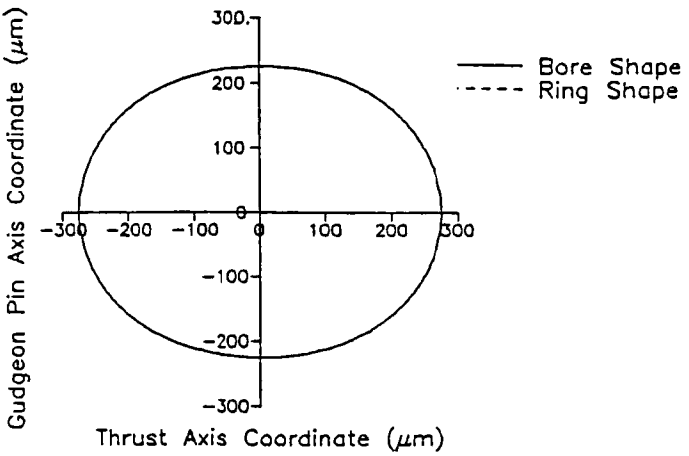
Piston ring conformity to a second order bore distortion of 25μm amplitude (Fourier amplitude constant A_2 = 25μm). The bore and ring shapes are superimposed on a reference circle of 250μm radius.

Figure 2.2 Case 1



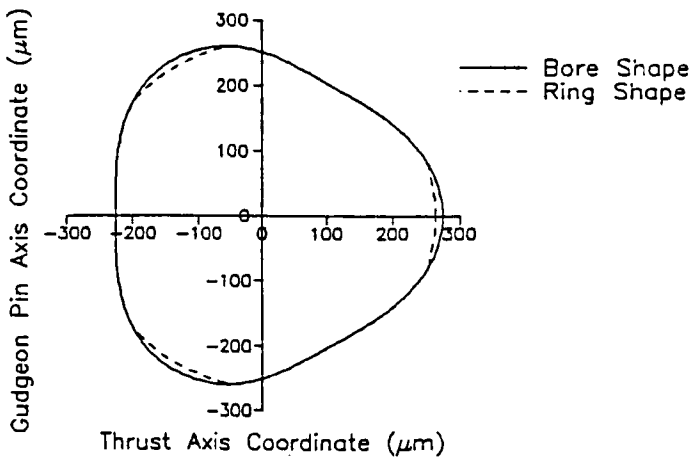
Piston ring conformity to a third order bore distortion of 25μm amplitude (Fourier amplitude constant A_3 = 25μm). The bore and ring shapes are superimposed on a reference circle of 250 μm radius.

Figure 2.3 Case 2



Piston ring conformity to a fourth order bore distortion of 25μm amplitude (Fourier amplitude constant A_4 = 25μm). The bore and ring shapes are superimposed on a reference circle of 250 μm radius.

Figure 2.4 Case 3



Piston ring conformity to a multiple order bore distortion (Fourier amplitude constants A_2 = +75μm; A_3 = -20μm; B_3 = +10μm; A_4 = +13μm; B_4 = +13μm). The bore and ring shapes are superimposed on a reference circle of 250μm radius.

Figure 2.5 Case 4

3. FUTURE DEVELOPMENTS

3.1 Introduction

The future directions in the design of piston assemblies for automotive use are likely to be guided by the requirement for fuel efficiency and the legislative controls upon exhaust emissions. The technology employed to meet these challenges will require new, and additional, information to be made available to the engine designer.

These developments may be considered in two different ways. Firstly, there is a need to relax some of the restrictions implicit in the analysis presented so far. A number of these will be indicated below as, dependent upon the operating conditions, they may have a noticeable effect upon the results obtained. Secondly, the more general view may be taken which considers materials, the manufacturing process and other changes in the design of the component parts.

3.2 Lubricant Viscosity

The majority of information regarding the viscosity of the lubricant within the piston assembly is determined from measurements made in a controlled environment and at known temperatures. Typically, oil extracted from the engine sump is taken to be representative of that used in the lubrication of the piston assembly. Clearly this does not give a true picture.

The hostile environment within the engine has generally precluded any measurement of the actual lubricant condition in the vicinity of the ring pack and piston skirt. Such information is likely to become known in the next few years and will be of considerable benefit when included in lubrication models of the piston assembly.

3.3 The Influence of Pressure

It is known that in some situations the influence of pressure can cause the lubricant viscosity to increase many fold. This piezoviscous effect can play an important part in maintaining a lubricant film through severe conditions. The added complexity of local elastic distortion resulting in conditions termed '*elastohydrodynamic*' lubrication has been noted as being of importance, Dowson, Ruddy and Economou (1983). The problems associated with the determination of the parameters used to represent this effect are identical to those of viscosity in the piston assembly.

3.4 Non-Newtonian Lubricants

Another of the difficulties associated with the analysis of the piston assembly is that lubricants may exhibit non-Newtonian effects. The is likely to be of particular importance in the study of the operation of ring packs where the lubricant films are small and hence the shear rates are high. Non-Newtonian behaviour is also likely when synthetic engine oils or multigrade lubricants with viscosity index (VI) improvers are used.

The consequence of using a non- Newtonian fluid are complex. In general a reduction in the lubricant films may be expected, all other factors being equal, but this may be balanced by other changes such as temperature.

3.5 Boundary Films and Lubricant Additives

Additive packages are common in most engine oils and serve a number of purposes such as to prevent foaming or to provide boundary lubricants. Their influence upon the analyses presented above is difficult to quantify particularly when the *'cocktail'* present contains many components. This problem is compounded due to the depletion of the additives with time and with the severity of engine operation.

In cases of true hydrodynamic operation their effect is likely to be small but where mixed lubrication is present they may influence the shear strength of any surface films developed. This is true, however, only for direct effects. The presence of an additive may lead to the formation of a boundary film which in turn will affect the wear of a particular component. This will then have an influence upon any hydrodynamic action and hence the two may be seen to be linked indirectly.

3.6 The Influence of Temperature

The influence of temperature upon the various components of the piston assembly is an important effect which must be accounted for in the analysis of engine components. Existing analytical techniques are capable of dealing with this problem and could be integrated with existing lubrication analyses. The additional level of complexity which would result, however, is usually regarded as being unacceptably involved.

At the heart of the problem is the determination of the surface temperatures of the various components. If these were to be determined separately, by finite element analysis for example, they could then be used as inputs to the main lubrication analysis. This would allow a much improved solution to be obtained without the need for repetitive iterations in order to obtain convergence through an engine cycle.

3.7 Deformation of the Piston Assembly

In the models of both the piston rings and the piston skirt it has been assumed that no deformation of the bodies takes place. This not only simplifies the problem but also removes some of the potential problems that would arise were local deformation of the components to take place (see Section 2, "Bore Distortion"). The ability to undertake the first steps in the full analysis is not in question (Dowson, Ruddy and Economou (1983), though the benefits of extending this further are not yet clear cut.

3.8 Wear of the Contacting Bodies

One of the well known features of engine tribology is that of "running-in". The initial surface profiles of components such as piston rings, cylinder liners, and to a lesser extent pistons, are modified by the removal of material. The modelling of this process is still at an early stage but future developments should eventually allow the shape of components to be correctly determined after a specified number of engine hours. This in turn will feed through into the hydrodynamic analyses where correctly defined lubricant film thicknesses will also help improve the accuracy of the end results.

3.9 Transient Effects

In all of the work considered in this section it has been assumed that the engine is operating at a steady state. This is a particularly gross assumption as the engine speed in most road vehicles is constantly changing. The resulting variation is surface velocity will influence some components more than others, and obviously the magnitude of any change can vary considerably.

3.10 Ceramic Components

One of the recent trends in the design of automotive engines has been the move towards the use of ceramics. These have high durability but are brittle when manufactured as solid components. The use of coatings is therefore seen as the most probable method of their introduction. Current knowledge suggests that existing additive packages are unlikely to prove satisfactory and that a new class of oil additives is therefore likely to be required. In addition, the influence of a ceramic faced component upon the counterface material must be considered. The *'running- in'* of the piston assembly is likely to be significantly different from that in a more conventional engine, and those wear products that are produced (both particulate and chemical) may lead to distress in the oil circulating system.

From an analytical point of view the current modelling ability is probably inferior to that of the more standard metal components. The wear of a ceramic is still not well understood as is its lubrication in the mixed/ boundary regimes. It is to be hoped, however, that research currently being undertaken will provide some of the answers to these uncertainties.

3.11 The Influence of Surface Texture

In the preceding sections it has been assumed that the rings, piston skirt and cylinder liner have perfectly smooth surfaces. This is clearly not realistic and better models, Rhode (1980), Patir and Cheng (1978) and Greenwood and Tripp (1971), can provide some means of investigating the effects of different surface textures. The ability of the currently available models to adequately represent the surfaces of the components in the piston assembly has, however, been questioned. The ability of surfaces, particularly the cylinder liner, to hold a lubricant may be important as may the local deformation of the surface features and their production by various finishing techniques, eg. plateau honing.

3.12 Blow-by, oil consumption and emissions

The increasingly tight requirements being placed upon engine manufacturers in respect of pollutant levels has lead to a shift in attention from friction reduction to emissions control. The development of a new engine is a very expensive exercise and hence it is particularly important to minimise the possibility of errors in the design at an early stage. Historically this process has often involved a progressive development supported by considerable experimental evaluation on engine test beds. Rapid changes in engine technology have, however, resulted in a greater reliance being placed upon analytical studies at an early design stage.

The modelling of piston assemblies has a known track record in the prediction of friction with reasonable agreement being obtained for this and the associated values of lubricant film thickness. The implicit requirement for accurate prediction of the inter- ring gas pressures can therefore be used to allow the blow-by of combustion chamber gases to be investigated. The situation regarding oil consumption and its influence upon exhaust emissions is considerably more complex. Very little experimental evidence is available regarding the oil flows around either the rings or the skirt. The situation regarding oil deposited on the liner walls after top dead centre firing is equally uncertain and hence any calculations should be viewed with caution unless validated against a large number of experimental measurements. A situation unlikely to exist when developing a new engine.

