

Selected Technical Papers



STP 1542

Rolling Element Bearings

Editors:

Yoshimi R. Takeuchi

William F. Mandler

Selected Technical Papers STP1542

Rolling Element Bearings

Editors:

Yoshimi R. Takeuchi

William F. Mandler



ASTM International
100 Barr Harbor Drive
PO Box C700
West Conshohocken, PA 19428-2959

Printed in the U.S.A.

ASTM Stock #: STP1542

Library of Congress Cataloging-in-Publication Data

ISBN: 978-0-8031-7528-0

This publication has been registered with the Library of Congress.
Library of Congress control number 2012036737

Copyright © 2012 ASTM INTERNATIONAL, West Conshohocken, PA. All rights reserved.
This material may not be reproduced or copied, in whole or in part, in any printed, mechanical, electronic, film, or other distribution and storage media, without the written consent of the publisher.

Photocopy Rights

Authorization to photocopy items for internal, personal, or educational classroom use, or the internal, personal, or educational classroom use of specific clients, is granted by ASTM International provided that the appropriate fee is paid to ASTM International, 100 Barr Harbor Drive, P.O. Box C700, West Conshohocken, PA 19428-2959, Tel: 610-832-9634; online: <http://www.astm.org/copyright>.

The Society is not responsible, as a body, for the statements and opinions expressed in this publication. ASTM International does not endorse any products represented in this publication.

Peer Review Policy

Each paper published in this volume was evaluated by two peer reviewers and at least one editor. The authors addressed all of the reviewers' comments to the satisfaction of both the technical editor(s) and the ASTM International Committee on Publications.

The quality of the papers in this publication reflects not only the obvious efforts of the authors and the technical editor(s), but also the work of the peer reviewers. In keeping with long-standing publication practices, ASTM International maintains the anonymity of the peer reviewers. The ASTM International Committee on Publications acknowledges with appreciation their dedication and contribution of time and effort on behalf of ASTM International.

Citation of Papers

When citing papers from this publication, the appropriate citation includes the paper authors, "paper title", J. ASTM Intl., volume and number, Paper doi, ASTM International, West Conshohocken, PA, Paper, year listed in the footnote of the paper. A citation is provided as a footnote on page one of each paper.

Foreword

THIS COMPILATION OF *Selected Technical Papers*, STP1542, *Rolling Element Bearings*, contains peer-reviewed papers that were presented at a symposium held April 13–14, 2011 in Anaheim, CA, USA. The symposium was sponsored by ASTM International Committee F34 on Rolling Element Bearings.

The Symposium Chairperson was Dr. Yoshimi R. Takeuchi, The Aerospace Corporation, Los Angeles, CA and the Co-Chair was Mr. William F. Mandler, Enceratec, Inc., Columbus, IN. The Associate Editor of the STP publication is Dr. Richard Neu and the Editors are Dr. Yoshimi R. Takeuchi and Mr. William F. Mandler.

Contents

Overview	vii
White Etching Crack Failure Mode in Roller Bearings: From Observation via Analysis to Understanding and an Industrial Solution J. Luyckx.	1
A Universal Bivariate Weibull Model for Static and Dynamic Fatigue Reliability Forecasting E. Y. Robinson	26
Roller Profile Development for an Axially Loaded, Single Row Spherical Roller Bearing in an Oscillating Application J. H. Cowles, Jr. and C. A. Houle	47
A Model to Estimate Separator Forces during Ball Speed Variations A. Leveille, P. Frantz, and G. Rosene	71
Bearing Thermal Conductance Measurement Test Method and Experimental Design Y. R. Takeuchi, S. E. Davis, M. A. Eby, J. K. Fuller, D. L. Taylor, and M. J. Rosado	92
Steel and Hybrid Spacecraft Ball-Bearing Thermal Conductance Comparisons Y. R. Takeuchi, S. E. Davis, and M. A. Eby	118
Resilient and Corrosion-proof Rolling Element Bearings Made from Superelastic Ni-Ti Alloys for Aerospace Mechanism Applications C. DellaCorte, R. D. Noebe, M. K. Stanford, and S. A. Padula	143
Evaluating the Impact of a Surprise Silicone Additive to a Synthetic Hydrocarbon Lubricant J. T. Hanks, D. W. Smith, C. J. Stevens, and R. E. Winkel	167

Overview

This book comprises a select collection of papers based on presentations given at the 2011 ASTM International Symposium on Rolling Element Bearings, on April 13–14, 2011 in Anaheim, CA. A total of twenty presentations provided insight into continuing advances in bearing technology. Of these, eight were chosen to be peer reviewed and selected for inclusion in this Special Technical Publication.

The symposium was the seventh in a series intended to share bearing technology developments at the international level. The first four symposia were sponsored by the REBG (Rolling Element Bearing Group). The seventh symposium is the third sponsored by ASTM International. Dr. Yoshimi R. Takeuchi served as symposium chair while the co-chair was Mr. William F. Mandler.

The Symposium's audience included bearing designers and developers, manufacturers of bearings and parts, material producers, researchers, and those interested in advanced bearing applications and bearing system development. The goal is to provide an overview of recent achievements in bearing technology and provide the engineer insight into ways this information could be used.

A global panel of experts was assembled to address various topics, including the introduction of novel testing methods together with acquired data. Other papers describe new analytical approaches for assessing the life of hybrid bearings and predicting cage instability, cover failure modes and their solutions, and convey means of designing rolling element bearings. In addition, this book contains unique test data on new materials, including advanced ball and race materials and the impact of lubricant impurities on performance. These subjects help create a technical knowledge base that enhances the bearing engineer's capabilities.

The papers contained herein demonstrate the commitment of the ASTM F34 committee to provide timely information to the rolling element bearing technology community.

Johan Luyckx¹

White Etching Crack Failure Mode in Roller Bearings: From Observation via Analysis to Understanding and an Industrial Solution

REFERENCE: Luyckx, Johan, "White Etching Crack Failure Mode in Roller Bearings: From Observation via Analysis to Understanding and an Industrial Solution," *Rolling Element Bearings* on April 13–15, 2011 in Anaheim, CA; STP 1542, Yoshimi R. Takeuchi and William F. Mandler, Editors, pp. 1–25, doi:10.1520/STP103908, ASTM International, West Conshohocken, PA 2012.

ABSTRACT: Some roller bearing applications are prone to the white etching crack (WEC) failure mode. The applications seem to have in common that they work under dynamic operating conditions. The specific feature of this failure mode is that the subsurface microstructure of a failed bearing contains modified material structures near cracks which are white after a nital etching test. In case of a WEC failure, the real lifetime of the bearing is much lower than the theoretical lifetime calculation. The hypotheses of fatigue overload, hydrogen, and accumulated plastic microstrain are evaluated and a root cause hypothesis is developed based on observations. The white etching material structures are interpreted as adiabatic shear bands generated by an impact load mechanism. We developed the root cause hypothesis that the dynamic operation of a roller bearing is generating a bearing internal pressure peak causing loads at high strain rate which result in material damage and initiate the WEC failure mode. Impact tests reveal a high sensitivity of through hardened martensitic and bainitic bearing steels for the adiabatic shear band failure mode. The origin of the bearing internal pressure peak is further explained based on available ElastoHydrodynamic Lubrication (EHL) experimental and simulation results. The generation of butterflies and WEC networks is interpreted as recrystallisation driven by high stress after many load cycles or a moderate stress combined with a high strain rate loading. The industrial experience is analysed from the perspective of the root cause hypothesis. The Weibull curve of a WEC bearing failure case is explained based on the material parameter full width at half-maximum (FWHM) at the

Manuscript received April 11, 2011; accepted for publication May 10, 2012; published online August 2012.

¹Hansen Transmission International nv, Lommel, 3920 Belgium.

Copyright © 2012 by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

raceway surface. The potential solutions of an optimised microstructure, black oxidized treatment, and hot assembly are identified based on positive industrial experience.

KEYWORDS: WEC, roller bearing, white etching

Introduction

Some roller bearing applications are prone to the white etching crack (WEC) failure mode. The specific feature of this failure mode is that the subsurface microstructure of a failed bearing contains modified material structures near cracks which are white after a nital etching test. The white colour is a chemical inert reaction of the modified material structure for nital. A typical example of such a material investigation for a martensitic bearing material is shown on the left side of Fig. 1. The same result is shown for a bainitic microstructure which is also characterised by development of an abnormal high butterfly density and butterflies in the material volume with low shear stress levels.

The failure mode of a WEC martensitic bearing failure is the development of axial hairline cracks followed by spalling. In a bainitic bearing variant, the subsurface WEC network will develop till big material parts will spall out of the raceway.

Detailed Investigations of White Etching Microstructure

The white etching microstructure has been investigated by many researchers. The reported results of these investigations are a higher hardness than the original material [1,2], the spheroidized primary carbides are dissolved [1–6], presence of nanocrystals [2–7]. The nanocrystals have a Body Centered Cubic (BCC) crystal lattice [1,2,4–6] and are identified as ferrite [2,4–6].

The white etching microstructure of a failed bearing has been investigated more in detail. The sample cutting was done with focussed ion beam (FIB) in the vicinity of the transition between the white etching microstructure and the original bearing material. This FIB sample is investigated by transmission

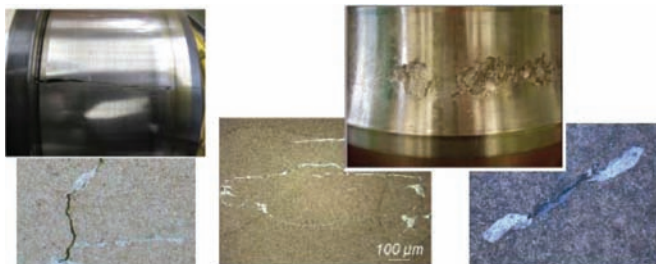


FIG. 1—WEC failed bearings and microstructure investigations. Left side is martensitic and right side is bainitic microstructure.

electron microscopy (TEM) and X-ray diffraction (XRD) analysis (Fig. 2). The carbides are indicated with white arrows and the dissolving process can be observed for the carbides in the white etching zone. There is a sharp transition between the original material matrix and the white etching microstructure.

The diffraction rings of area A corresponds with the original grain size which is much larger than the nanocrystals of area B. The BCC crystallographic structure is identified in both areas.

We can conclude that the material observations are in line with the results of other researchers. Based on the dissolving process of the carbides, we interpret that the white etching area is progressing in a direction so that the original bearing material is transformed and the sharp transition between both areas is moving progressively.

The typical white etching microstructure is observed in several other components. A few examples are fish eyes after very high cycle fatigue tests [8], chips and surfaces after machining of steels [9–13], in failed gears [14,15], in material samples after an impact test [16–18] and on corrugated surfaces of rail treads [19,20].

WEC Bearing Failure Characteristics

WEC bearing failures are observed in several roller bearing applications, with different bearing types and from different bearing suppliers. Some examples out of the public domain are automotive alternators [3,21], belt driven pulley systems [21] for engine accessories [6,22,23], bearings applied in automotive powertrains [24], generators and gear units of wind turbines [25]. The dynamic operation conditions of the bearing application is a common feature of the observed WEC bearing failures. The dynamic excitations in automotive pulley

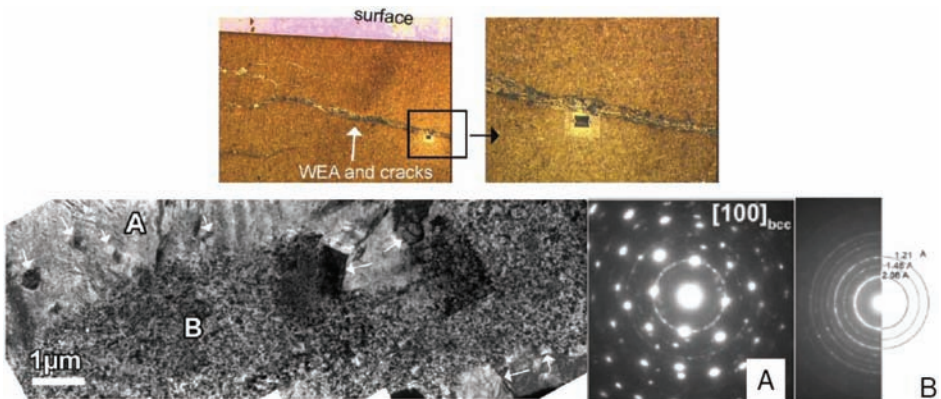


FIG. 2—FIB sample preparation (upper), TEM and XRD analysis of areas A and B (lower).

systems can be generated from the combustion engine vibrations or belt resonance which creates an oscillation of the rotational speed superposed on the average pulley speed [6,23].

In a drive system or transmission, the WEC failure mode is observed on a specific bearing location and not on the other positions. The failures occur much before the L10 life of the bearing and are classified as premature failures. The failure rate can increase up to the complete population of a bearing application.