REFERENCES

Azevedo, M.N., and Filho, S.C., (1988), 'Lightweight Pistons.', SAE Paper 880729

Bishop, G.R., and Leavitt, A.H., (1975), 'Performance Simulation of a Diesel Piston and Ring System.', SAE Paper 750786

Brombolich, L.J. and McCormick, H.E., (1988), 'Tribological aspects of cylinder kit systems in adiabatic engines', SAE Paper 880018.

Daskivich, R.A., (1980), 'Sensitivity of new-engine oil economy to cylinder bore characteristics', Surface Roughness Effects in Hydrodynamic and Mixed Lubrication, ed. S.M. Rohde and H.S. Cheng, ASME pub., pp. 1-17.

Dowson, D., Ruddy, B.L., and Economou, P.N., (1983), 'The Elastohydrodynamic Lubrication of Piston Rings', Proc. R. Soc., Lond, **A386**, pp 409-430

Dunaevsky, V.V., (1990), 'Analysis of distortions of cylinders and conformability of piston rings', STLE Tribology Trans., **33**(1), pp 33-40.

Fesser, R.C., (1988), 'Small engine piston ring design for reduced oil consumption', SAE Paper 881327.

Frisch, S., (1988), 'Analysis of Heavy-Duty Diesel Piston Including Material, Air Gap, and Thermal Barrier Coating', SAE Paper 880671

Green, A.B., (1969), 'Inter Visual Studies of Piston-Cylinder Dynamic Oil Film Behaviour.', Wear, **13**, pp 345-360

Greenwood, J.A., and Tripp, J.H., (1971), 'The Contact of Two Nominally Flat Surfaces', Proc. Inst. Mech. Engrs., **185**, pp 625- 633.

Hill, S.H. and Newman, B.A., (1984), 'Piston ring designs for reduced friction', SAE Paper 841222.

Kaiser, H.J., Schmillen, K., and Spessert, B., (1988), 'Acoustical Optimisation of the Piston Slap by Combination of Computing and Experiments', SAE Paper 880100

Knoll, G.D., and Peeken, H.J., (1982), 'Hydrodynamic Lubrication of Piston Skirts', Trans. ASME, J. Lub. Tech., **104**(4), pp 504-509

Li, D.F., Rhode, S.M., and Ezzat, H.A., (1982), 'An Automotive Piston Lubrication Model.', ASLE Trans., **26**(2), pp 151-160

Loenne, K. and Ziemba, R., (1988), 'The Goetze cylinder distortion measurement system and the possibilities of reducing cylinder distortions', SAE Paper 880142.

Meir, A., (1952), 'Zur Kinematik der Kolbengerausche.' ATZ, **54**(6), June, pp 124-128

Mierbach, A., Dueck, G.E. and Newman, B.A., (1983), 'Heat flow through piston rings and its influence on shape', SAE Paper 831283.

Munroe, R. and Parker A., (1975), 'Transverse Movement Analysis and its Influence on Diesel Piston Design.', SAE Paper 750800

Rhode, S.M., (1980), 'A Mixed Friction Model for Dynamically Loaded Contacts with Application to Piston Ring Lubrication', Proceedings of the 7th Leeds/ Lyon Symposium on Tribology, pub. Westbury House

Oh, K.P., Li, C.H., and Goenka, P.K., (1987), 'Elastohydrodynamic Lubrication of Piston Skirts.', Trans. ASME, J. Trib., **156**(2), pp 356-362

Patir, N., and Cheng, H.S., (1978), 'An Average Flow Model for Determining the Effects of Three Dimensional Roughness on Partial Hydrodynamic Lubrication', Trans ASME, J. Lub. Tech. **100**, pp 12-17

Takiguchi, M., Machida, K., and Furuhama, S., (1988), 'Piston Friction Force of a Small High Speed Gasoline Engine.', Trans. ASME, J. Lub. Tech., **110**(1), pp 112-118

Engine Tribology / C.M. Taylor (Editor)

SOLID MATERIALS

K. Holmes
Leyland Daf Technical Centre, Leyland, PR5 1SN

1. INTRODUCTION

In previous chapters you have read about many of the factors affecting the tribology of engines. At this stage, therefore, readers should have a good background in the lubrication regimes which exist in different parts of an engine, how these parts of the engine are designed in order to minimise friction and wear, and how different types of wear occur in engines.

What happens in all the major moving parts of an engine and how to analyse their lubrication or dry contact should now be familiar to you. However, in addition to using the design analysis available these days, the choice of material for the various moving parts of an engine is equally important. The pressures on Engine Designers and Development Engineers these days is more intense than ever, and a knowledge of the materials available to complement modern lubricants and analysis techniques is vital. For example, what could be worse than having the best cam and tappet contact conditions and a high quality lubricant, only to find that the engine seizes up due to a basic material incompatibility between the two parts.

Within this chapter, details of materials available for various engine parts will be described, and new developments in the field of engine materials will also be introduced for possible future use. Finally, details of how to analyse engine failures due to tribological problems will be discussed, as this aspect is vital to the Designer and Development Engineer alike.

2. ENGINE DEVELOPMENT TRENDS

2.1 History of Engine Materials

The internal combustion engine has now been with us for over 100 years. Much of the technology of engine materials in the early years of I.C. engines came from steam engines, as they had similar moving parts and the need for good lubrication in a hot, hostile environment. Due to the reciprocating engine's use in such widely varying applications ranging from lawn mowers, through cars, trucks, aircraft, to large stationary power generators and

ships, different materials have suited the different applications. For example, the widely different cost base between a mass produced Austin 7 car and a Rolls-Royce Merlin-engined Spitfire led to different solutions to the materials used in their design. The need for light weight for an aircraft engine also played a part!

However, in the modern era of reciprocating engines, gas turbines have taken over in some areas and electrical power is becoming more pronounced in other areas. Therefore, the I.C. engine is generally used in a narrower field of application than, say, 30 years ago, and the engine materials used have also narrowed. In addition, materials technology from gas turbines is now being developed for reciprocating engines.

To give an example of how materials in engines have changed we can look back at main bearings. In the past, many main bearings were actually cast in situ, using white metal (lead-based or tin-based). Nowadays, these materials would not withstand the bearing pressures or higher temperatures in an engine and, therefore, copper-lead or aluminium-tin alloys are generally used in the form of a steel-backed bearing shell. Modern machining techniques have also helped to make this possible.

2.2 Current Trends in Engine Design

In understanding the materials used in the moving parts of engines nowadays, some of the trends in engine design and performance need to be analysed. In order to do this, it will help to classify different I.C. engine applications, as below.

(A) Car engines (petrol and diesel)
(B) Commercial vehicle engines (diesel)
(C) Small stationary engines (petrol and diesel)
(D) Large stationary engines (diesel).

The following matrix outlines the influences on these various types of engine, due to legislative, technical and commercial pressures.

	A	B	C	D
Exhaust Emission legislation	✓	✓		
Noise legislation		✓	✓	
Environmental pressure groups	✓	✓	✓	
Higher powers	✓	✓		✓
Better fuel economy	✓	✓		✓
Reduction in fuel quality	✓	✓	✓	✓
Reduced maintenance		✓	✓	
Lower manufacturing cost	✓	✓	✓	
Competition from other power sources			✓	✓

These needs have led to changes in engines as follows:

(i) Improved combustion to give better fuel economy and lower emissions means higher cylinder pressures and higher temperatures.
(ii) Lower emissions can be helped by lower oil consumption, which generally means less lubricant in high temperature areas.
(iii) Increased use of turbocharging has led to higher temperatures within engines, and higher oil temperatures.
(iv) Lower engine speeds (for fuel economy and noise reasons) in diesel engines means less "hydrodynamic" effect.
(v) Lower sulphur fuels mean less corrosion effects, but also less boundary lubricant effect.
(vi) Unleaded petrol means reduced "dry lubricant" effect on valves and rings.
(vii) Trend to lighter engines has increased the use of aluminium, which is tribologically worse than cast iron.

There have been other influences on engine design, but it can be seen already that changes to traditional engine materials have been made necessary due to the rapid changes in engine performance.

3. LUBRICATION MECHANISMS AND MATERIAL TYPES

Before moving on to details of specific parts of the

engine, it is wise to recap on the wear and lubrication mechanisms possible, and how they influence material type.

3.1 Dry Contact

Dry contact can occur in parts of the engine due to high loads and/or high temperatures causing lubricant film breakdown. In addition, start-up from cold and initial running-in of an engine are both cases where dry contact is likely to be present.

In these cases, it is important that the two contacting materials are compatible in terms of adhesive wear. Differences in hardness between the two materials also need to be carefully watched in order to minimise abrasive wear.

3.2 Hydrodynamic Lubrication

During hydrodynamic lubrication, there should be no contact between surfaces, therefore material choice should be less vital. However, there are other considerations such as corrosive influences due to the effect of a hot lubricant. Erosion of the materials can also be a problem, particularly if small solid particles are present in the lubricant flow.

3.3 Elasto Hydrodynamic Lubrication

In the case of e.h.l., both surfaces need to be able to withstand the extreme pressures present and, therefore, high strength and good fatigue resistance are necessary properties. One exception to this is the case of rotary oil seals where the rubber or polymeric seal usually operates in the e.h.l. mode. The soft material does need good fatigue, and compatibility with the counterface against which it runs.

3.4 Material Types

Some of the basic reasons for engine material choice have been summarised here, and it is now appropriate to consider some of the materials used in engines.

Generally, there are two types of contact in engine surfaces. One is where there is a large area of contact, such as would be found in a main or big end bearing, or between piston and liner. In this case it is usual to have a hard-soft combination of materials, which will allow for limited initial wear of the softer material, during running-in of the engine, or to accept slight tolerance misalignments.

Secondly, there are small areas of contact, such as a line or point contact on a cam/tappet or a gear pump. Here, it

is usual to use a hard-hard combination of materials, very often metals which have been heat treated or otherwise processed to give a hard surface. One important consideration here is that the materials used should be tribologically compatible. For example, a chrome-plated surface rubbing against a similar surface would give a hard-hard combination, but they would seize up after a short time due to their ability to "fuse" together easily. It would be better to have two dissimilar hard surfaces. One other problem to consider with hard surfaces is that they also produce hard wear particles, which can then cause damage to other parts of the engine.

Materials which are used for tribological applications usually are made up of hard and soft components. The hard element is there to give strength and the soft element is there to provide conformity to misalignment and rubbing. Therefore, a simple rule to adopt in choosing materials is to think about the balance needed between strength and conformability and to initially look at the component strength needed.

Cost of materials is also a critical factor in many industries, and this is where relatively cheap materials which can be treated to improve their wear resistance are used, e.g. induction hardened steel for crankshafts.

At this stage it is appropriate to understand typical wear rates and friction values for various materials, as shown in the tables below.

Material	Specific wear rate
Aluminium	10^{-10} m^3/Nm
Mild steel	10^{-12} m^3/Nm
Hard steel	10^{-14} m^3/Nm
Tungsten carbide	10^{-16} m^3/Nm
Polyethylene	10^{-16} m^3/Nm

The units used here reflect wear volume per sliding distance and load, but the critical factor is the great variability in wear between materials, e.g. hard steel wears at 1/100th the rate of mild steel.

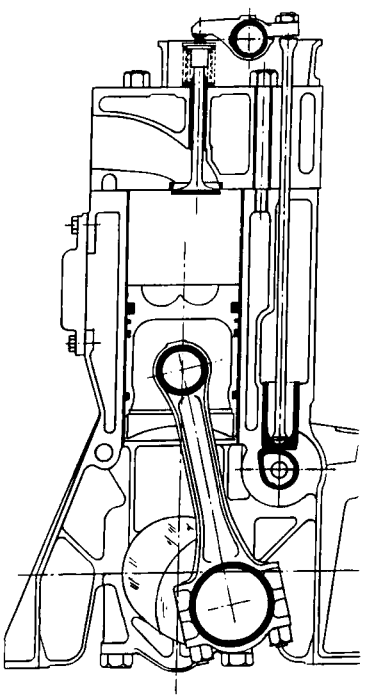
Material	Friction coefficient
Aluminium	1.0
Stainless steel	1.0
Mild steel	0.3 - 0.7
PTFE	0.1
Lubricated situation	< 0.05

Friction coefficient is the ratio of the force needed to slide one surface over another, and the normally applied load.

Unfortunately, friction and wear are not quite as simple as this, as these results are quoted under standard test rig conditions. If, for example, the results were measured at a high temperature or higher pressure, then the wear of some materials would increase dramatically due to local melting. Friction can also alter significantly with changing conditions.

4. APPLICATION OF MATERIALS TO ENGINES

Looking now at the wearing surfaces of an engine, as shown in Fig. 1, each contact is different and therefore needs different considerations for choice of materials. The following sections outline materials used for particular parts of the engine.



4.1 Pistons, Piston Rings and Cylinder Bores

The piston-ring-cylinder bore contact is a complex tribological situation where it is generally assumed that the contact is lubricated and that the piston and ring run on a film of oil. However, theoretical analysis of piston rings now shows that oil films break down under certain

conditions in the engine, such as at the ring reversal areas and during initial start-up. Increasing engine temperatures and cylinder pressures, combined with the need for lower oil consumption, have exaggerated the times when oil films may break down. Therefore, the requirements for materials in this area can be summarised as below:

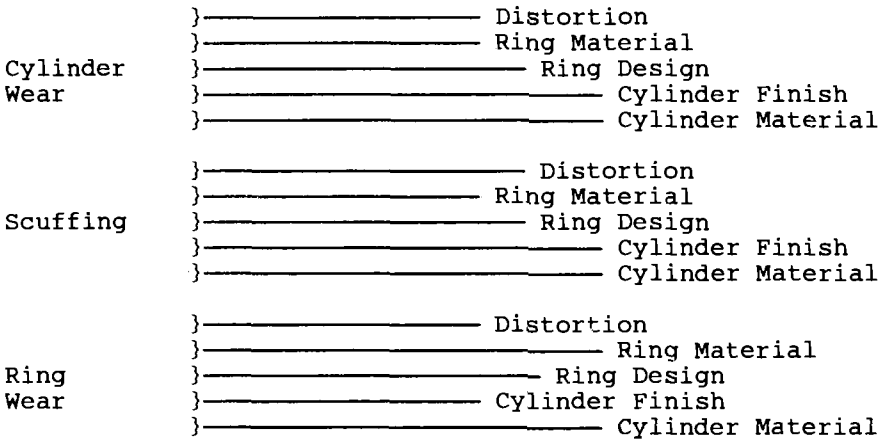
Pistons	Piston Rings	Cylinders
Light weight.	Ductility.	Low cost.
Easy to machine.	Resistance to	Machinability.
Low cost.	dry contact.	Compatible with
Strength at high	Low friction.	ring material.
temperature.	Easy to machine.	Low wear.
	Low wear.	Good heat transfer.

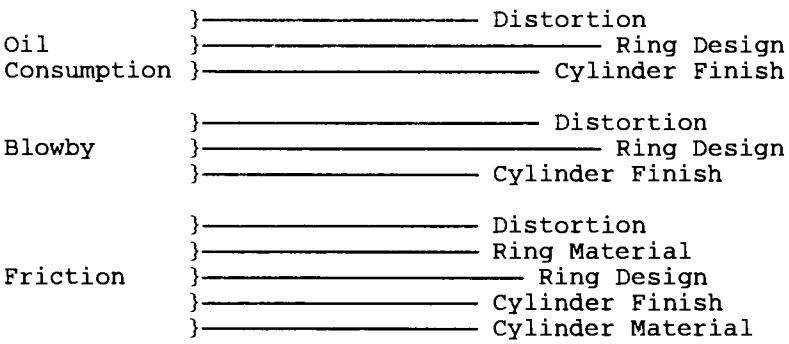
The contact regions are generally large, but areas such as oil control rings have line contacts. Traditionally, engines have used cast aluminium pistons and cast iron cylinder bores. Piston rings are usually cast iron, treated with chromium to reduce wear.

However, modern engines, depending on their type, could have aluminium cylinders, cast iron pistons, and advanced coatings on the piston rings.

Depending on the severity of engine usage and degree of engine performance, the degree of wear resistance will vary.

Fig. 2 illustrates the relative importance of material choice for cylinders and piston rings compared with other factors such as ring pack design. As can be seen, cylinder material particularly affects wear and scuffing resistance; ring material affects ring wear principally.





For cast iron cylinder liners or cylinder bores, normal casting grade of iron will be the lowest performance available. For engines of higher performance, cast irons of greater hardness will be needed, but the level of free ferrite at the rubbing surface needs to be limited during manufacture. The grade of iron could vary from 3% Carbon content with 0.2% chromium and 0.5% Nickel. Stronger, more severe duty cylinder liners would have a higher nickel content and contain 0.5% Molybdenum. Phosphorus content can also be used to give varying grades of wear resistance. For aluminium bores, wear performance can be improved by using a cast iron cylinder liner, or by the use of a chrome-based or nickel-based surface coating. Most companies will have literature available, or their own standards, for materials used in engines. These, and the literature from company suppliers, should be used to give indication of properties.

Modern engines are increasingly looking for lower friction and lower oil consumption, and materials or treatments are being developed to improve these areas in the piston/bore contact region. In some cases, nickel-based coatings are used in preference to chrome-based coatings to help lower friction. One step further is to incorporate low friction coefficient materials, such as PTFE, in the coatings. Care has to be taken, though, that the wear rate is not compromised.

Piston rings are usually made of cast iron, with higher levels of chromium and molybdenum (1%) to give more ductility. Wear resistance is also improved, but it is more common to apply a surface coating to the piston rings in reducing wear further. Materials used include chromium, molybdenum, or even copper, where high initial conformability may be needed. This has also led to the use of two-stage coatings on rings to combine wear resistance with good initial bedding-in of the engine.

Pistons are normally made of aluminium with a 10-20% silicon content and 1-3% nickel, the higher alloy content

giving higher strength at elevated temperatures. Some high rated engines use grey cast iron pistons, or even a combination of a cast iron piston crown, with an aluminium piston skirt. One further trend in the material used for pistons is the use of the squeeze-casting process which gives higher strength. It also enables ceramic piston bowl reinforcement to be incorporated in the design, this providing further high temperature strength in the hottest parts of the piston.

4.2 Crankshafts and Bearings

The main, big end and small end bearings in an engine are simpler to consider than the piston-cylinder contact. Bearings have relatively large contact areas during start-up, but the vast majority of their life is spent with a full hydrodynamic film between the bearing material and the crankshaft, or gudgeon pin.

Properties which are important for engine bearings are:

Conformability
Strength
Corrosion resistance (acidic and alkaline)
Erosion resistance
Cost
Low wear

Bearings nowadays are thin shells, usually with a steel backing layer to give strength and tighter tolerancing. The steel shell is overlaid with a bearing material, usually tin-aluminium or lead-bronze, the choice of material and the thickness of this layer determining the fatigue strength. A further thin overlay is then put onto the bearing material, this being softer to provide a degree of initial conformability. The following table gives typical properties for a range of engine bearing materials.

Material	Elemental Composition	Fatigue Strength	Brinell Hardness
White metal	90%Sn, 7%Sb, 3%Cu	14-17 MN/m^2	20-30
	60%Cu, 40%Pb	20 MN/m^2	30-40
Lead Bronze	80%Cu, 10%Pb, 10%Sn	45 MN/m^2	80-100
Tin-Aluminium	80%Al, 19%Sn, 1%Cu	32 MN/m^2	30-40
	93%Al, 5%Sn, 2%Ni	45 MN/m^2	60-80

Where possible, Tin-Aluminium is used as it is cheaper than other materials and is more resistance to corrosion, but it

tends to seize up more easily, hence the use of a soft overlay to provide conformability. Lead bronze has good fatigue strength and erosion resistance, but is more susceptible to corrosion by weak acids, again illustrating the need for a corrosion resistance overlay. Guidance is given below, where *** means most suitable and * means least suitable.