The typical Weibull curve of a WEC bearing failure case (Fig. 3) has a shape factor in the range of 2.5 to 3.5 which is a significant deviation compared with the Weibull of a normal bearing fatigue life. The shape factor above 1 means that after a limited lifetime, the failure rate is increasing with time and a significant part of the total population is failing in a small time frame. Another feature of this WEC failure case, is a survival population of about 40% after 20 months.

The Weibull curve shows that the L10 life is about 6 months which is much lower than the calculated lifetime of more than 100 years. The low real lifetime and observation of survival population are leading to the conclusion that the WEC failure mode is not a normal bearing fatigue problem, it is a robustness problem. In case of a bearing fatigue problem, the expectation is that the complete population will fail which is not in line with the observation of a survival population.

There are several patents or articles [21,23,24] claiming a technical solution for WEC based on a dedicated experience. A lot of research at different parties is ongoing. There is no agreed understanding of the WEC failure mode



FIG. 3—Weibull curve of a WEC bearing failure case.

within the technical world of bearing suppliers and appliers. In case that a WEC failure mode is observed, the design engineers are implementing new bearing types or bearing configurations based on best practice.

We will evaluate different hypotheses with available observations and analyse the industrial experiences to propose a root cause hypothesis, develop a WEC failure model and identify the possible industrial solutions.

Hypothesis Evaluation

Fatigue Overload Hypothesis

The available test data at the bearing suppliers are showing that white etching bands (left side of Fig. 4) can be generated at high stress levels and after sufficient load cycles [26–28]. An overview of the correlation between contact stress, inner ring revolutions and the microstructural alterations is presented at the right side of Fig. 4 [26].

The bearing test results leads to the interpretation that the white etching microstructures in the WEC failed bearing are generated by a fatigue overload in the bearing application. The hypothesis of fatigue overload is in contradiction with the observations of the WEC bearing failure mode on only 1 bearing location in a drive or transmission system and also no failures on other loaded components such as gears, shafts and housings. The presence of a fatigue overload is also in contradiction with the lack of observation of typical fatigue material degradation via measurement of residual stress and full width half max values. Measurement campaigns performed with instrumented bearings in the application, are confirming that the external bearing loads are in line with the applied specifications for bearing design resulting in the conclusion that the applied Hertzian stress level is according the calculated stress level.

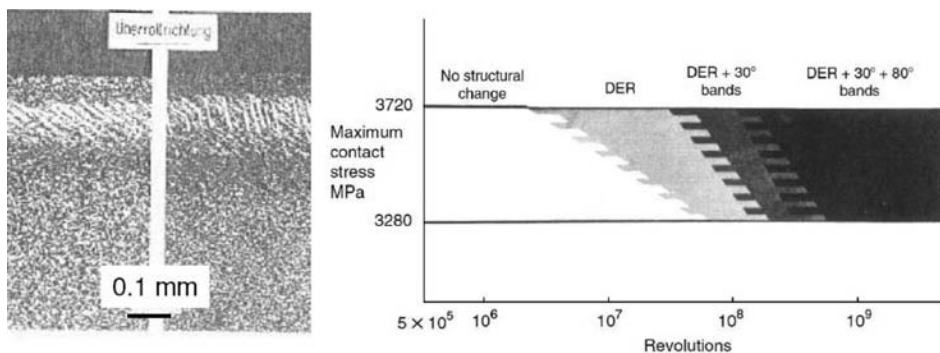


FIG. 4—white etching bands [27] and microstructural alterations as function of contact stress and inner ring revolutions [26]. (DER = dark etching region).

This leads us to the conclusion that the white etching microstructures and high WEC bearing failure rates cannot be explained by a fatigue overload and thus also not only by the applied Hertzian stress level. We will explain later that the Hertzian stress is an element of the total loading system but there is also an additional material loading system active.

Hydrogen Hypothesis

Hydrogen is considered by a lot of researchers as the main driver for the WEC failure mode [3,21,22,24,29]. The hydrogen hypothesis is based on hydrogen generation in the oil film followed by diffusion in the bearing material. This process will continue until a sufficiently high hydrogen content is realised to initiate a hydrogen damage mechanism resulting in cracks and the typical white etching microstructure of the WEC failure mode.

We performed hydrogen measurements on failed and unfailed inner rings from bearing applications prone to the WEC failure mode (Fig. 5). The hydrogen observations are classified by green blocks for max 1 ppm and red above 1 ppm. The borderline of 1 ppm is a global agreed reference for potential negative impact of hydrogen [24].

The hydrogen measurements are interpreted that the hydrogen content in the bearing material is a result of hydrogen generation at the raceway surface and transport from the raceway towards the core after the initiation of damage at the raceway. The hydrogen level in the bearing material is below the critical level of 1 ppm before the bearing failure. The presence of local atomic hydrogen cannot be discussed based on these observations.

Kino and Otani [24] have reproduced WEC's in carbonitrided, quenched and tempered SCM420H after 2.6×10^7 stress cycles. The combination of a stress level of 3.6 GPa, an operating temperature of 120°C and a specific oil is resulting in a global hydrogen level of 4.2 ppm [24] and leads to the B-type of

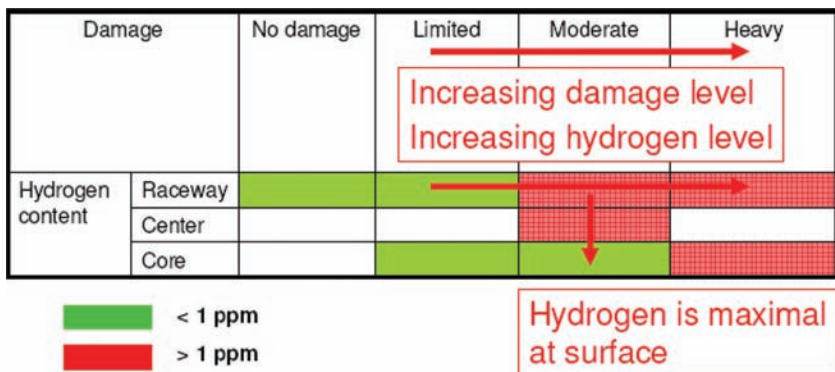


FIG. 5—Hydrogen observations in bearing applications prone to the WEC failure mode.

white structure which is similar with subsurface WEC network. Vegter and Slycke [29] have reproduced WEC networks in martensitically hardened and tempered ASTM 52100 at 3.2 GPa and with hydrogen charging up to 4.8 ppm after 1.5 to 4.0×10^6 revolutions.

The hydrogen observations show a lack of hydrogen in the bearing in the early state of failure (see Fig. 5, no damage and limited). The second discrepancy is the high Hertzian stress levels of 3.2 [29] and 3.6 GPa [24] to generate the WEC subsurface network. The measurement campaigns have supplied the proof that the applied load levels do not exceed the loads applied for the bearing design. Table 1 is a comparison between the applied test conditions of the hydrogen research [24,29] and a WEC failure case with a 50% failure rate after 18 months. The calculated Hertzian stress level of this bearing application is at 1.3 GPa.

A bearing test with hydrogen precharged 6309 bearings at 1.3 GPa stress and 6000 RPM, will accumulate after 100 h a total of 1.8×10^8 stress cycles and an expected failure rate of 50% based on field data. This is in fact a very fast and representative test to confirm that hydrogen charging is the root cause for the WEC failure mode. Up to now, there are no results reported about tests with hydrogen charged bearings loaded in the range between 1 and 1.5 GPa and oil temperatures between 50 and 80°C which are the application parameters of the example mentioned in Table 1.

The hydrogen hypothesis is also not able to predict a survival population as noticed in the Weibull curve of a WEC failure case (Fig. 3).

We will explain further that there is in fact an additional loading mechanism active. The impact of this additional loading mechanism can be simulated in some extent by charging the bearing with hydrogen but this is insufficient because there is still the need for Hertzian stress levels which are much higher than the stress level of the bearing application (see Table 1). The conclusion is that hydrogen is somewhat an accelerator for the WEC failure mode but it cannot explain the observed WEC failures at Hertzian stress levels between 1 and 1.5 GPa.

Accumulated Plastic Microstrain

Vincent et al. [30] developed a model where they proved that dislocations generated by the difference in stiffness between inclusion and surrounding bearing

TABLE 1—Comparison of bearing test and application parameters for WEC failure mode.

Application	Contact stress/ 1.3 GPa	Stress Cycles/ 1.8×10^8	WEC failure rate
Bearing test rig and hydrogen charging [29]	246%	11%	100%
Bearing test rig [24]	277%	28%	40%
Bearing application	100%	100%	50%

material, are moving towards an area which corresponds with the orientation, shape, and size of butterfly wings relative to the inclusion (Fig. 6). The physical interpretation is that the cyclic loading of the inclusion, resulted in an accumulated plastic microstrain distribution corresponding with the white etching area of the butterfly.

White etching microstructure is interpreted by many researchers [1–6,28–31] as localised plastic microdeformation and this is in line with the proposed root cause hypothesis in the Root Cause Hypothesis section.

Root Cause Hypothesis

Interpretation of White Etching Material Structures

Different WEC failure cases have been analysed and the change from martensitic towards bainitic material without modification of bearing type or other elements of the bearing application, resulted in the typical WEC material observations for respectively martensitic and bainitic bearings (see Fig. 1). This is in fact the proof that there is an identical loading mechanism driving the WEC failure mode and causing two different material responses depending of the bearing material type.

The axial hairline crack development in a cylindrical and taper roller bearing both out of martensitic material, is another similarity which allowed us to

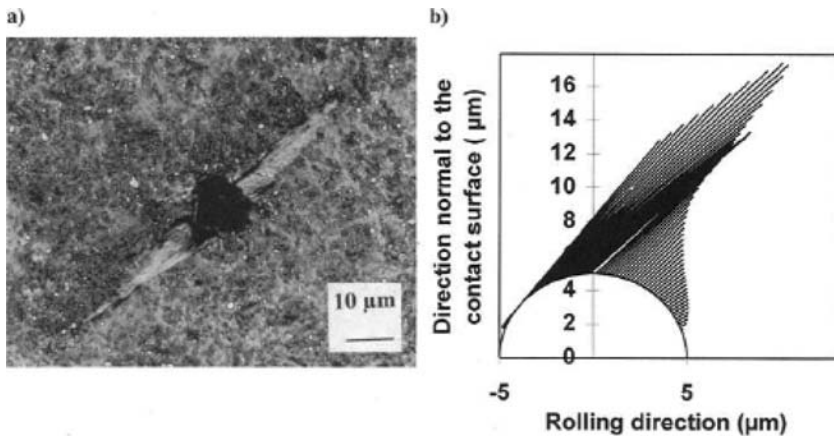


FIG. 6—[30] Comparison between an experimental white etching area (WEA) and the corresponding calculated domain swept by dislocations ($P_H = 4.5$ GPa); (a) optical micrograph of the WEA observed around an aluminium oxide inclusion found at 0.6 mm beneath the surface (the shear stress is max at 0.38 mm beneath the surface); (b) calculated domain (only the half system is represented).

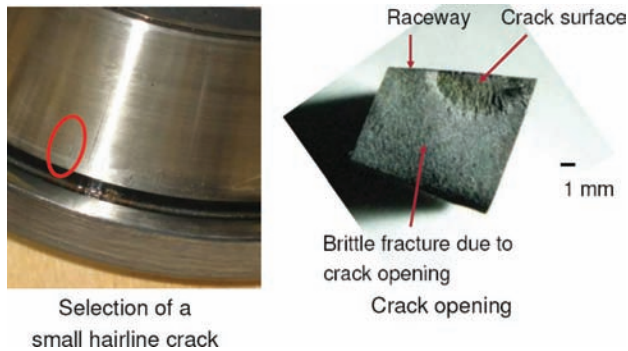


FIG. 7—Initial crack selection and crack opening.

reduce 4 WEC failure cases with two different materials, three different bearing types on three different shafts, towards one general problem. A more detailed study of one WEC bearing failure case has the potential to generate fundamental understanding of the WEC failure mode. An initial axial hairline crack of a martensitic taper roller bearing has been selected and opened (Fig. 7).

Scanning Electron Microscope (SEM) investigation of the crack surface reveals the presence of a semicircular crack just below the raceway (Fig. 8 upper part). This semicircular crack has a flat appearance and is classified as transgranular because there are no grain boundaries observed at the crack surface.

The etching test performed on a perpendicular section of the crack surface is presented in Fig. 8 lower part. A surprising observation is that the crack has two propagation directions, one is perpendicular with the raceway and the other is parallel or tilted relative with the raceway. The white etching material structures are observed below the semicircular crack surface and only in the part of the crack parallel or tilted with the raceway. There are beside the crack also white etching zones parallel with the raceway which do not belong to the main crack system (Fig. 8, lower part, right side).

We selected a crack with clear observations to enable a good interpretation. In some crack investigations you will find less developed cracks, smaller segments with white etching areas and sometimes there are local influences of inclusions. The observations of Fig. 8 are typical and systematic for a WEC failure case in martensitic material.