	Load Capacity	Resists Seizure	Resists Corrosion	Shape Tolerance	Max. Temp.
Pb White Metal	*	***	**	***	130°C
Sn White Metal	*	***	***	***	130°C
Pb - Bronze	**	*	**	**	180°C
Sn - Al (+Cu)	**	*	**	*	180°C
Sn - Al + overlay	**	**	***	**	170°C
Cu - Pb + overlay	**	**	***	**	170°C
Ph. Bronze	***	*	**	**	220°C

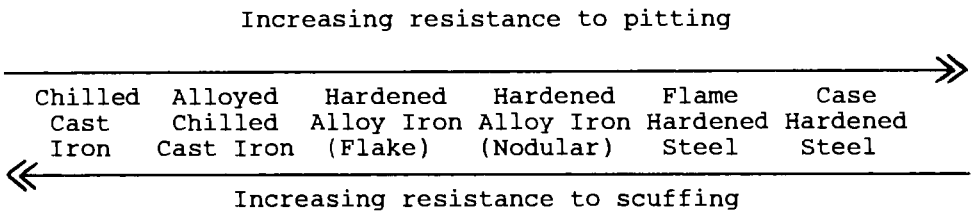
Shafts designed for running against engine bearings are made of heat treated steels or spheroidal graphite irons. The shaft should be approximately three times the hardness of the bearing material, but shaft fatigue strength is also a major consideration. In general, shafts for light and medium-duty engines can be induction or flame hardened, but for heavy-duty engines, nitriding is used to provide the necessary strength. Cost again is a consideration here, as the nitriding process gives a harder shaft, which is more costly to machine to the correct size and profile.

4.3 Valve Train Components

Valve train components cover a wide range, such as rocker shafts, reciprocating bearings, rocker ends, and cams and tappets. In particular, cams and tappets tend to cause the most problems in modern engines. The trend towards overhead camshaft engines and complex valve geometries, coupled to cost reductions in design and materials have led to the increasing problems in the valve train region.

Cams, tappets and rocker ends are line or point contact regions, usually with limited lubrication, and therefore need a hard-hard material combination. Due to these parts of the engine being at the end of the oil supply line, very little oil is available during start-up and therefore materials need to be scuff resistant. Pitting resistance is also important in order to cope with the high fatigue

pressures present. It is usual, therefore, to use chilled cast irons or hardened steel combinations in cams and tappets, and the diagram below illustrates the pitting and scuffing resistance of these materials.



Normally the tappet is made to be more resistant to pitting, and the cam to be more scuff resistant. Experimental work, coupled to past experience, will guide the Engine Designer in choice of materials in this area.

New developments in valve train materials include the use of ceramics as tappets, the best material under test conditions so far being a mix of aluminium oxide and silicon nitride. Coatings of titanium nitride are also being tried for their wear resistance in valve train components.

One very advanced development which could become normal is to have a "dry" top end to the engine, obviating the need for lubricant on the valve train components. The benefits of this are reduced engine complexity and need for gaskets, hence reliability should improve, but the penalties are increased friction and the cost of the materials needed. Further work is needed to prove the concept and to reduce material costs.

5. MATERIAL ADVANCES

Developments in the areas of ceramics, plastics and surface coatings could revolutionise engine materials over the next 20 years, although the general feeling is that the process will be very gradual.

For example, we are beginning to see structural plastics being considered for engines, but not for moving parts, therefore the engine could become a complex mixture of materials, some parts specifically developed for strength, some for wear resistance, some for light weight. This leads to problems with joining techniques for dissimilar materials, and to design difficulties such as the simulation of material interfaces in Finite Element programs.

Ceramics in their present state of development can be used for localised piston reinforcement, small turbocharger rotors, and for severe wear surfaces. Larger scale use of ceramics will depend on improved consistency of the material, allied to better methods of non-destructive testing of the material.

Research work also exists for the development of surface coatings and treatments for engines. Here again, much will depend upon the consistency with which materials can be coated, and the development of non-destructive test methods. Although coatings are already in use in engines to a significant degree, future efforts will concentrate on ceramic coatings, chemically-deposited coatings, and applications of plasma heat treatments.

6. FAILURE DIAGNOSIS

Despite the best intentions of using the correct material and optimising the design of tribological components in engines, it is inevitable that Designers and Development Technicians will be involved in the analysis of failures or the monitoring of wear in engines. Therefore, a knowledge of the analysis techniques and a study of common failures is essential.

6.1 Oil Analysis

A great deal of knowledge about the tribological condition of the engine can be found from a thorough study of the oil condition. Indeed, regular oil sampling and analysis can be used to predict wear rates or to study the feasibility of uprating engines.

The principal method of oil analysis is spectrographic oil analysis, where an oil sample is burnt and the concentration of particular metallic elements can be quantified by the absorption or emission of light frequencies in the flame. Below is a list of elements which can be analysed as an aid to engine wear diagnosis, together with typical wear limits for engines and possible sources of the wear metal. It must be stated, however, that different engines will give different maximum allowable concentrations, and experience or regular sampling is needed to give a complete picture of engine condition.

Element	Concentration Limit	Causes
Aluminium	40 ppm	Bearing wear, piston wear, dirt contamination.
Copper	40 ppm	Bearing wear, oil cooler corrosion
Lead	100 ppm	Bearing wear (Lead-Bronze)
Chromium	40 ppm	Piston ring wear, cylinder liner wear, coolant contamination.
Iron	100-120 ppm	Cylinder wear, tappet/ cam wear, crankshaft wear.
Tin	20 ppm	Bearing wear
Silicon	20-30 ppm	Cylinder wear, dirt contamination (can be present in some oils).
Sodium	50 ppm	Coolant contamination.

Experience and commonsense both help in analysis. For example, if regular analyses show a sudden change in contamination rate of one or two elements, that should be used to identify a potential failure.

Also, if tin and aluminium bearings are used in the engine, they will bias the results. If high values for both tin and aluminium are seen, then a bearing wear or fatigue problem should be suspected. Similarly, if high levels of iron and silicon were seen, it would be a sign of ring or cylinder wear due to dirt contamination. Also note that where a silicon carbide liner treatment is used, the silicon level would normally be higher than for a standard engine (up to 100 ppm).

6.2 Piston/Cylinder Failures

The most common piston and cylinder failures are complete seizure, and high wear giving high oil consumption.

Piston/bore seizures are often too catastrophic to the engine to enable analysis of the seized bore to be useful. However, it is helpful to analyse the condition of the other cylinders and pistons in the engine. Most information is often gained by looking at marking on the piston crown and skirt, and the cylinder bore. These will indicate where the seizure may occur, and thus give guidance as to the cause. Examples are given below:

Marking on the piston crown	- piston crown too large, cylinder bore distortion at the top, piston balance wrong.
Marking on the piston skirt	- lower cylinder bore distortion, piston skirt too large, piston balance wrong, engine timing wrong, dirty assembly.
Top piston ring marking	- Inadequate ring lubrication, incompatible material with cylinder bore, wear of ring in ring groove.
Cylinder bore marking at regular intervals around the bore	- Cylinder bore distortion due to high engine temperatures, inadequate cooling, high duty cycle.
Cylinder bore shiny in patches, slight scoring	- Bore polishing due to poor lubricant, or poor ring design.
Cylinder bore smooth and no wear on rings	- Bore glazing due to improper engine running-in.

6.3 Bearing/Shaft Failures

Although bearing seizures can take place due to lack of lubrication or overloading, other failure mechanisms such as material fatigue or corrosion are also common.

Visual examination of a worn or failed bearing, in conjunction with the shaft against which it has run, will often show the mode of failure. However, in some cases it may be necessary to use an optical microscope to show more detailed areas of damage. Finally, the use of an electron microscope can give a 3-D image of the failed region, and can also be used to look at the materials in discreet areas of the surfaces. For example, small particles embedded in a bearing can be analysed in detail, in order to determine the source of the particles. In a similar way, examination of a worn bearing can indicate if one of the metallic elements of the bearing has been leached from the surface by corrosion.

Some examples of bearing damage and their causes are given below.

Rough surface on the crankshaft	- corrosion, inadequate oil film or inadequate oil supply, poor machining.
Small particles embedded in the bearing	- dirt contamination, wear from other parts of engine causing damage.
Removal of bearing overlay locally (often gives 2-coloured bearing)	- initial misalignment of crank and bearing, excessive loading, overlay too thin.
Loss of bearing material in large pieces	- overloading, overheating due to excessive usage, wrong material selection.
Light and dark patches on back of bearing shell	- dirt on back of bearing during assembly, under-tightening or over-tightening of bearing cap.
V-shaped marks on bearing surface	- cavitation caused by poor design (often of position of oil feed), erosion of bearing by particles from crankshaft (nitrided crankshafts).
Pitting damage (small holes) on bearing surface	- corrosion, due to sulphur compounds or acids in lubricant, inadequate overlay protection of bearing.

6.4 Cam/Tappet Failures

Due to the lubrication mechanism of cams and tappets being the same as gears, i.e. elasto-hydrodynamic lubrication, the types of failure are the same as for gears. Failures can be categorised as surface pitting, wear, or scuffing (and possible seizure). Visual examination, or optical microscope examination is normally sufficient in analysing cam and tappet failures. On occasions examination of the wear debris in the engine can reveal more about the failure if more than one form of damage has occurred.

Typical failure features and causes are shown below.

Small holes in the surface, often joined to other, roughly triangular holes	- pitting damage caused by overloading of the material, poor material choice, design gives too high a contact stress.
Polished appearance to faces	- wear caused by inadequate oil films, chemical attack by lubricant, dirt particles in oil.
Scored appearance on cam and/ or tappet	- scuffing damage due to poor lubricant supply, overheating of oil, poor surface finish on cam or tappet, wrong material choice.
Scoring marks down the side of tappets	- inadequate clearance between tappet and engine block, dirt contamination on assembly.

As already stated, experience of failures plays a major part in understanding their causes, combined with using the correct analysis techniques. It is also useful, though, to study typical failure pictures, as can be found in ASTM reference books.

7. ACKNOWLEDGEMENTS

The author acknowledges the knowledge gained from members of two major initiatives on engine materials supported by the UK Government. These are the CARE (Ceramic Applications in Reciprocating Engines) and CAST (Coatings and Surface Treatments) projects.

Engine Tribology / C.M. Taylor (Editor)
1993 Elsevier Science Publishers B.V.

Engine lubricants

J.C. Bell

Shell Research Limited, Thornton Research Centre, P.O. Box 1, Chester CH1 3SH

1. INTRODUCTION

The subject of lubricants for engines is too large to be treated comprehensively in a single chapter. This chapter aims to give at least a superficial overview of the areas that are likely to be of greatest interest to the expected readership. The main emphasis is on crankcase oils for automotive engines, although much of it is also relevant to engines in stationary and marine applications. For the sake of brevity, several important topics have been omitted, together with much of the detail of those included. For those who wish to fill some of these gaps, Reference 1 contains a wealth of detailed information on the subject.

The lubrication of the engine appears frequently to have been an afterthought in the engine design process. It is encouraging to note that in recent years, perhaps as a result of unwanted surprises in the past, there has been a movement towards greater interaction between the designers of engines and components, and those who develop engine lubricants. It is hoped that this publication will be effective in promoting this trend.

2. FUNCTIONS OF AN ENGINE LUBRICANT

Lubricants for both gasoline and diesel engines need to fulfil several important functions in addition to the best known, that of controlling friction and wear. They protect the engine against rusting and other types of corrosion even before a vehicle goes into service. The lubricant is a primary medium for cooling the piston, where it also plays a part in completing the seal between the piston, piston-ring and cylinder surfaces against leakage of the combustion gases. In long-term use, it provides the only means of keeping the internal surfaces of the engine clean and in the shape in which they were designed to operate, by minimising the deleterious effects of combustion products and the many other contaminants that can enter the engine.

The lubricant must perform these functions for periods of hundreds of hours between oil changes, under a wide variety of climatic and operating conditions. High-speed driving, or earth moving and mining operations, under heavily loaded conditions can cause extreme thermal stressing and oxidation of the oil. During city driving or local delivery service, when the engine remains relatively cold, acidic combustion products and partially burned fuel dilute the oil and cause corrosion and sludge problems. Even in temperate climates it may occasionally be necessary to start the engine at sub-zero temperatures. The oil must remain fluid under all these conditions to perform its functions, ideally maintaining the right physical properties to perform them efficiently.

3. LUBRICANT COMPOSITION

The lubrication requirements of modern automotive engines cannot be met by simple mineral oils alone, or even by synthetic fluids. Both require the addition of a number of chemical additives to improve their performance in specific areas. The typical constituents of lubricants for gasoline and diesel engines are indicated in Table 1. The major component is the base fluid, which is usually a high quality, solvent-refined mineral oil. For certain high performance applications, e.g. high power output, long oil-change intervals, low starting temperatures, etc., it may be necessary to use more highly processed mineral oils or synthetic chemical products.

The base fluid is supplemented with typically 5 to 25 per cent of additives, which perform the wide variety of roles indicated in Table 1. The materials and their functions will be explained more fully in subsequent sections. There may be up to 10 to 15 different additives in a finished lubricant. In the formulation of lubricants, great care is taken to obtain the correct balance of properties, some of which conflict with others, often having to use more than one additive of each type to obtain the desired result.

Table 1: Typical ranges of composition of engine crankcase lubricants

Function	Component	Concentration, %m
Friction	Viscosity Index Improver	0 - 6
and	Anti-wear additive	0.5 - 2
Wear	Friction Reducer	0 - 2
	Rust/Corrosion Inhibitors	0 - 1
Contamination	Antioxidant	0 - 1
and	Detergent	1 - 10
Cleanliness	Ashless Dispersant	2 - 9
Maintaining	Pour Point Depressant	0 - 0.5
Fluid Properties	Anti-foam additive	0 - 0.001
Base oil (mineral and/or synthetic)		75 - 95

4. CONTROL OF FRICTION AND WEAR

Generally, it is desirable to minimise both friction and wear: the former for good fuel economy and cold starting, and the latter for long-term durability. This is the basis of what follows, but it should be remembered that there are instances in which low friction is undesirable, for example, when the engine oil is used to lubricate integral gearbox or tractor transmissions where synchromesh rings or clutch packs may require a high coefficient of friction.

4.1 Lubricant Viscosity

The most effective way of minimising wear is to ensure that the surfaces of rubbing components are separated by a thick oil film. Here, the viscosity of the oil plays a predominant role, whether in the hydrodynamic or elasto-hydrodynamic regime. Thus, high viscosity oils would be favoured simply to prevent wear. Unfortunately, this results in poor cold-starting and fuel economy because of the high friction associated with high viscosity oils.

This conflict is illustrated in Fig.1. Starting at very low viscosity (thin oil films) we have the high friction and wear associated with the boundary lubrication regime. As the viscosity increases, both friction and wear diminish as the system passes through the mixed (boundary/ hydrodynamic) regime until the point where significant boundary contact ceases. Friction then increases in a hydrodynamic manner with viscosity whilst wear may continue to slowly diminish.

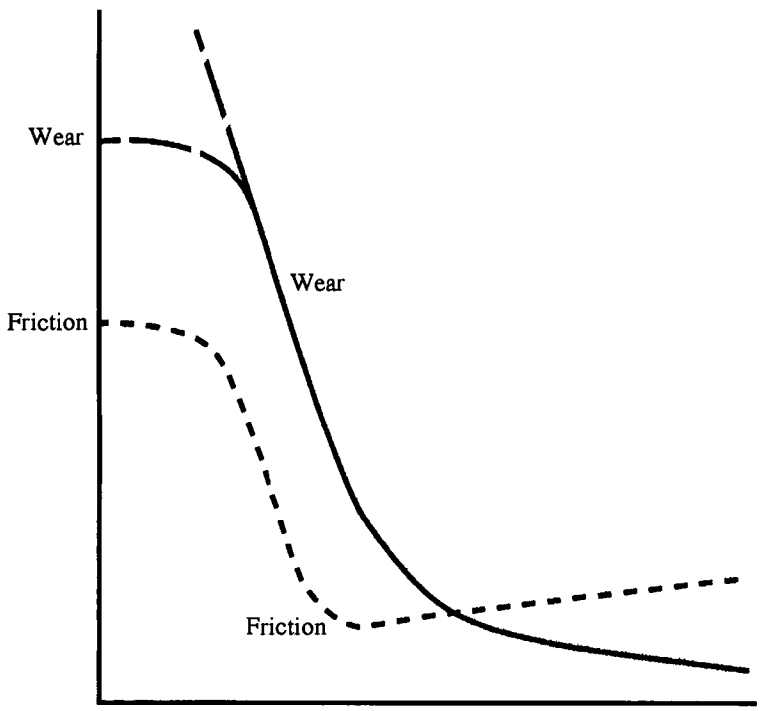


FIG. 1 - Schematic variation of friction and wear with oil viscosity - other conditions constant

The ideal oil film thickness would be the minimum that avoided boundary contact. However, because of the differing dynamic loading and velocity cycles experienced by all the principal lubricated contacts, e.g. bearings, piston/ring/cylinder and cam/follower, the optimum viscosity for each component varies throughout the cycle, and the range is not the

same for the different components. Thus, there is no single optimum viscosity for an engine, even when operating at constant conditions, and even less so under the usual wide range of conditions experienced in normal service.

Energy losses due to friction are cumulative over all components and over the period of operation, whereas the life of an engine can be seriously reduced because of the wear of a single component, possibly in a very small number of events. To maintain an adequate safety margin, oil viscosities substantially greater than the optimum for low friction have been recommended in the past. Recently, there has been a move to lower viscosities to obtain fuel economy benefits. This has been made possible by improvements both in engine technology and in our understanding of the lubricant factors that affect the balance between friction and wear.

The high sensitivity of the viscosity of oils to temperature presents a significant challenge. The viscosity of a high-quality paraffinic mineral oil decreases by a factor of about 400 between -5°C, a moderate cold-starting temperature, and 120°C, a fairly typical sump oil temperature. The temperature of the oil in the main and big-end bearings can be much higher, and that in the little end and around the piston rings higher again. In contrast, in North America and Scandinavia the ability to start at temperatures as low as -30°C may also be required. To facilitate cold starting and avoid engine seizures due to poor oil pumping, the SAE J300 engine oil viscosity classification (Table 2) includes requirements for cold cranking resistance (ASTM D 2602) and borderline pumping characteristics (ASTM D 4684).

Table 2: SAE J300 March 1991 version. Engine oil viscosity classification[a]

SAE Viscosity Grade	Low-Temperature (°C) Viscosities		Viscosity[c] (cSt) at 100°C	
	Cranking[b](cP) max.	Pumping[c] (cP) max. with No Yield Stress	min	max
0W	3250 at -30	30 000 at -35	3.8	-
5W	3500 at -25	30 000 at -30	3.8	-
10W	3500 at -20	30 000 at -25	4.1	-
15W	3500 at -15	30 000 at -20	5.6	-
20W	4500 at -10	30 000 at -15	5.6	-
25W	6000 at -5	30 000 at -10	9.3	-
20	-	-	5.6	<9.3
30	-	-	9.3	<12.5
40	-	-	12.5	<16.3
50	-	-	16.3	<21.9
60	-	-	21.9	<26.1

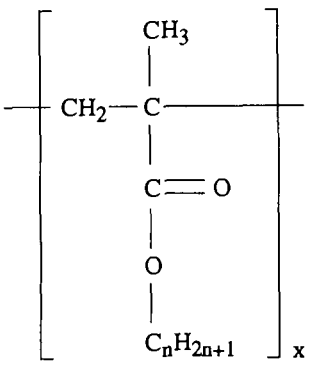
Note: 1 cP = 1 mPa.s; 1cSt = 1 mm^2s^{-1}

(a) All values are critical specifications as defined by ASTM D3244.
(b) ASTM D 2602 (multi temperature version cold cranking simulator).
(c) ASTM D 4684 (mini rotary viscometer TP-1) **Note:** The presence of any yield stress detectable by this method constitutes a failure regardless of viscosity.
(d) ASTM D 445 (capillary viscometer).