The total crack pattern is interpreted as being caused by an impact load generating a fast semi circle cleavage fracture under tensile stress, fast stepped perpendicular crack growth deeper in the material under tensile stress and fast developing parallel cracks caused by shear stress (see Fig. 9). This interpretation of the crack pattern is resulting in the classification of the white etching areas as an adiabatic shear band caused by an impact load.

The impact load is generating a high strain rate and this mechanical energy is nearly completely transferred into very local heat. As the local heat

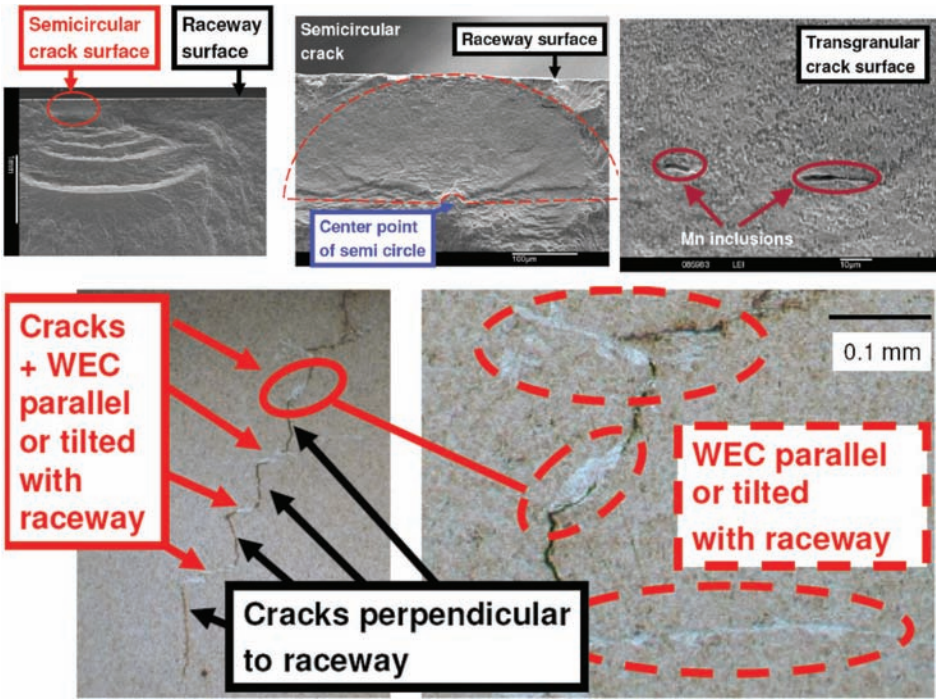


FIG. 8—SEM investigation of the opened crack (upper) and etching test performed on perpendicular section (lower).

generation process is very fast, there is no time for heat transfer and for that reason the process is named adiabatic. It is known from observations [1] and impact tests [16], that a shear stress component will enable the adiabatic shear failure mode with nanograins and white etching material structures. Due to the

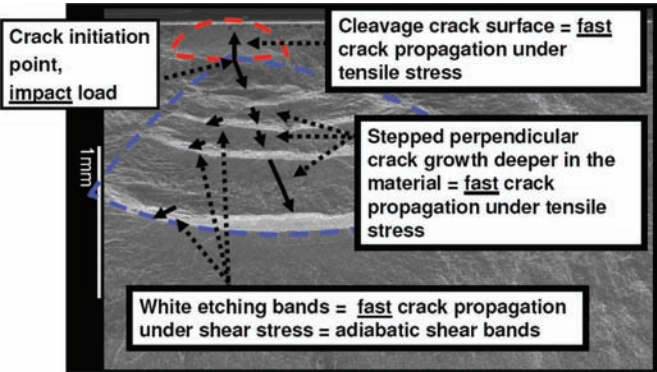


FIG. 9—Interpretation of total crack pattern.

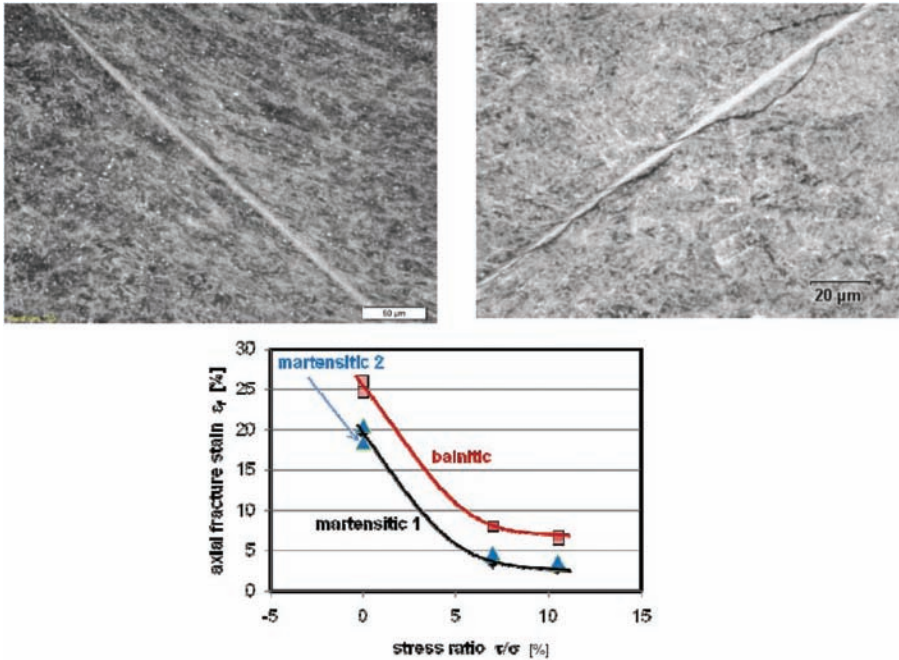


FIG. 10—Adiabatic shear band after an impact test performed on samples out of bearing steel (upper). Axial fracture strain as function of ratio shear stress to compression stress (lower).

lack of heat transfer and the shear stress component, the white etching bands obtained after impact tests are classified as adiabatic shear bands.

The adiabatic shear band in through hardened martensitic and bainitic bearing steel is observed after an impact test performed on test samples manufactured via wire erosion out of standard bearings (Fig. 10 upper part).

The high sensitivity of through hardened martensitic and bainitic bearing steels for the adiabatic shear failure mode is proven experimentally by the observations of shear band localisation with pure compression impact load (Fig. 10, upper part, right) and significant reduction of axial fracture strain at a low ratio of shear to compression stress (Fig. 10 lower part).

The interpretation that the white etching material structures in WEC failed bearings is an adiabatic shear band, is in line with assumptions of a recrystallisation process of other researchers [3–5,7,28,29,31]. The adiabatic shear band hypothesis is directly supported by other researchers [7,23] and the interpretation of white bands after a bearing endurance test [32]. The high deformation rate and heat development during machining of steels is the driver that in chips and also at the machined surface, adiabatic shear bands are generated [9–13]. The similarity in the diffraction patterns (very fine grains and BCC structure)

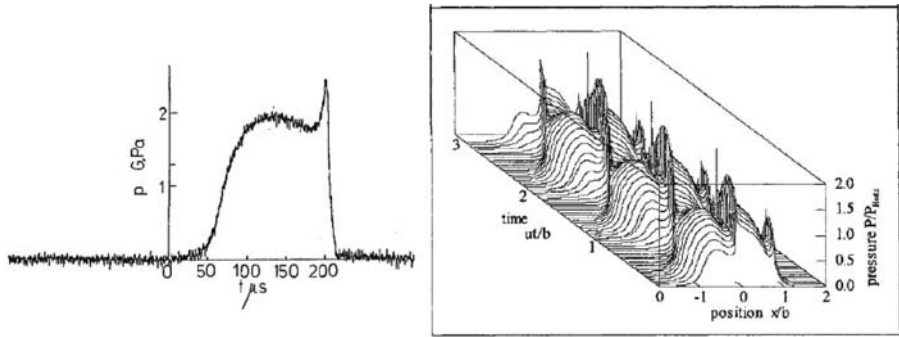


FIG. 11—Measured pressure peak as function of time in a bouncing ball experiment [33] (left) and simulated pressure peak superposed on Hertzian pressure for varying load [34] (right).

between metal chips and butterfly wings in martensitic bearing steel is mentioned in Ref [2].

Impact Load

The patent in Ref [23] is pointing to an external impact load as driver to create an adiabatic shear band and causing also an indentation on the raceway surface. The lack of indentation and surface damage on the raceway close to the surface crack associated with the semicircular crack, is interpreted that there was no excessive external load to drive the generation of the adiabatic shear band. The impact load which is driving the WEC failure mode is very probably generated in the oil film between roller elements and rings.

There are measurements and simulations available of film separation and squeeze effects in the EHL contact [33–37] which predict or confirm that under varying slip and load conditions, pressure peaks are generated. Figure 11 left [33] is an example of the pressure measured as function of time in a bouncing ball experiment. The pressure maximum at about 130 μ s is the pressure when the oil film thickness is minimal and the pressure peak at about 200 μ s is just before the ball and the plate are separated.

The pressure measurement possibilities are limited due to the short lifetime of the applied film sensor technology and test conditions which can be generated on test rigs. These limitations are leading the EHL experts to perform simulations and visual measurement of oil film thickness to generate more understanding about pressure, temperature and separation in an EHL contact.

Figure 11 right [34] is an example of a simulation which reveals that pressure peaks are superposed on top of the Hertzian pressure. The simulated operation condition is a variable load which is resulting in pressure peaks at inlet that are transported by the oil film. The horizontal axis shows the oil film

contact (left side is inlet) and the diagonal axis represents the time relation via transformation of position x which is divided by the half width of the oil film b and multiplied with film speed u . EHL tests reveal that forced excitation is generating oil film perturbations at the inlet which will travel through the contact at the average speed [35,36].

A similar analysis can be done for varying slip conditions. EHL pressure measurements under constant slip conditions [37] are confirming that the pressure peak at the outlet is increasing with increasing slip (Fig. 12 left).

The dynamic operating condition of varying slip during operation is simulated [34] and the resulting pressure variations are shown in Fig. 12 right. The change in slip is creating a change of oil film thickness which is then transported into the oil film. The change in oil film thickness has been simulated in an experimental setup on a two roller test rig. One roller has on the surface a waviness pattern which created the oil film thickness variations. The experimental results in Fig. 13 [38] are a proof that the oil film thickness variations are resulting in significant pressure variations.

The discussed EHL experimental results and simulations are supporting the assumption that there are internal pressure peaks generated in the oil film of a bearing under dynamic load conditions. The pressure peaks will create a fast increase of the load and causing a high strain rate resulting in the generation of adiabatic shear bands [32]. The pressure peak is associated with a speed profile in the oil film causing a tensile stress at the surface [39]. The pressure peaks are in fact an additional load mechanism generating loads at high strain rate which result in material damage and initiate the WEC failure mode.

This root cause hypothesis can now be applied to explain the material response which is otherwise only observed at much higher stress levels as discussed in the Fatigue Overload Hypothesis section or a result of hydrogen and an increased stress level (the Hydrogen Hypothesis section). The development of WEC networks and a high butterfly density outside the high stress areas in bainite can be explained with the same root cause hypothesis, i.e., the Hertzian pressure with addition of pressure peaks.

Butterfly and WEC Network

The generation process of a butterfly is interpreted as a two stage process [32]. There is first the accumulation of plastic micro deformation which can be visualised via the dislocation distribution presented in Fig. 6 [30]. The local accumulated plastic deformation is reducing the capability of the material to resist against recrystallisation.

The second step is the development of white etching butterfly wings which is created by recrystallisation via two possible mechanisms, i.e., high stress after many load cycles or a moderate stress combined with a high strain rate loading. The butterflies observed in bearings after high load tests are

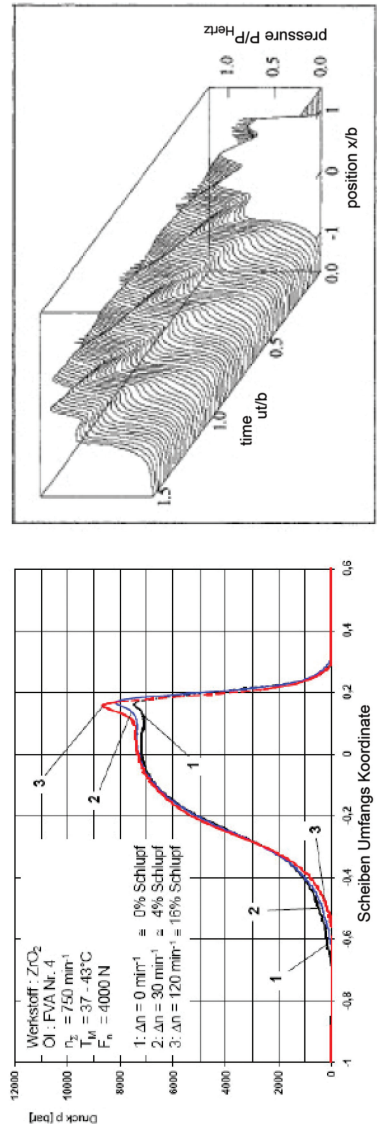


FIG. 12—Measured pressure in EHL contact for different slip values [37] (left) and simulated pressure peak superposed on Hertzian pressure for varying slip [34] (right).

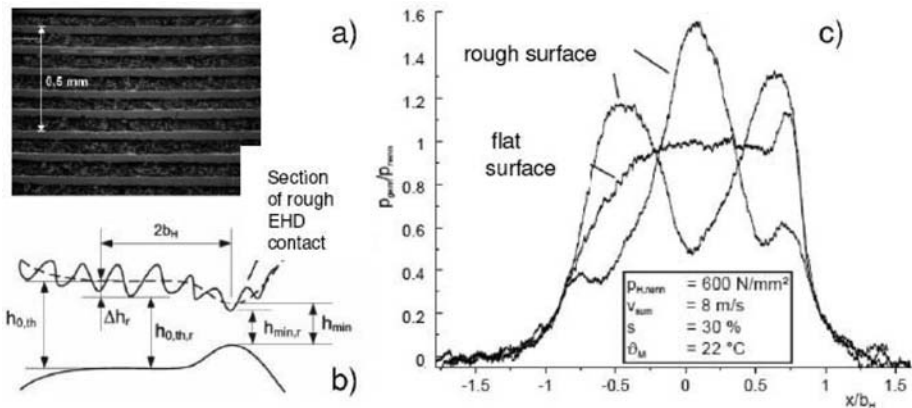


FIG. 13—(a) Roller with waviness; (b) sketch of EHL contact; (c) measured pressure [38].

interpreted as being generated as consequence of contact stress above 2.2 GPa and time [40]. A limited butterfly development in bearing applications with contact stresses in the range between 1 and 1.5 GPa, can be explained by a local stress riser from the inclusion geometry or interaction with other inclusions [41]. A high density of butterflies in bearing applications with contact stresses in the range between 1 and 1.5 GPa is explained by the high strain rate caused by the pressure peak. The repeating action of pressure peaks will result in additional plastic deformation and butterfly development at locations with a much lower Hertzian stress level than observed in bearing tests where butterflies are generated with high stress after many load cycles.

The crack will appear during the development process of the butterfly when the local conditions for crack initiation and propagation are favourable. This may occur before or after the butterfly wings are developed. The crack tip and the transition between white etching material and original bearing material are in fact generators of new dislocations and plastic micro deformation. If the pressure peaks are sufficiently high, then the butterfly will start to propagate and may end up in a WEC crack network.

A WEC crack network will start from material features which will enhance the local development of an adiabatic shear band via plastic deformation and strain hardening. If the local material conditions are favourable for WEC initiation, then there is a high probability for further development of the WEC network resulting in the bearing failure as seen in bainitic material (Fig. 1, right).

The material properties of bainitic bearing steel are able to accumulate for a long period the effects of the pressure peaks and result in a high butterfly density outside the high stress area and WEC networks. The martensitic bearing steel has a lower accumulating effect and high butterfly densities in WEC failed martensitic bearings are for the time being only observed in combination

with a high inclusion density. The different material response of martensite and bainite is supported by the observed difference in impact load robustness of both bearing steels as presented in Fig. 10.

Summary of Root Cause Hypothesis

Roller bearing applications with dynamic operating conditions will create internal pressure peaks which are superposed on the Hertzian pressure. The pressure peak is resulting in a impact load mechanism which is generating tensile and shear stresses at a high strain rate.

The impact load mechanism is able to create material damage in martensitic bearing steel via the semi circle cleavage crack, cracks perpendicular with the raceway and adiabatic shear bands (with white etching material structures) tilted or parallel with the raceway. This material damage is the start point for development of an axial hairline crack which is followed by raceway spalling and bearing failure.

The contact stress and the impact load mechanism will develop in bainitic bearing material plastic deformation around local material features. The impact loading will transform the initial plastic deformed areas into white etching areas and butterflies. The continued application of the impact load will result in the further growth of the white etching areas till a WEC network forms and the development of butterflies at depths outside the high shear stress areas. The WEC networks will develop till the raceway is spalling which is then the end of the bearing life.

The root cause hypothesis for the WEC failure mode in martensitic and bainitic bearing steel is represented in Fig. 14.

The industrial experience will now be analysed to verify this root cause hypothesis and to identify possible industrial solutions for the WEC failure mode in roller bearings.

Industrial Experience

General Experience

The industrial experience is based on gear units applied in wind turbines. The reference is a total of 20 000 gear units manufactured in a time frame of 20 years.

The risk for WEC bearing failures is linked with the roller size of the bearing. A heavier roller will increase the dynamic energy in the squeezed oil film and generate increased pressure peaks which are experimentally observed in [33]. The appearance of WEC failures in much smaller bearings of automotive applications [3,6,21–24], can be explained by the increased dynamic excitations of belt driven engine auxiliaries and automotive powertrain applications.

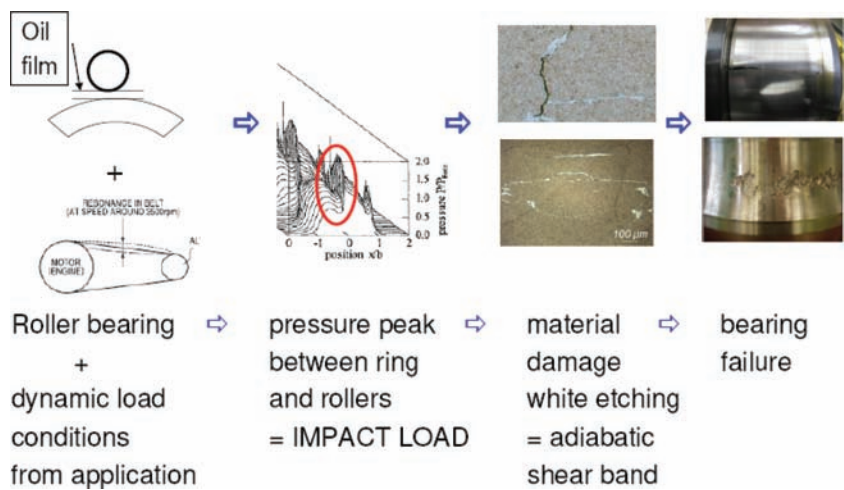


FIG. 14—Root cause hypothesis of the WEC failure mode.

Spherical roller bearings and angular contact ball bearings are more prone to the WEC failure mode. The contact between roller elements and raceways of both bearing types is such that there are two pitch lines with slip free contact and outside there is kinematic slip in the oil film due to internal geometry as presented in Fig. 15 [32].

The kinematic slip will increase the pressure peak (see Fig. 12 left). The pitch line positions of a spherical roller or ball bearing in operation will vary somewhat under dynamic load and internal bearing dynamics. The change of pitch line position is resulting in a variation of slip in the EHL oil film and enhancing the generation of pressure peaks as explained in the Impact Load section.

Cylindrical bearings under dynamic axial load conditions will be prone to variations of roller tilting which is causing variation of the borders of the EHL

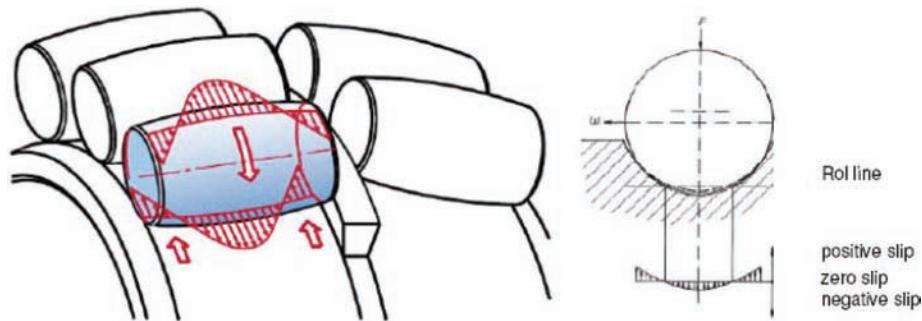


FIG. 15—Contact between roller/ball and raceway with two pitch lines and kinematic slip.

contact. This mechanism will result in pressure peaks at inlet that are transported by the oil as explained in the Impact Load section.

Experience Based on Introduced Changes

We observed a reduction of the WEC failure rate from 40 % to 2.7 % after 14 months operation when the bearings have been replaced by non optimised case carburised versions without any other change. The positive impact of a case carburised microstructure can be explained by the presence of a significant amount of retained austenite which has a FCC crystallographic lattice. Impact testing performed on a mix of FCC and BCC crystal lattice types in Ti, reveal reduced sensitivity for the adiabatic shear failure mode if a sufficient amount of the metal matrix is FCC [16]. The comparison of different case carburised bearing microstructures reveals the observation of features that may reduce the impact loadability. The material analysis and additional performance data are applied to identify an optimised case carburised microstructure to avoid the WEC failure mode in roller bearings.

By analysing the field performance of gear units, two extraordinary observations are detected. The WEC failure rate after 2 years field operation reduced from over 40% to zero failures for two changes without any other modification in the bearing application. This extraordinary performance improvement is noticed after introduction of black oxidised bearings and after bearing assembly at 130°C. The bearing assembly temperature is linked with the applied bearing heating process. Due to differences in bearing heating processes, we were able to put the WEC failure mode on, off and on again and this for a statistically relevant population which is a proof for the effectiveness of the parameter “bearing assembly temperature.” More detailed analysis did reveal that bearing assembly temperature is a significant parameter to avoid the WEC failure mode. The total observation is summarized in Table 2.

Impact of Bearing Assembly at 130°C

The bearing material has been stabilized for several hours at a temperature of 160°C or higher before the hard machining operation. We expect that the

TABLE 2—*Observations of WEC failure rate and correlation with bearing assembly temperature.*

Bearing Assembly Temperature	Failure Rate	Period	Gearbox Population
100°C	40%	14 months	400
110°C	12%	14 months	
110 ↔ 120°C	7,3%	15 months	41
130°C	0%	26 months	34

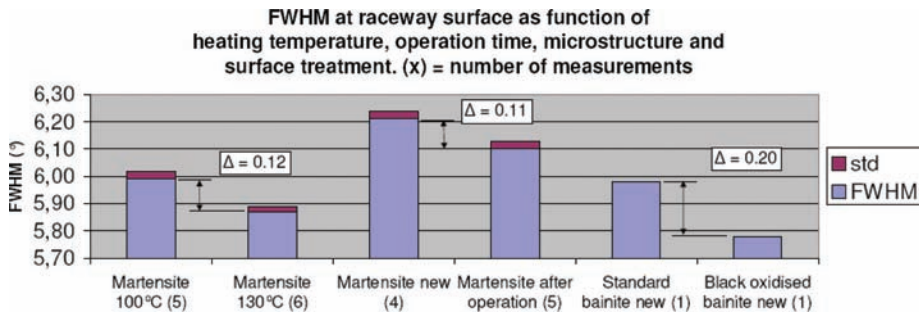


FIG. 16—Comparison of raceway surface FWHM values.

heating of the bearing up to 130°C will change only the material parameters of the hard machined surfaces.

The microstructural changes of the raceway surfaces have been investigated via XRD measurements of the full width at half maximum (FWHM) of the (211) ferrite reflection on Cr-K α radiation. We have cut two material samples out of the same bearing ring and both samples have been heated in a similar way as a bearing assembly process at respectively 100°C and 130°C. The raceway surface FWHM values of an inner ring have been compared between a loaded and not loaded zone after field operation. The FWHM measurement of the raceway surface has been repeated on a standard bainitic and a black oxidised bainitic bearing. All the measurement results are shown in Fig. 16.

The bearing assembly at 130°C is generating a FWHM reduction of 0.12 which is considered small but identified as a relevant change of the microstructure [42] and in line with the expected range between 0.05 and 0.18 [43].

The overall conclusion is that bearing assembly at 130°C, operation time in the field and the black oxidising process are generating reduced FWHM values at the raceway surface. The black oxidising is a process of about 20 min in a bath between 120°C and 140°C which has a similarity with bearing assembly at 130°C. The impact of operation time is interpreted as the consequence of a surface load mechanism [44,45]. A research program is ongoing to compare the physical parameters of the raceway surfaces with original and reduced FWHM value.

WEC Failure Model

The reduction of raceway surface FWHM with operation time is applied to interpret the WEC Weibull curve. A Weibull curve of the first 66 bearing applications is calculated to avoid the influence of bearing applications with limited operation time (Fig. 17).