To reduce this temperature sensitivity, high-molecular-weight polymers, known as Viscosity-Index Improvers are added to the oil. (Viscosity Index, or VI, is an empirical index

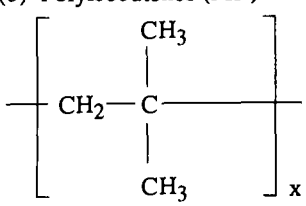
of viscosity-temperature sensitivity, higher values indicating lower sensitivity). Representative chemical structures of the most widely used classes of VI Improvers are shown in Fig.2. The

(a) Polymethacrylates (PMA)



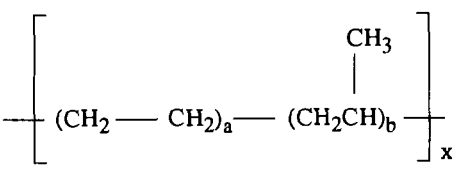
These are polymerised esters of methacrylic acid. They normally exhibit pour-point depressing activity. Dispersant properties can be obtained by incorporating polar groups in the molecular structure.

(b) Polyisobutenes (PIB)



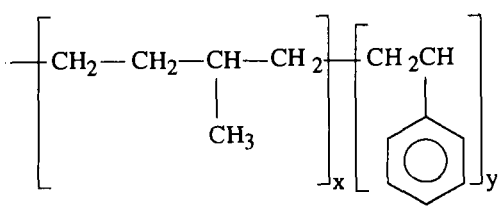
These are non-dispersant polymers and they have no effect on the pour point of formulated lubricants. They have limited use in modern formulations.

(c) Olefin co-polymers (OCP)



These are usually co-polymers of ethylene and propylene. Dispersant properties can be obtained incorporating polar groups in the molecular structure.

(d) Styrene/diene co-polymers



The molecular weight distribution is optimised to give good shear stability in crankcase applications. Its uniquely effective thickening power in solution gives an overall thickening efficiency that is superior to other polymers of equivalent shear stability.

FIG. 2 - VI-improvers

polymer causes more thickening at high temperatures, giving the oil "multigrade" characteristics. This is illustrated schematically in Fig.3. "Multigrade" oils fall into a higher viscosity band at 100°C than that of the straight "W" grade oil. For example, a SAE 10W40 oil has a viscosity of less than 3500 cP at -20°C but must have a viscosity greater than 12.5 cSt (about 10.5 cP) at 100°C, rather than 4.1 cSt for a SAE 10W oil (Table 2). Note that for practicality in quality control, Kinematic Viscosity (= Dynamic Viscosity/Density) is quoted for the high temperature part of the SAE classification.

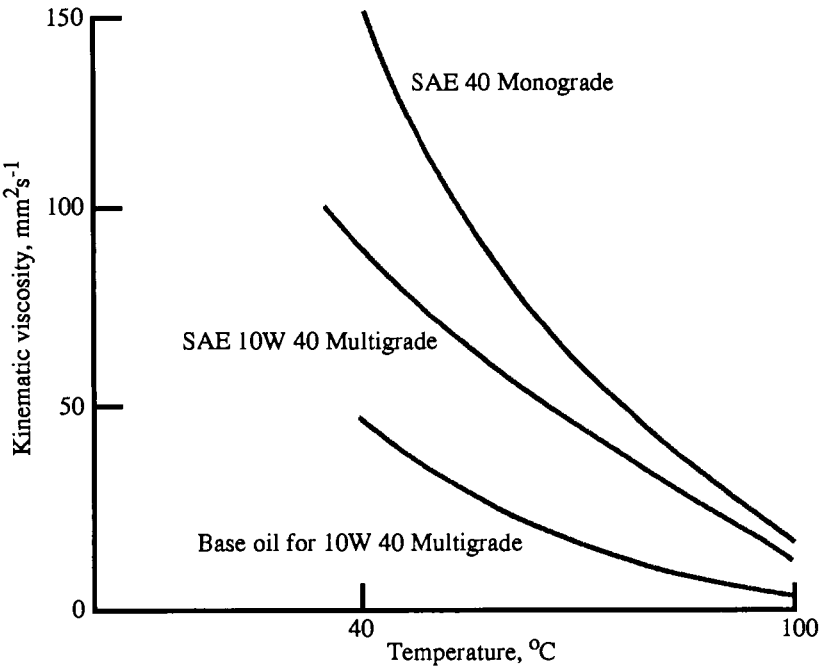


FIG. 3 - Viscosity/temperature behaviour of multigrade oils

The incorporation of polymeric VI improvers confers measurably non-Newtonian properties on the oil. Under high shear rate conditions, above about 10^6 s^{-1}, the effective thickening provided by the polymer can diminish significantly, as illustrated in Fig.4. This reduces friction in the critical contacts of the engine, where shear rates can be of the order of 10^7 s^{-1} or greater, but may also be expected to reduce the oil film thickness. This effect can be offset with certain polymers that confer visco-elastic properties on the oil, which has been shown to enhance the oil film thickness in engine bearings (Ref.2). Thus, the balance between frictional efficiency and wear protection can be optimised by the judicious selection of the types of polymer and viscosity grades. The European CCMC classification includes a limit of 3.5 cP minimum for high-temperature, high-shear viscosity (HTHSV) at 150°C and 10^6 s^{-1}. At the time of writing the current SAE J300 classification (revised February 1992) includes the HTHSV limits shown in Table 3. Further revisions of SAE J300 are imminent.

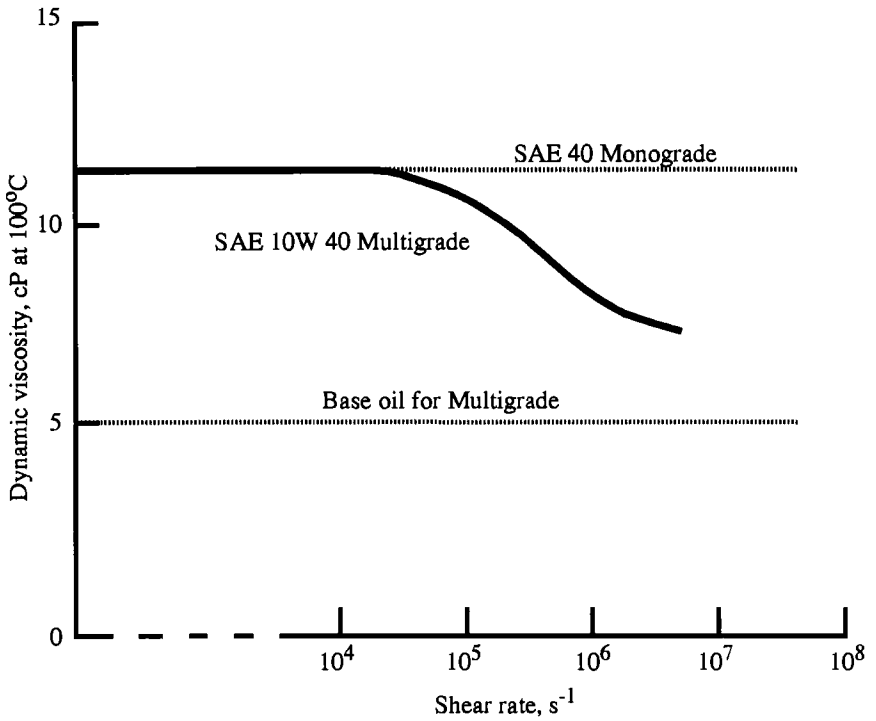


FIG. 4 - Viscosity/shear rate behaviour of multigrade oils

Table 3: High temperature high shear viscosity limits of the SAE J300 classification

SAE Viscosity Grade	High Temperature High Shear Viscosity[e,f] at 150°C and $10^6 s^{-1}$ (cP) min.
0W	2.4
5W and 10W	2.9
15W, 20W and 25W	3.7

(e) ASTM D 4624, ASTM D 4683, ASTM D 4741, CEC l-36-A-90.
(f) High shear viscosity specifications not required for single grade oils.

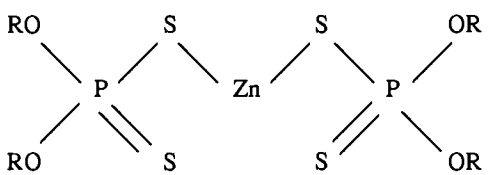
Oils containing polymeric additives also suffer permanent viscosity loss during shearing because of degradation of the polymer molecules by the extremely high shear stresses experienced by the lubricant in concentrated contacts, eg. cam/follower, oil pump, gears. The

reduction in viscosity again depends on the molecular structure of the polymer, its concentration, and the severity of the contacts in the engine. In this case, viscosity loss is cumulative and thus dependent on the duration of service. Most specifications for multi-grade oils include requirements for permanent shear stability assessed in the Bosch diesel injector rig test (DIN 51382). VI improvers also suffer thermal and oxidative degradation, but, generally, these can only be evaluated effectively in engine tests.

4.2 Anti-wear additives

Some contact between the sliding surfaces inevitably occurs at critical points in the operating cycle, particularly during running-in, transient operation, start-up and stopping. Limited wear protection is offered by oxide films on the metal surfaces, and by sulphur compounds in the oil but anti-wear additives are nowadays essential to control wear to acceptable levels in modern engines.