We identify three different phases in the bearing life of this application prone to WEC failure mode :

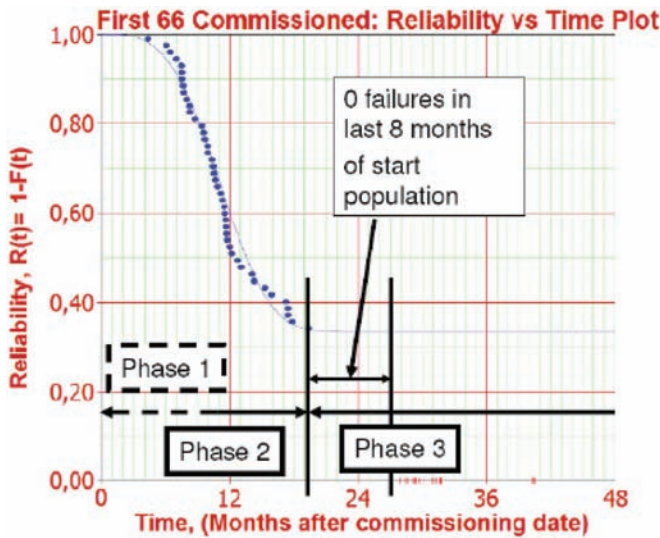


FIG. 17—Weibull of first 66 bearing applications.

Phase 1: potential damage initiation phase which is limited in time till the surface FWHM value is reduced by bearing operation.

Phase 2: failure development phase with observation of bearing failures up to some time after start of the bearing application.

Phase 3: survival phase if no damage was generated during phase 1.

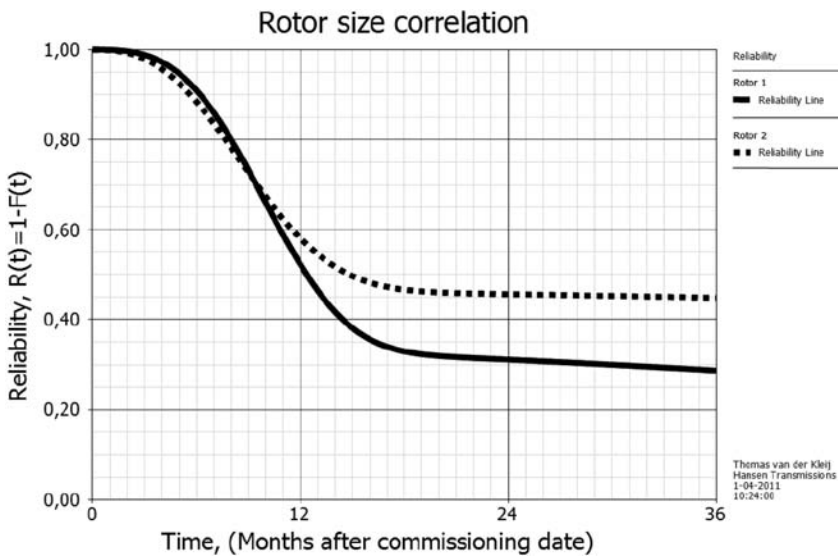


FIG. 18—WEC Weibulls of identical gear unit applied with two rotor sizes.

The Weibull curve of a WEC failure case may deviate from the graph shown in Fig. 17 due to limited occurrence of the dynamic operation conditions or influence of applied loads on the failure development.

The loads will have an impact on the probability for dynamic operation conditions during phase 1. Also load level combined with load cycles will control speed of FWHM reduction at the raceway surface (end of phase 1) and speed of failure development during phase 2. The correlation between loads and the parameters of the WEC Weibull can be visualized by a Weibull analysis per load category. The influence of a different rotor for a WEC bearing failure case in a wind turbine gear box is presented in Fig. 18. The applied gear units are identical and the different rotor size is resulting in different failure rates at the end of phase 2.

The proposed WEC failure model allows us to assess the failure risk and define optimized service strategies.

Summary and Potential Industrial Solutions

The features of WEC failed bearings have been explained. The results of electronic microscopy and X-ray diffraction analysis of the white etching microstructure are presented. The WEC bearing failure characteristics are analysed.

The root cause hypotheses of fatigue overload and hydrogen have been analysed and are rejected based on available observations. There is a limited impact of hydrogen on the WEC failure mode identified but not sufficient to explain the field failures at stress levels in the range between 1 to 1.5 GPa. White etching microstructure is interpreted by many researchers as localised plastic micro deformation.

Different WEC failure cases have been generalised and a detailed analysis of a martensitic WEC failed bearing is discussed. The white etching material structures in a WEC failed bearing are interpreted as adiabatic shear bands generated by an impact load mechanism. We developed the root cause hypothesis that the dynamic operation of a roller bearing is generating a bearing internal pressure peak causing loads at high strain rate which result in material damage and initiate the WEC failure mode. The root cause hypothesis has been verified with impact tests and the available industrial experience. There are no elements in contradiction with this root cause hypothesis.

The FWHM value at the raceway surface is identified as a major parameter to control the WEC failure mode. The WEC failure model is defined with three different phases in the bearing life, i.e., potential damage initiation phase, failure propagation phase, and survival phase. The proposed WEC failure model allows to assess the failure risk and define optimized service strategies.

The industrial experience to avoid the WEC failure mode is presented. The identified potential industrial solutions can be classified in two categories:

- (a) additional raceway surface treatment such as black oxidising or bearing assembly at 130°C;
- (b) bearing material with optimised case carburised microstructure.

The field observations and developed root cause hypothesis are leading to the conclusion that roller bearings have a conceptual endurance limitation in case of dynamic operation conditions. This conceptual limitation is enhanced in case of big size roller bearing applications. The proposed potential solutions of black oxidising, bearing assembly temperature at 130°C and optimised case carburised microstructure will increase the system robustness of big size roller bearing applications such as wind turbine gear units. A similar effect is expected for small bearing applications with high dynamic excitation.

Further research is required to generate a better understanding of the WEC failure mode so that the effectiveness of the discussed potential industrial solutions can be assessed. The observation of a failure stabilisation in the WEC Weibull after some time, is increasing the confidence level that the proposed potential industrial solutions are a long term solution.

References

- [1] Schlicht, H., 1973, "Über die Entstehung von White Etching Areas (WEA) in Wälzelementen," *HTM*, 28, pp. 112–123 (About the Generation of White Etching Areas in Rolling Contact).
- [2] Becker, P. C., "Microstructural Changes Around Non-metallic Inclusions Caused by Rolling-contact Fatigue of Ball-bearing Steels," *Met. Technol.*, June, 1981, pp. 234–243.
- [3] Hiraoka, K., Nagoa, K., and Isomoto, T., "Study on Flaking Process in Bearings by White Etching Area Generation," *J. ASTM Int.*, Vol. 3, No. 5, 2006, Paper ID JAI14059.
- [4] Grabulov, A., Ziese, U., and Zandbergen, H.W., "TEM/SEM Investigation of Microstructural Changes Within the White Etching Area Under Rolling Contact Fatigue and 3-D Crack Reconstruction by Focused Ion Beam," *Scr. Mater.* 57, 2007, pp. 635–638.
- [5] Grabulov, A., and Zandbergen, H. W., "TEM and Dual Beam (SEM/FIB) Investigations of Subsurface Cracks and White Etching Area (WEA) Formed in a Deep Groove Ball Bearing Caused by Rolling Contact Fatigue (RCF)," Proceedings of VHCF-4 Conference, Belgrade, Serbia, 2007.
- [6] Harada, H., Mikami, T., Shibata, M., Sokai, D., Yamamoto, A., and Tsubakina, H., "Microstructural Changes and Crack Initiation with White Etching Area Formation Under Rolling/Sliding Contact in Bearing Steel," *ISJI Int.*, Vol. 45, No. 12, 2005), p. 1897–1902.
- [7] Shibata, M., Gotoh, M., Oguma, N., and Mikami, T., Proceedings of the International Tribology Conference, Yokohama, Japan, 1995, pp. 1351–1356.

- [8] Sakai, T., "Crack Initiation Mechanism of Bearing Steel in Very High Cycle Fatigue," *Proceedings of European Conference of Fracture*, Alexandroupolis, Greece, July 3–7, 2006.
- [9] Miller, S. F., Blau, P. J., and Shich, A. J., "Microstructural Alterations Associated With Friction Drilling of Steel, Aluminum, and Titanium," *J. Mater. Eng. Perform.*, Vol. 14(5), Oct 2005, pp. 647–653.
- [10] Molinari, A., Musquar, C., and Sutter, G., "Adiabatic Shear Banding in High Speed Machining of Ti-6Al-4V: Experiments and Modeling," *Int. J. Plasticity*, Vol. 18, 2002, pp. 443–459.
- [11] Clos, R., Schreppel, U., and Veit, P., "Chip Formation and Strain Localization in 100Cr6," 1st Colloquium Process Scaling, Bremen, Germany, Oct 28–29, 2003, pp. 1–9.
- [12] Li, J. G., Umemoto, M., Todaka, Y., and Tsuchiya, K., "A Microstructural Investigation of the Surface of a Drilled Hole in Carbon Steels," *Acta Mater.*, Vol. 55, 2006, pp. 1397–1406.
- [13] Ramesh, A., and Melkote, S. N., "White Layer Formation in Machining of Hardened Steel: Experiments and Modeling," *High Speed Machining of Hard/Super Hard Materials Conference*, Nov 7–11, 2003, Singapore.
- [14] Michaelis, K., "Gear Failures Pitting," Gear Research Centre TU München, Course at the Univ. of Ljubljana.
- [15] Alban, L. E., "Systematic Analysis of Gear Failures," American Society for Metals Gauga County, Ohio, 1985.
- [16] Meyer, L. W., and Krueger, L., "Investigation of Titanium Alloys Under Biaxial Impact Loading," Final Technical Report, Nordmetall GbR, Stollberg, Germany, August 25, 1996.
- [17] Wei, Q., Kecskes, L., Jiao, T., Hartwig, K. T., Ramesh, K. T., and Ma, E., "Adiabatic Shear Banding in Ultrafine-grained FE Processed by Severe Plastic Deformation," *Acta Mater.*, Vol. 52, 2004, pp. 1859–1869.
- [18] Yang, Y., Xiong, J., and Yang, X., "Microstructure Evolution Mechanism in Adiabatic Shear Band in TA2," *Trans. Nonferrous Met. Soc. China*, Vol. 14, No. 4, Aug 2004.
- [19] Baumann, G., Fecht, H. J., and Liebelt, S., "Formation of White-etching Layers on Rail Treads," *Wear* Vol. 191, 1996, pp. 133–140.
- [20] Wang, L., "Microstructure and Residual Stress State in the Contact Zone of Rails and Wheels," thesis, Berlin, 2002.
- [21] Ohki, C., and Goto, T., Compressor Bearing and Compressor Component, U.S. Patent 7585114 B2, filed March 26, 2004, issued September 8, 2009.
- [22] NSK info brochure, "Long-Life Bearings for Engine Accessories."
- [23] Umeda, A., Shiga, T., and Ihata, K., Apparatus and Method for Determining Patterns of Damage Being Caused in Rolling Contact Element, U.S. Patent 2007/0044543, filed July 26, 2006, issued March 1, 2007.

- [24] Kino, N., and Otani, K., "The Influence of Hydrogen on Rolling Contact Fatigue Life and its Improvement," *JSAE Rev.* Vol. 24, 2003, pp. 289–294.
- [25] McVittie, D., "Wind Turbine Gearbox Reliability, The Nature of the Problem," Wind Turbine Reliability Workshop, Sandia National Labs, Albuquerque, NM, October 3–4, 2006.
- [26] Swahn, H., Becker, H., and Vingsbo, O., "Martensite Decay During Rolling Contact Fatigue in Ball Bearings," *Metall. Trans A*, Vol. 7A, 1976, pp. 1099–1110.
- [27] Zwirlein, O., and Schlicht, H., "Werkstoffanstrengung beim Wälzbeanspruchung - Einfluß von Reibung und Eigenspannungen," *Z. Werkstofftech.*, Vol. 11, 1980, p. 1.
- [28] Voskamp, A. P., "Material Response to Rolling Contact Loading," *J. Tribol.*, Vol. 107, 1985, pp. 359–366.
- [29] Vegter, R. H. and Slycke, J. T., "The Role of Hydrogen on Rolling Contact Fatigue Response of Rolling Element Bearings," *J. ASTM Int.*, Vol. 7, No. 2, 2006, Paper ID JAI102543.
- [30] Vincent, A., Lormand, G., Lamagnère, P., Gosset, L., Girodin, D., Dudragne, G., and Fougères, R., "From White Etching Areas Formed Around Inclusions to Crack Nucleation in Bearing Steels Under Rolling Contact Fatigue," *Bearing Steels: Into the 21st Century*, ASTM STP 1327, J. J. C. Hoo and W. B. Green, Eds., ASTM International, West Conshohocken, PA, 1998, pp. 109–123.
- [31] Moreira de Freitas, D. F., "Formation de Phase Blanche en Fatigue de Roulement," (Generation of White Etching Microstructure During Bearing Fatigue Testing), *Scr. Metall.*, Vol. 17, 1983, pp. 683–686.
- [32] Schlicht, H., "Über Die Adiabatic Shearbands und Die Entstehung der Steilen Weißen Bänder in Wälzlagern," (About the Adiabatic Shear Bands and the Generation of "High-angle White Bands" in Roller Bearings), *Materialwiss. Werkstofftech.*, Vol. 39, No. 3, 2008, pp. 217–226.
- [33] Safa, M. M. A., and Gobar, R., "Pressure Distribution Under a Ball Impacting a Thin Lubricant Layer," *J. Tribol.*, Vol. 108, 1986, pp. 372–376.
- [34] Hooke, C. J., "Dynamic Effects in EHL Contacts," *Proceedings of the 29th Leeds-Lyon Symposium on Tribology*, York, UK, September 3–6, 2002, Elsevier B. V., Vol. 41, 2003, pp. 69–78.
- [35] Kalogiannis, K., Mares, C., Glovnea, R. P., and Ioannides, S., "Elastohydrodynamic Film Thickness Response to Harmonic Vibrations," International Multi-Conference on Engineering and Technological Innovation, 2008, <http://www.iiis.org/CDs2008/CD2008SCI/IMETI2008/PapersPdf/F687SR.pdf> (04/2011).