Zinc dialkyl dithiophosphates (ZDTPs)(Fig.5) are used almost universally for this purpose. The mechanism by which they prevent wear is still not well understood, but there is evidence that they are capable of preventing both severe adhesive wear (e.g. scuffing during run-in) by the formation of thin sulphide reaction layers, and long-term wear by the deposition of thick (~ 0.1 μm) phosphate films. The former function is often referred to as "extreme pressure" or "EP" protection and the latter as "anti-wear" protection. Certain detergent additives also provide some protection against wear (Section 5.3).



The R group may be alkyl or aryl

FIG. 5 - Zinc dithiophosphate extreme pressure/anti-wear additives

The effectiveness of ZDTPs depends to a considerable extent on the nature of the R groups (Fig.5), which control the rate of thermal decomposition and subsequent reactivity. They may be either alkyl or (occasionally) aryl groups. By judicious combination of different chain lengths and structures, it is possible to obtain wear protection over the desired range of operating temperatures and materials. ZDTPs also perform an important role in protecting the oil against oxidation (Section 5.2) and here again the balance of performance and effective life of the additive is largely determined by the combination of R groups.

Current good quality crankcase lubricants have Zn and P contents in the range 0.05 to 0.2%m. Some specifications impose limits on the maximum phosphorus content that may be used. These may become more restrictive in future because of concern about the poisoning of catalytic exhaust emission control systems, although phosphorus appears to be only one of a number of causes of catalyst deterioration. The removal of phosphorus from engine lubricant formulations could lead to considerable difficulties in maintaining satisfactory wear protection.

4.3 Friction reducers

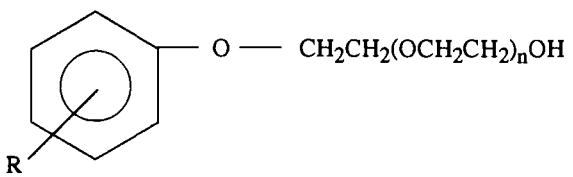
Recent years have seen strenuous efforts to reduce the fuel consumption of automotive

engines for both environmental and conservation reasons, particularly in gasoline engines in North America. The frictional properties of the lubricant have an important influence on the overall friction losses in both engines and transmission systems, as do their viscometric characteristics. A variety of esters, amides and metal soaps are available that can reduce boundary friction very significantly in simple laboratory tests (up to 30% reduction); this is found to translate into fuel economy benefits of up to about 4% in controlled comparative tests. A classification for "Energy Conserving" oils has been introduced, which is based on the US Sequence VI test and requires the demonstration of fuel consumption improvements of 1.5% (EC I) or 2.7% (EC II) relative to a reference oil in a 3.8l Buick V-6 gasoline engine under standardised operating conditions (Ref.1).

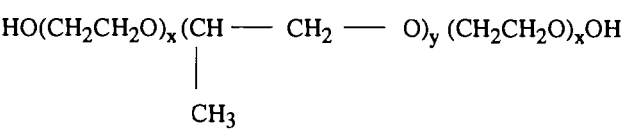
4.4 Corrosion inhibition

Corrosion of metal surfaces can be a problem both during delivery and in short trip, low-temperature operation, when water and acidic combustion products can condense into the crankcase. Rusting of ferrous surfaces is most common, but the corrosion of copper-lead and aluminium-tin bearing alloys must also be considered. Corrosion can cause seizure of vital components, such as valve lifters and bearings, and accelerates wear on sliding surfaces. The entrainment of the corrosion products, particularly oxide particles, in the oil can cause abrasive wear elsewhere in the engine.

Corrosion can be minimised by (a) neutralising acidic products before they reach the metal surface, and (b) preventing access of the corrosive material to the surface with a strongly adsorbed passivating film. The former is usually achieved by the overbased detergent and the dispersant additives (Section 5) but, under severe conditions, it can be necessary to enhance this protection with supplementary additives designed specifically to protect the metals at risk. Examples of two important classes of anti-rust supplements are shown in Fig.6.



(a) Alkylphenol ethoxyethanol in which n = 4 to 9 and R is often an iso-octyl or nonyl group



(b) An ethylene oxide/propylene oxide (EO/PO) block copolymer in which x could be about 4 and y could be about 55

FIG. 6 - Anti-rust additives

5. ENGINE COOLING AND CLEANLINESS

5.1 Volatility, Oxidation and Contamination

The most arduous duty of an oil, from the chemical viewpoint, is the cooling and lubrication of the piston surfaces, in particular, the surfaces under the piston crown, the piston pin bearing,

and the piston rings and grooves. In these areas, temperatures can rise typically to around 250°C and exceptionally to 300°C. The direct effect of this heating is the loss of some of the more volatile components of the lubricant, chiefly from the base oil, causing an increase in viscosity. To limit the adverse effects of volatility, lubricants are subjected to controls on the evaporative losses at high temperature; typically in the DIN 51581 standard in which maximum losses of between 13 and 25 % after 1 hour at 250°C are required, depending on the visosity grade and on the specifying organisation. This is achieved in the base oil manufacturing and quality control process.

At piston temperatures, the oil is subject to severe oxidation and some degree of thermal degradation. This results in the breakdown of the original molecular structures, the products of which polymerise causing oil thickening and the formation of gums and lacquer deposits on the piston and other engine surfaces. In the piston ring belt region and on the cylinder wall, the oil is also exposed to the induction charge and partially quenched combustion products. In addition to air, fuel and water, these gases contain partially oxidised hydrocarbon radicals and oxides of nitrogen, which create a particularly severe environment for oxidation that can lead to the formation of heavy carbonaceous deposits in the ring grooves and upper lands. This can result in ring sticking or blocking, loss of oil control and gas sealing, and in the extreme, piston seizure. Fortunately, the exposure of the oil to these conditions is extremely short.

The oil carries with it the products of oxidation and combustion as it drains back down into the sump, where the oxidation reactions continue for many hours at a much lower rate, catalysed by the active surfaces of wear debris, bearing alloys, etc.. Some of the blowby gases passing the piston rings also condense and become entrained in the sump oil, further increasing the level of contamination. Organic acids and oxides of nitrogen are the principal contaminants in gasoline engines, leading to corrosion and sludge formation. These problems are further aggravated by the closed crankcase ventilation and exhaust gas recirculation systems that are required to meet current and future anti-pollution standards. In diesel engines, soot and strong acids are more prevalent, causing oil thickening, sludges and corrosive wear. In both cases, significant quantities of environmental dust are ingested with the intake air, some of which finds its way into the oil, despite conventional air and oil filtration. Filtration is an important subject in its own right. Improved oil filtration can substantially reduce engine wear and increase component life (Ref. 3).

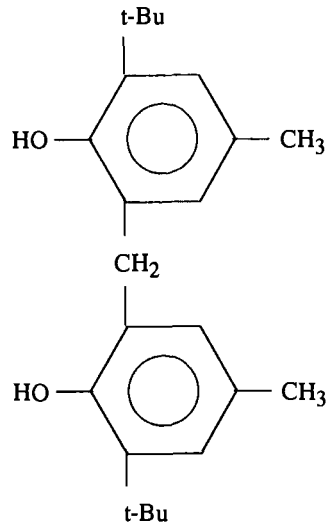
5.2 Resistance to Oxidation

The resistance of a lubricant to oxidation is governed in the first place by the oxidational stability of the base oil. For engine lubricants, mineral oils with good inherent oxidational stability are needed. This is usually achieved by the manufacture and selection of oils with minimum contents of aromatic and other unsaturated hydrocarbons, which tend to initiate the chain of oxidation reactions most rapidly. The content of unsaturated compounds can be further reduced by chemical processes after conventional refining to break unsaturated bonds by introducing more hydrogen into the molecular structure, e.g. hydrotreatment. In especially demanding applications, it may be necessary to employ synthetic ester fluids, some of which possess an inherent protection of the more vulnerable sites within their molecular structures, and which, in general, have better defined chemical composition than mineral oils.

The base oil properties alone are insufficient for the majority of engine applications. Antioxidant additives are used to improve them, ZDTPs (Section 4.2) being the most common. For good long-term oxidation protection, a ZDTP with high thermal stability is preferred. This can conflict with optimum anti-wear properties, so a combination of ZDTPs having a range of thermal stabilities is often required. This is obtained by control of the combination of R groups present (Fig. 5). A variety of phenol and amine compounds (Fig.7) are also available as

supplementary antioxidants, and a degree of antioxidancy is provided by certain types of detergent, alkyl salicylates being particularly effective in this respect. (Section 5.3). It should be noted, however, that the deficiencies of a poor quality base oil can rarely be compensated for by the use of heavy antioxidant treatments.