- [36] Kaneta, M., Ozaki, S., Nishikawa, M., and Guo, F., "Effects of Impact Loads on Point Contact Elastohydrodynamic Lubrication Films," Proceedings of the Institution of Mechanical Engineers, Part J, Sage Publications, *J. Eng. Tribol.*, Vol. 221, Number 3, 2007, pp. 271–278.
- [37] Sperrfechter, T., "Keramische Bauteile im Elastohydrodynamische Kontakt," (Ceramic components in elastohydrodynamic contact) Institut für Keramik im Maschinenbau, Universität Karlsruhe (TH), 1998, ISSN 1436-3488.
- [38] Kreil, O., "Einfluss der Oberflächenstruktur auf Druckverteilung und Schmierfilmdicke im EHD-Kontakt," (Influence of the surface texture on the pressure distribution and oilfilm in a EHD contact) Oliver Kreil, TU München, Diss, 2009, http://deposit.d-nb.de/cgi-bin/dokserv?idn=99328146x&dok_var=d1&dok_ext=pdf&filename=99328146x.pdf (04/2011).
- [39] Schlicht, H., and Broszeit, E., "Die Werkstoffbeanspruchung im Wälzkontakt bei Hoher Flächenpressung, Ermittelt Nach Hertz und Nach der EHD-Strömungshypothese," (Material stressing at rolling contact with high specific load determined by Hertzian- and elastohydrodynamic (EHD) flow hypothesis), *Materialwiss.Werkstofftech.* Vol. 38, No. 4, 2007, pp. 255–262.
- [40] Lund, T., "Sub-surface Initiated Rolling Contact Fatigue - Influence of Steel Matrix, Non-metallic Inclusions and Operating Conditions," *8th International ASTM Symposium on Bearing Steel Technologies*, Vancouver, May, 2009, ASTM International, West Conshohocken, PA.
- [41] Vincent, A., Fougères, R., Lormand, G., Dudragne, G., and Girodin, D., "A Physically Based Endurance Limit Model for Through Hardened and Surface Hardened Bearing Steels," *Bearing Steel Technology, ASTM STP 1419*, J. M. Beswick, ASTM International, West Conshohocken, PA, 2002.
- [42] Gegner, J., "Post-Machining Thermal Treatment (PMTT) of Hardened Rolling Bearing Steel," MMT, 2006, pp. 2-66–2-75, http://www.ariel.ac.il/management/research/pf/zinigrad/mmt/MMT-2006/Service_files/papers/Session_2/2-066_po.pdf (04/2011)
- [43] Gegner, J., "Evidence and Analysis of Thermal Static Strain Aging in the Deformed Surface Zone of Finish-machined Hardened Steel," *Powder Diffr. Suppl.*, Vol. 24, No. S1, June 2009, pp. 45–50.
- [44] Nierlich, N., and Gegner, J., "Material Response Models for Sub-surface and Surface Rolling contact fatigue," Proc MMT 2006, Vol. 1, Chap. 1, pp. 182–192, http://www.ariel.ac.il/management/research/pf/zinigrad/mmt/MMT-2006/Service_files/papers/Session_1/1-182_ma.pdf (04/2011).
- [45] Gegner, J., Nierlich, W., and Brückner, M., "Possibilities and Extension of XRD Material Response Analysis in Failure Research for the Advanced Evaluation of the Damage Level of Hertzian Loaded Components," *Materwiss. Werkstofftech.*, Vol. 38, No. 8, 2007, pp. 613–623.

Ernest Y. Robinson¹

A Universal Bivariate Weibull Model for Static and Dynamic Fatigue Reliability Forecasting

REFERENCE: Robinson, Ernest Y., "A Universal Bivariate Weibull Model for Static and Dynamic Fatigue Reliability Forecasting," *Rolling Element Bearings* on April 13–15, 2011 in Anaheim, CA; STP 1542, Yoshimi R. Takeuchi and William F. Mandler, Editors, pp. 26–46, doi:10.1520/STP104267, ASTM International, West Conshohocken, PA 2012.

ABSTRACT: A universal statistical model based on a bivariate Weibull distribution is presented to forecast the reliability of bearing rolling contact fatigue (RCF) life under various applied load levels. The model is derived from work done in the field of composite material static fatigue (also known as stress rupture), and a statistically significant database, including hundreds of observations, confirmed the model predictions. Although the detailed failure mechanisms in RCF (lubricant starvation, contamination, centrifugal forces, differential component growth, ball tracking, surface smoothness, etc.) that determine the RCF life of ball bearings are complex and the definition of failure rests on arbitrary test acceleration changes, the static fatigue model has attractive features that might be useful for RCF life data correlation and reliability forecasts, as illustrated in this article. Ultimately, the value of a forecast model rests on replicating experiments that confirm model predictions. There are two premises underlying the model. First, the fatigue life is represented by a Weibull distribution, characterized by the Weibull slope (shape factor), which is denoted here as b and which is normally assumed to be invariant with life. Second, the applied load parameter, for example, the median strength-normalized contact stress, also reflects a Weibull distribution at a *hypothetical constant time*. These assumptions lead to the convenient use of median-normalized variables for the life and load parameters and underlie what is here called the *reliability forecast chart*, giving a comprehensive overview of reliability for service life at a constant operating load or stress. We address details of this model with a graphical presentation of static and RCF data and data analyses that apply the form of the *static* fatigue model. In addition, a method of estimation of the two shape factors is addressed, as these underlie the trade, at constant reliability, between load and life (the

Manuscript received August 9, 2011; accepted for publication March 26, 2012; published online July 2012.

¹The Aerospace Corporation, Altadena, CA 91001.

Copyright © 2012 by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

so-called load-life exponent). Model variants that address progressive over-arching damage, such as corrosion, can also be drawn from the prior static fatigue model. For further development and application of the model, more data are required.

KEYWORDS: bivariate, Weibull, distribution, rolling, contact, fatigue, model

Nomenclature

- b = Weibull life shape factor (also Weibull slope of fatigue life data)
- C = constant
- d (subscript) = design survival probability
- Gl/Ep = glass strand/epoxy composite
- i = i th ranked observation
- Kev/Ep = Kevlar/epoxy composite
- Lx = fraction $(1 - x)$ survival life, e.g., L1 means 99 % survive
- m = Weibull strength shape factor
- o (subscript) = Weibull scale factor
- R = load parameter
- RCF = rolling contact fatigue
- RFC = reliability forecast chart
- r (subscript) = a reference value
- S = survival probability
- SNHB = silicon nitride (ball) hybrid bearing
- t = time or life parameter
- 50 (subscript) = median value
- Γ = function of the incomplete exponential integral
- σ = stress parameter

Introduction

The well-known Weibull statistical distribution [1,2], originally put forward to account for the strength scatter of brittle materials and the size effect, was early applied to fatigue life data [3] in a single variate form (life), as it did not naturally account for the load-life exponent that relates how life trades against applied load at constant reliability. Lundberg and Palmgren [4] tried to account for the effect of load via a shear stress based model, but their a priori load-life exponent prediction did not agree with experimental data. Many authors, e.g., Zaretsky [5], have addressed this issue and put forward a variety of load-life exponents. None of these has entered the mainstream of application, primarily as a result of the scarcity and scatter of test data. The very long fatigue life of modern hybrid bearings comprising ceramic balls (silicon-nitride) on hard steel races means that often tests are suspended. Zaretsky, in Ref 6, notes a discussion with Weibull wherein he proposes a bivariate exponential form very similar to the formulation presented here, which he evidently never published.

In this paper we propose a universal bivariate model for static fatigue and for hybrid bearing rolling contact fatigue (RCF) life addressing two variates. fatigue life and applied load. although we freely choose the load and then experimentally test the life distribution by replicating fatigue tests at constant load. We either have the strength distribution for b or use trial values of m (in Eq 11) to find a graphical fit to the data in order to establish the strength. The form of this model is rationalized in the section titled “The Bivariate *Static* Fatigue Reliability Forecast Chart” and is identical to the proposed dynamic (rolling contact) fatigue model. The mathematical form of the model (Eqs 11 and 13) underlies a comprehensive chart of reliability contours, with ordinate $\log(\text{load})$ and abscissa $\log(\text{life})$. We have named this the reliability forecast chart (RFC). The probability contours form a family of straight parallel lines with a slope defined by the ratio $-b/m$ of the Weibull life distribution shape factor to the strength distribution shape factor, a somewhat surprising result. We use this ratio to “back out” the strength shape factor when suitable strength distribution measurements are not available. Figure 3 depicts the RFC as applied to extensive *static* fatigue (stress rupture) data, with a close fit between data and model. The RFC may then be used to select the operating stress for a given reliability over a required service life.

A compact analytical model to forecast the *static* fatigue life of fiber composite materials (e.g., pressure vessels) subject to steady long-term pressure load was developed in the 1960s and applied to aerospace composite structures, as described in “The Bivariate *Static* Fatigue Reliability Forecast Chart” and in Refs 7 and 8.

The large differences between the failure mechanisms of RCF and static fatigue, as well as vastly more complex RCF failure mechanisms (lubricant starvation, contamination, centrifugal forces, differential component growth, ball tracking, surface smoothness, etc.), and the susceptibility of dynamic fatigue failure to many variable and subtle effects contribute to large and irregular data scatter. We attempt to apply the static fatigue model and the RFC in order to predict the fatigue lifetime of rolling component element bearings, which are speculative and vulnerable to fundamental arguments but still possibly useful.

Large, consistent databases are needed in order to quantify the way in which critical variables control the life and probability of bearing fatigue failure under specified loads (or load duty cycles) and within a certain time period. The exponential trade between load and life, given by Eq 11, is a term that gives a fixed probability while trading load and life.

This model (Eq 13) utilizes normalized variables and specifically calls for median normalized variables. So, in order to estimate the two shape factors, we need to test the rolling contact fatigue life at more than one load level in order to get the slope of the median contour, for which Eq 13 gives the ratio b/m . The observed lifetime distribution shape factor b is estimated from the

observed life distribution by the commonly used Weibull slope, and the strength shape factor is “backed out.”

We apply this model to various types of bearings and illustrate how a reliability forecast chart can be constructed. The number of data is small and the results are variable, as found by Zaretsky [6]. Nevertheless, the model has attractive flexibility and will provide accurate reliability estimations for modern hybrid bearings, given more test data.

Weibull Distribution Considerations

Weibull developed his eponymous distribution to address certain problems in the ceramics industry, which he worked in in Sweden. His concerns were the wide scatter of brittle strength data and the evident size effect, whereby large-sized specimens exhibited significantly lower strengths.

He also considered the fact that brittle strength was governed by various flaws distributed throughout the material, and that a fatal combination of tensile stress and local flaws causes failure. He proposed the survivability form presented in Eq 1 to account for the tensile stress (raised to an exponent m) distributed throughout the material volume under tension

$$S = e^{-\int (\sigma/\sigma_o)^m dV} \quad (1)$$

where:

S = cumulative probability of survival,

σ = tensile stress,

m = a shape factor,

σ_o = a scale factor, and

V = volume subject to tension.

If the material is a unit volume under uniform tension, the survival probability distribution takes the following very simple and commonly used form

$$S = e^{-(\sigma/\sigma_o)^m} \quad (2)$$

This form has been found to be applicable to static fatigue life prediction (stress rupture), RCF life forecasting, and a variety of other applications (see references).

Weibull's work, originally published in 1939 in a Swedish engineering journal [1,2], did not take the world by storm until 1951 with the publication of an article in the *Journal of Applied Mechanics* [3] entitled “A Statistical Distribution Function of Wide Applicability.” That publication led to wide dissemination and usage of the Weibull model on all sorts of data, some far removed from the strength of brittle materials, or even from weakest link mechanisms.

The view has been “if it works, use it,” and one of the first areas to use this form was the fatigue life distribution of rolling element bearings. In both static and dynamic fatigue life, the issue is how to properly incorporate the effect of applied load on the fatigue life.

Weibull General Scale Factor

Before dealing further with the simultaneous effects of load and time on life reliability, it is appropriate to change the Weibull distribution scaling factor. The scaling factor in Eq 2 corresponds to the specific survival probability $S = 1/e = 0.37$, and neither this probability nor the magnitude of the scale factor is a general indicator of the data central value. Although the mean appears to be an attractive scale factor, it is complicated by the fact that the probability corresponding to the mean is a gamma function of the shape factor, as noted in Eq 3

$$S_{\text{mean}} = e^{-\{\Gamma[1+(1/m)]\}^m} \quad (3)$$

where:

Γ = gamma function of the incomplete exponential integral, and
 m = Weibull modulus.

This expression for S_{mean} depends on m and is also neither intuitive nor a convenient normalizing factor.

Use of the median as a scale factor provides an indication of the central value, independent of distribution parameters, and clarifies design safety factors as fractions of the median.

Using an arbitrary reference survival value S_r and a corresponding variable (life or load) in Eq 2

$$\ln(S_r) = -\left(\frac{\sigma_r}{\sigma_o}\right)^m \quad (4)$$

Solving for σ_o in terms of the reference percentiles in Eq 4 and rewriting in terms of the reference percentile leads to

$$S = S_r^{(\sigma/\sigma_r)^m} \quad (5)$$

Equation 5 is an alternative scaled expression of the Weibull distribution, in a less familiar form. For median scaling, S_r is set as 0.5, and the observed median variable σ_{50} is used to scale all the observed data

$$S = 0.5^{(\sigma/\sigma_{50})^m} \quad (6)$$

Equation 6 is the median-normalized Weibull distribution and also leads to the well-known straight line relation for estimating the Weibull shape factor from data

$$\log \left[\frac{\ln(S_i)}{\ln(0.5)} \right] = i \log \left[\frac{\sigma_i}{\sigma_{50}} \right] \quad (7)$$

Note also that Eq 7 provides an estimate of the Weibull modulus from each individual i th observation.

The Appendix lists a number of useful formulations with the median-normalized Weibull distribution and introduces the application of the individual computation of the shape factor from each observed data point. This is addressed in the data analysis and tells us a lot about the data trends, and it can indicate the presence of a lower bound for fatigue, an endurance limit.

The Bivariate *Static Fatigue Reliability Forecast Chart*

In the 1960s, an extensive experimental program was undertaken at the Lawrence Radiation Laboratory (LRL) to measure the stress rupture life of fiber reinforced materials requiring the long-term storage of highly pressurized gases. The composite materials investigated included mainly glass, Kevlar, and carbon fiber composites. The data generated from this test program include many hundreds of glass fiber composite strands subject to several different sustained loads for median life periods ranging from a few minutes to several years, as shown in Fig. 1. Figure 1(a) presents static fatigue life at a load of 479 ksi, 84 % of the strength, and a median life of 14 min (0.23 h) with a shape factor $b=0.9$. Figure 1(b) presents static fatigue life at a load of 426 ksi, 75 % of the strength, and a median life of 6.5 h, with the same shape factor $b=0.9$. Figure 1(c) presents static fatigue life at a load of 373 ksi, 65 % of the strength, and a median life of 513 h, again with the same shape factor $b=0.9$.

The distribution is fitted with a Weibull modulus of 0.9. The accumulated static fatigue data trends were analyzed and used to define a bivariate model (parameters for both life and load) based on the Weibull distribution [1,2]. The heuristic rationale of the model was based on the assumed relation of the survivability to an arbitrary exponential function H of the load (R) and the time (t)

$$\ln(S) \approx H(R, t) \quad (8)$$

Assuming the function H is separable yields

$$\ln(S) = g(R)f(t) \quad (9)$$

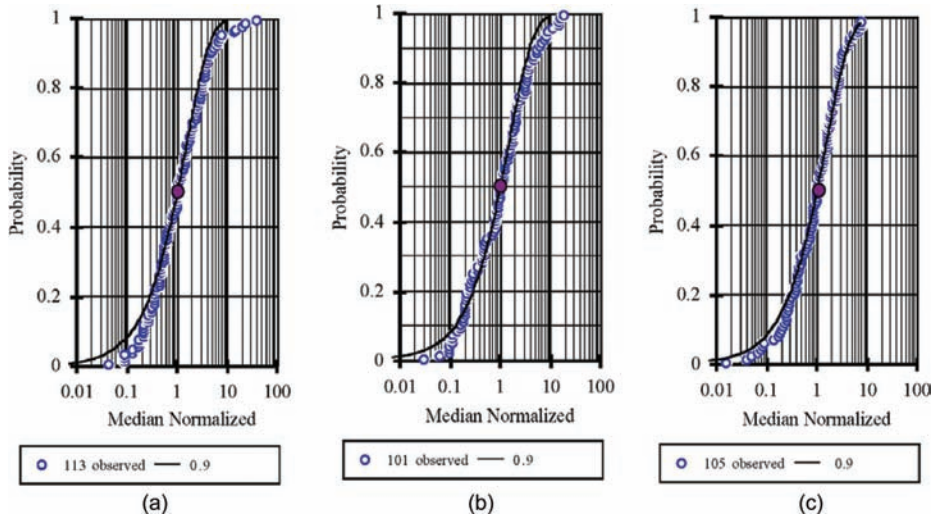


FIG. 1—(a) Static fatigue life at a load of 479 ksi, 84 % of the strength, and a median life of 14 min (0.23 h) with shape factor $b=0.9$. (b) Static fatigue life at a load of 426 ksi, 75 % of the strength, and a median life of 6.5 h, with the same shape factor $b=0.9$. (c) Static fatigue life at a load of 373 ksi, 65 % of the strength, and a median life of 513 h, again with the same shape factor $b=0.9$.

And further, if these functions are exponentials

$$\ln(S) = CR^m t^b \quad (10)$$

The bivariate median-normalized model is derived as illustrated for Weibull's original, single variate form Eq 5, by eliminating C

$$S = 0.5(R/R_{50})^m (t/t_{50})^b \quad (11)$$

Weibull's scale factors, t_o or σ_o , called characteristic values, are replaced by the observed median data R_{50} for strength and t_{50} for life. This is a convenient form of the normalized Weibull distribution. The median is the chosen scaling factor, determined directly from the center of any set of data as a central value that is distribution free and statistically reliable.

In fact, however, any percentile may be used as the normalizing factor; this is useful when a small number of failures have been observed in a large parent population. With S_r as the survival probability at the normalizing factor, the general normalized Weibull distribution is the double exponential

$$S = S_r (R/R_r)^m (t/t_r)^b \quad (12)$$

This is the universal fatigue model, applicable to both static and dynamic fatigue. The LRL static fatigue data consider several load levels with hundreds of stress rupture times from minutes to several years. Each data group, comprising about 100 life observations, was scaled to its median, and the normalized distributions were compared (refer to Fig. 1). These remarkable data gave clear evidence that the distribution shape is of the Weibull form and, for glass fiber/epoxy composite strands (Gl/Ep), shows *invariance* with time. This shape invariance ranges over median lifetimes of 15 min to far more than 513 h. A partial group of lifetimes was observed at 58 % load with an estimated median life of 113 000 h (~ 12.9 years). Probing the data via a graph displaying individual estimates of the Weibull life shape factor b (using Eq 7), displayed in Fig. 2, verifies the existence of an essentially invariant shape characteristic, indicating an invariant Weibull modulus throughout the data. An analysis of Kevlar-epoxy strand static fatigue data in the Appendix shows very different characteristics via this graphical display and leads to a hypothetical static fatigue model to account for overarching progressive damage, or corrosion, causing reduced dynamic fatigue life.

In log form, the static fatigue model of Eq 12 underlies the RFC, which plots log of load versus log of life shown in Fig. 3

$$\log\left(\frac{R}{R_r}\right) = -\left(\frac{b}{m}\right) \log\left(\frac{t}{t_r}\right) + \left(\frac{1}{m}\right) \left[\log \left[\frac{\ln\left(\frac{1}{S}\right)}{\ln\left(\frac{1}{S_r}\right)} \right] \right] \quad (13)$$

Figure 3 shows an enormous amount of data at three load levels, with smaller numbers at still lower loads (and longer life). The life shape factor as derived

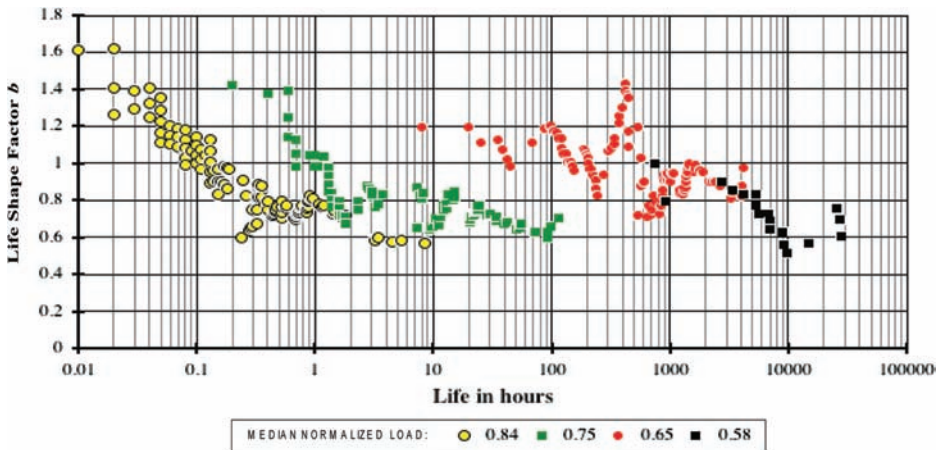


FIG. 2—Shape factor b trends for Gl/Ep composite static fatigue.

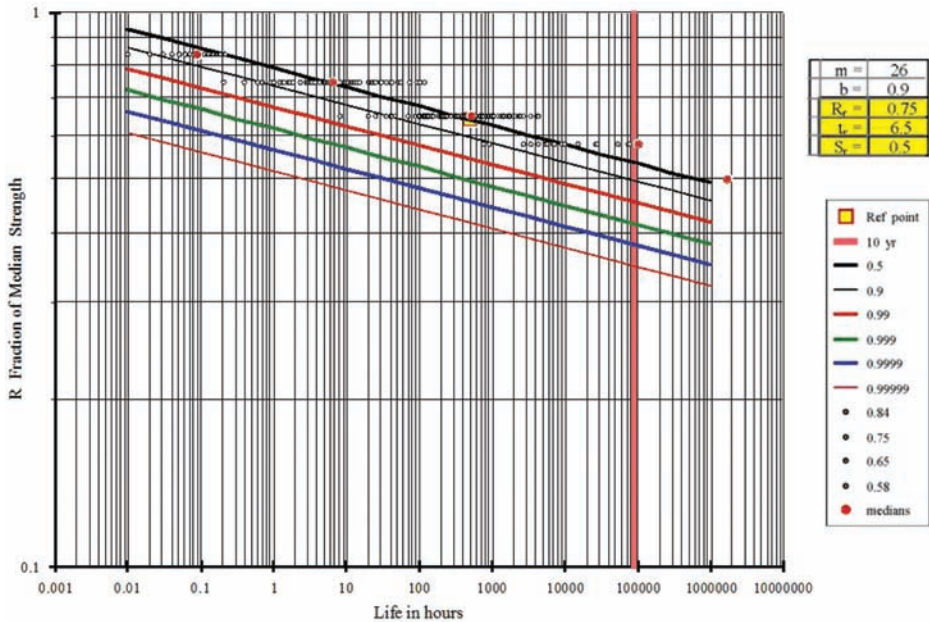


FIG. 3—Static fatigue reliability forecast chart for Gl/Ep composite.

in Fig. 1 is $b = 0.9$. The strength shape factor, $m = 26$, is derived from the slope of the observed probability contours in Fig. 3.

In order to construct the chart, data are needed at two or more different load levels (in this case, five load levels) so as to establish the slope of the probability contours. This slope is given by $-b/m$ in Eq 13 and is the ratio of the underlying shape factors. The individual median lives also define the slope of the 50 % contour and are another approach to the observed slope. As its name implies, this is a forecast chart that gives a comprehensive view of the data, the reliability contours, and the allowable loads for achieving the desired reliability at a given life. An “anchor point,” or reference point, is needed, such as in Fig. 3; $R_r = 0.65$, $t_r = 513$ h, and $S_r = 0.5$, which is the median for the $R = 0.65$ group, as shown in Fig. 1 and plotted in Fig. 3.