(a)



(b)

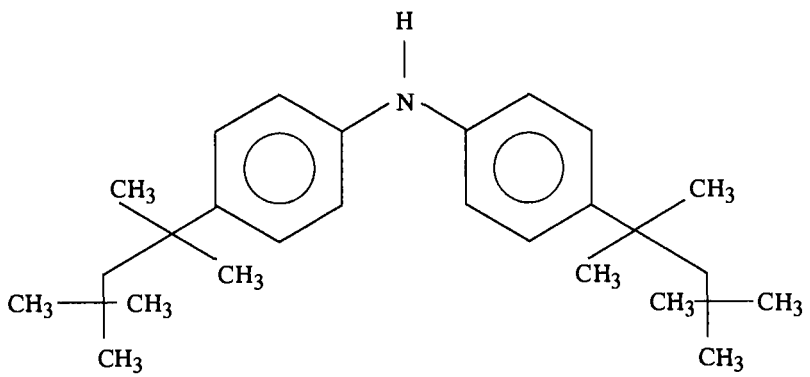


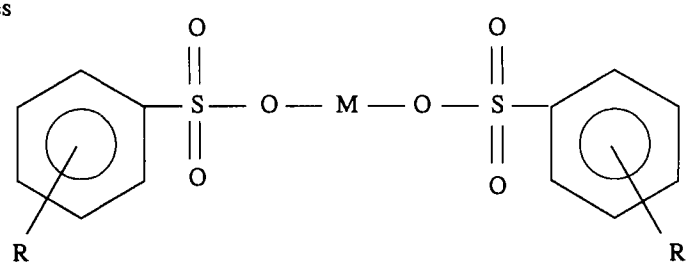
FIG. 7 - The chemical structure of a typical hindered bis-phenol (a) and a di-phenyl amine (b) antioxidant additive

5.3 Dealing with Deposits

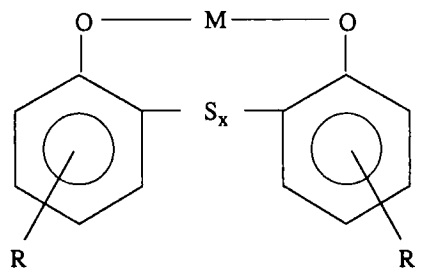
The build-up of carbonaceous deposits on the hotter parts of the piston can cause serious deviations from the design geometry, which leads to various kinds of malfunction ranging from progressive loss of oil consumption control to high rates of blow-by and eventual failure of lubrication. Certain organo-metallic compounds containing calcium or magnesium are effective in controlling high-temperature deposits. They are usually referred to as detergent

additives because it was originally thought that their main function was to clean off deposits after they had formed, although it is now believed that their ability to maintain material in suspension at high temperatures may be at least equally important.

(a) Sulphonates



(b) Alkyl phenates



(c) Alkyl salicylates

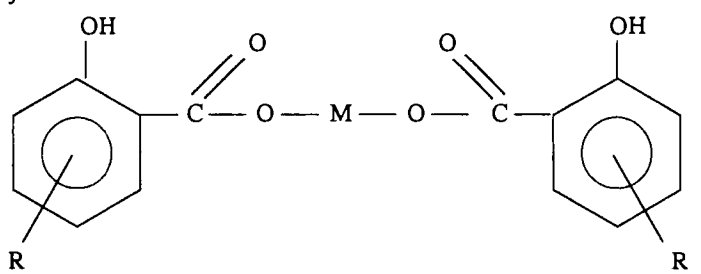


FIG. 8 - Typical structures of the main classes of detergent additives

The structures of the three main types of detergent: alkyl salicylates, phenates and sulphonates, are shown in Fig.8. The phenyl groups also provide antioxidant properties, thus inhibiting deposits at their source. The detergent can be made more alkaline, or "overbased", by the inclusion of excess metal carbonate, the structure being stabilised by micelle formation. This enables them to neutralise acids in the entrained combustion products, thereby providing protection against corrosive wear and rusting. The former is particularly important in diesel engines, which generate substantial amounts of strong acids derived from sulphur in the fuel. Thus, diesel engine oils tend to have higher alkalinity than motor oils. Alkalinity is expressed as Total Base Number (TBN), in units of mg of KOH/g. Motor oils generally have values

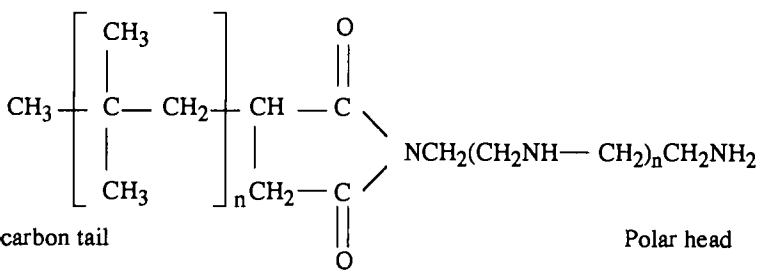
between 5 and 10, automotive diesel oils up to about 15, whilst medium-speed diesel engines running on residual fuels may require up to 30 mg KOH/g. Cylinder oils for large 2-stroke diesels may range up to 70 or 80 mg KOH/g.

5.4 Keeping it all in Suspension

The accumulation of insoluble material in the oil should ideally be maintained in suspension until it can be removed by the oil filter. Unfortunately, detergents are less effective in dispersing contaminants at the lower temperatures in the oil sump, particularly when the engine is stopped at ambient temperatures, or under cold stop-start conditions. The oxidised gums and lacquers, wear debris, soot, etc. then tend to deposit out as sludges, often in combination with condensed water. This can cause the blockage of oil passages, the oil pump inlet and the rocker cover, and, in particularly sooty engines, gel formation, with serious consequences for the supply of lubricant to critical parts.

Ashless (ie.non-metallic) dispersant molecules have a polar "head" which adsorbs strongly on the insoluble particles and a long hydrocarbon "tail" to obtain good oil solubility. They are effective at low temperatures and are available in a variety of complex chemical structures to confer differing combinations of polar attraction and oil solubility. Typical structures of the three most widely used types are shown in Fig.9. Some modern engines with exhaust emission control devices, especially those having poor crankcase ventilation, appear to present increasing problems of sludge build-up, particularly when operated over a widely varying temperature range. These difficulties can be resolved by the use of more highly active dispersants, but interference of such strongly polar materials with anti-wear film formation means that delicate compromises must be made to optimise the overall performance of the whole lubricant formulation.

(a) Mono-succinimides



(b) Bis-succinimides

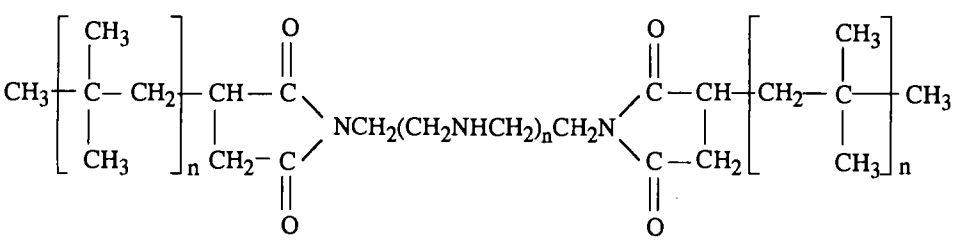
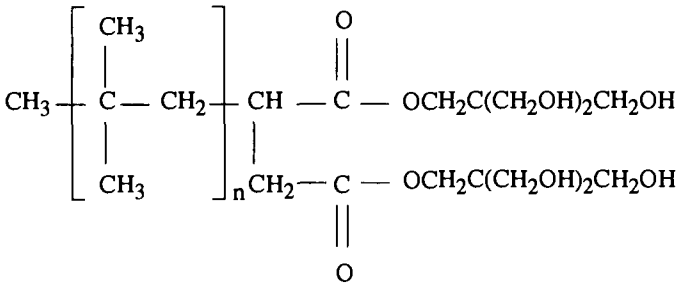


FIG. 9 - Representative structures of ashless dispersant additives
(continued on following page)

(c) Succinate esters



This PIB/MALA/PENTA type dispersant is an example of a succinate ester in which the polar head is obtained using pentaerythritol in place of the polyamine.

FIG.9 (contd....) - Representative structures of ashless dispersant additives

6. PERFORMANCE EVALUATION OF ENGINE LUBRICANTS

The performance of automotive engine lubricants is evaluated for development, specification and quality control purposes in a vast variety of laboratory glassware, rig and engine tests. These tests are intended to evaluate the properties described above, and a great many others that it has not been possible to discuss here. They are described more comprehensively in Reference 1.

The range of tests selected for the assessment of lubricant formulations during the development process will depend to a large extent on the performance specifications that are considered appropriate for the type of engines, vehicles and service duties that the lubricant will need to satisfy. These have tended to be dominated by the American Petroleum Institute (API) specifications for gasoline engine oils (API S-series) and diesel oils (API C-series) and, to a lesser extent, by military specifications, and the corresponding European classifications, the CCMC G-series and D-series, respectively. However, many engine and vehicle manufacturers have their own specification requirements, which may include tests specific to their own needs, in addition to the standardised industry tests, and these are becoming increasingly important. Many of these specifications require evaluation in both gasoline and diesel engine tests to safeguard the use of a lubricant in mixed fleets of vehicles. It should be noted that the performance required in these tests, and the test procedures themselves, can change quite frequently in response to the evolution of engine design, environmental legislation and other factors.

The purpose of standardised specification tests is to evaluate particular well-defined aspects of performance, at a level of severity equal to, or exceeding, that expected to be encountered in service, in as reproducible a manner as possible. In many cases, considerable efforts are made during the development of a test method to ensure that the problems that occur in service are adequately simulated in the final standard procedure. However, it is most unlikely that the full demands of vehicle operation, both foreseen and unforeseen, can ever be reproduced with a

combination of laboratory tests. Thus, to fully evaluate a lubricant, it is still essential to conduct extensive well-controlled trials in vehicles under a wide range of real service conditions.

The process of development and approval evaluation is far too costly therefore to be employed for quality assurance purposes. Once a lubricant formulation has been approved, quality control can normally be ensured with a much smaller range of laboratory tests. These will usually include tests for viscosity at high and low temperature, pour point for cold-flow properties, and a measurement of the appropriate metals' contents to ensure that the correct additive concentrations have been used (e.g. Ca,Mg,Zn).

7. ACKNOWLEDGEMENTS

The author is grateful to colleagues at Thornton Research Centre and in Shell International Chemical Company for their help in preparing this article, and to Shell Research Ltd. for permission to publish it.

REFERENCES

1 Shell International Chemical Company, 1992, "Reference Data for Crankcase Oils".
2 Bates, T.W., Williamson, B., Spearot, J.A. and Murphy, C.K., 1986, "A Correlation Between Engine Oil Rheology and Oil Film Thickness in Engine Journal Bearings", SAE Technical Paper 860376.
3 Needelman, W.M. and Madhavan, P.V., 1988, "Review of Lubricant Contamination and Diesel Engine Wear", SAE Technical Paper 881827.

This Page Intentionally Left Blank