All-steel Bearings

In the search for very large amounts of bearing fatigue data, the extensive survey by Lieblein and Zellen [9] was examined. They undertook a very comprehensive review of deep groove all-steel rolling element bearings acquired from several bearing manufacturers, amounting to 213 total test groups, with the year of manufacture ranging from 1936 to 1951. The size of these groups varied but typically included 12 to 30 bearings per test group, amounting to several thousand bearings being tested for fatigue. The Weibull modulus was estimated for each test group by the ratio of L50 (the life that half the

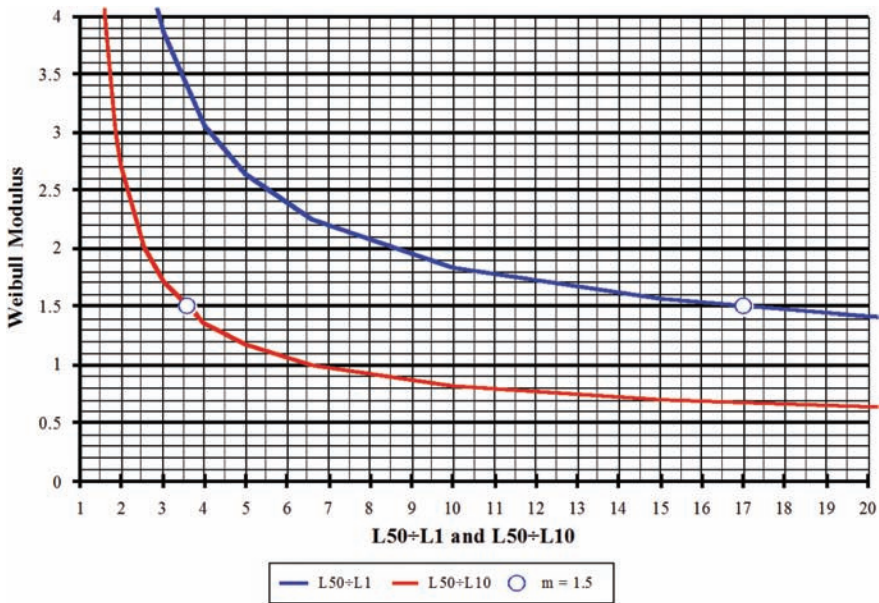


FIG. 4—Estimating the Weibull modulus from $L50/L10$ and $L50/L1$ ratios.

population survive) to L10 (life that 90 % survive), in accordance with Fig. 4, in which L1 is the life that 99 % survive.

Figure 5 shows a large variation of the Weibull modulus, computed via the L50 to L10 ratio, ranging from 0.5 to 4.5, with an overall median of 1.4. This scatter runs throughout the data in Ref 9 among similar bearing sizes, similar test conditions, and different types of bearings. Such a variation of shape factors calls into question the utility of the data and can have severe impact on long-term forecasts of very high reliability. The pooled median-normalized data shown in Fig. 5(b) do not represent a classic Weibull shape. The design choices are high design reliability (i.e., a high factor of safety), or extensive

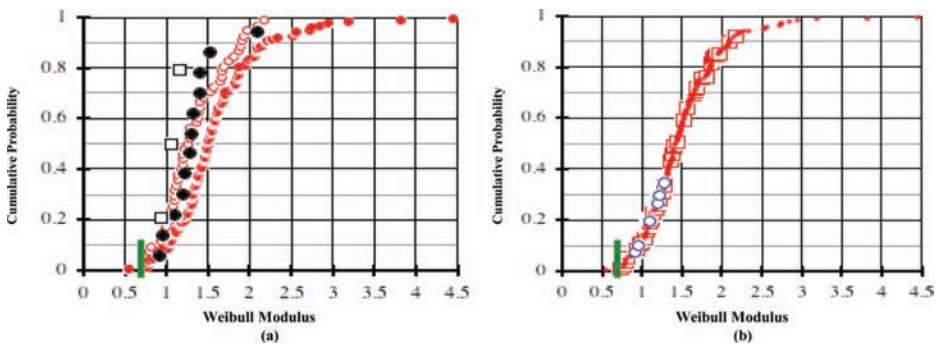


FIG. 5—All-steel RCF fatigue data. (a) All-steel fatigue data from four companies [9]. (b) Pooled and median-normalized data from (a).

testing of the operational bearing under realistic conditions, to define a consistent reliability forecast based on observed operational data.

In more recent fatigue testing of all-steel bearings (52100) at Aerospace, two groups of fatigue failures, tested years apart, were observed, and both data sets were found to fit a Weibull modulus of 0.7, as shown in Fig. 6. Despite this shape consistency, the shape factor $b = 0.7$ is at the extreme low end of the large Lieblein and Zellen database, as shown by the vertical bar in Fig. 5.

As part of the Aerospace RCF data analysis, size 101 52100 steel bearings were tested in two different rigs, several years apart, in order to compare the statistics of fatigue life. The results, shown in Fig. 6, were consistent and somewhat surprising. Both sets of data were well fitted with a Weibull life shape factor $b = 0.7$, a value at the extreme low end of the Lieblein–Zellen database. Figure 6(a) shows the individual data and the pooled data. The life difference shown in Fig. 6(a) between old and new data, although consistent, is not regarded as significant. The data show consistent life distributions, even though tests were conducted years apart on different test rigs. The individual life shape factors and the pooled data shape factor are both well represented by the $b = 0.7$ shape factor.

Because the design concern is with early failure at low probability, placing as large a number of bearings as possible under test and analyzing the consequences of the early failures for the lifetime distribution can conserve testing time. With this approach it is not necessary to test more than half a test population to failure, and it is possible that not even that many will be needed. As an example, if we had only the first 6 failures (out of 11 on test), the estimated shape factor would still be 0.7.

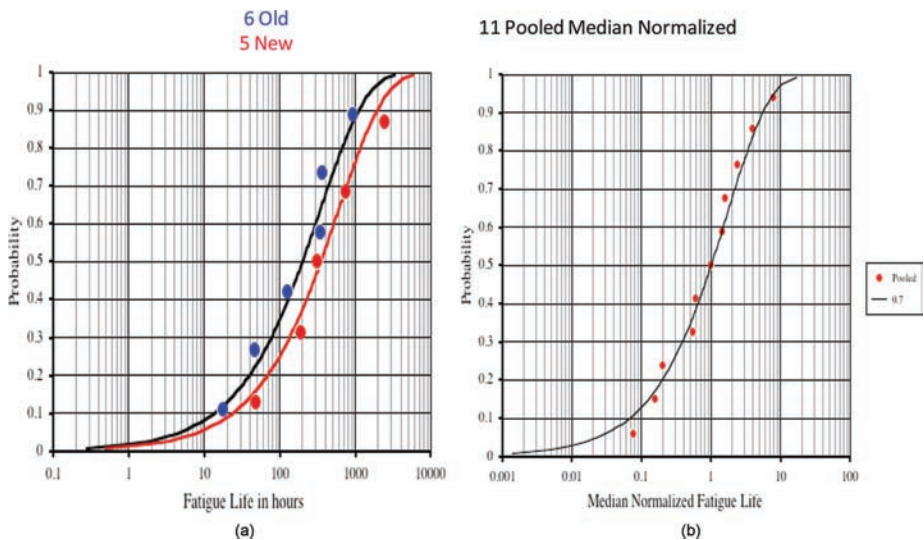


FIG. 6—(a) Test machine comparison with all-steel size 101 bearings at 400 ksi, 5000 rpm. (b) Pooled data median, normalized, also fitted with $b = 0.7$.

Silicon Nitride Ball Hybrid Bearing Fatigue and Reliability Forecast Chart

Some fatigue tests of size 207H silicon nitride ball hybrid bearings (SNHBs), described in Fig. 7, provide data at two load levels that allow the construction of an RFC.

At a 1500 lb load (270 ksi), six sets of four bearings survived for 2500 h without failure. Assuming that this is an expected outcome (a 50-50 chance) implies that the reliability for each set of four bearings at these conditions is $\Gamma 0.5^{(1/6)} = 0.89$, giving one point on the forecast chart

$$S = 0.89, \quad R = 270 \text{ ksi}, \quad t = 2500 \text{ h}$$

As noted in Fig. 7, the six “no fail” observations give a point on the RFC of 2500 h at 270 ksi with a survivability of $0.5^{(1/6)} = 0.89$. The reference strength of silicon nitride balls is taken as 540 ksi.

Four sets of four of these 207H bearings were sudden-death fatigue tested at 3000 lb (335 ksi) and 5000 rpm, as shown in Fig. 7

$$1620, 2376, 3672, 3715 \text{ h; median} = 3024 \text{ h}$$

The probability distribution for these SNHB data is shown in Fig. 8, in which the failure probabilities are the median rank, computed as recommended by Johnson [10], for rank i and sample size N , by

$$F_i = 1 - S_i = (i - 0.3)/(N + 0.4)$$

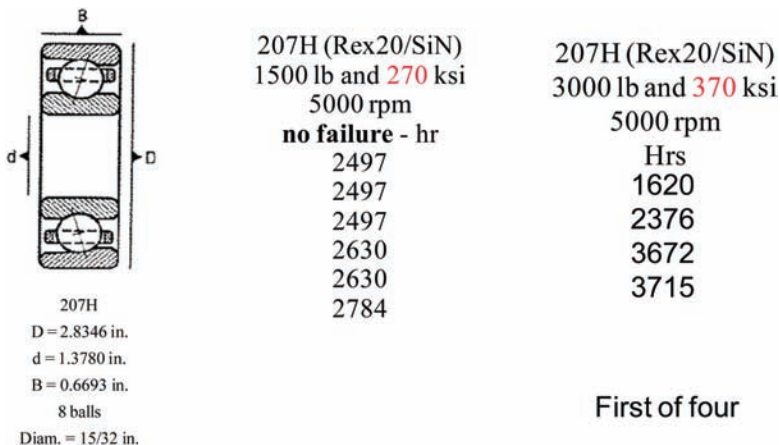


FIG. 7—SNHB size 207H rolling contact fatigue tests (1998); sudden-death, groups of four.

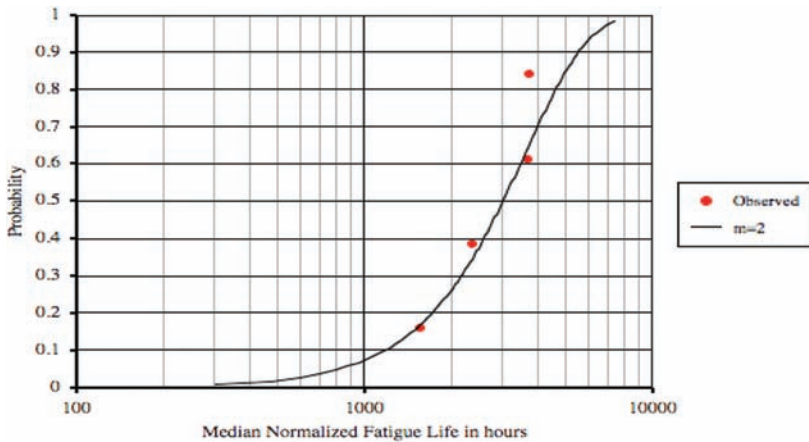


FIG. 8—Fatigue life of groups of four 207H bearings tested (sudden-death) at a thrust load of 3000 lb (375 ksi), having a median life of 3024 h at 5000 rpm.

and the Weibull shape factor was estimated (as in Fig. 8) from curve fits as $b = 2$.

This gives a reference point for the forecast chart in Fig. 9 as

$$S = 0.5, \quad R = 335 \text{ ksi}, \quad t_{50} = 3024 \text{ h}, \quad b = 2$$

The RFC (Fig. 9) shows the 50 % contour passing through the median at 335 ksi (least-of-four) survival probability at 0.89, consisting of 270 ksi (50 % of median strength) and six No Fail, at 2500 h, so $S \geq 0.89$, shown just short of the $S = 0.9$ contour. Using Eq 13 and altering the input m to get the best fit to the slope produces $m = 7$ and $-b/m = 0.28$ for the load-life exponent. Figure 9 provides a rational prediction of reliability versus load and life for the 207 H hybrid bearings under test. Although the data are sparse, this chart predicts 99 % reliability at about 12 % of the reference strength for a service period of over 10 years.

Another Forecast Chart Estimate for Silicon Nitride Balls

In Ref 8, Parker and Zaretsky provide experimentally measured fatigue life distribution data on silicon nitride balls from tests in a ball fatigue test machine conducted at two loads (shown in Fig. 10). Data are provided at 620 ksi and 800 ksi. Both loads show similar Weibull slopes ($b \approx 1$) for life. One data set includes the median, and the other tests reach about 30 %.

A fatigue RFC can be constructed as follows:

Reference point: load $R_r = 0.8 \text{ msi}$, life $t_{50} = 18 \times 10^6$ cycles, probability $S_r = 0.5$.

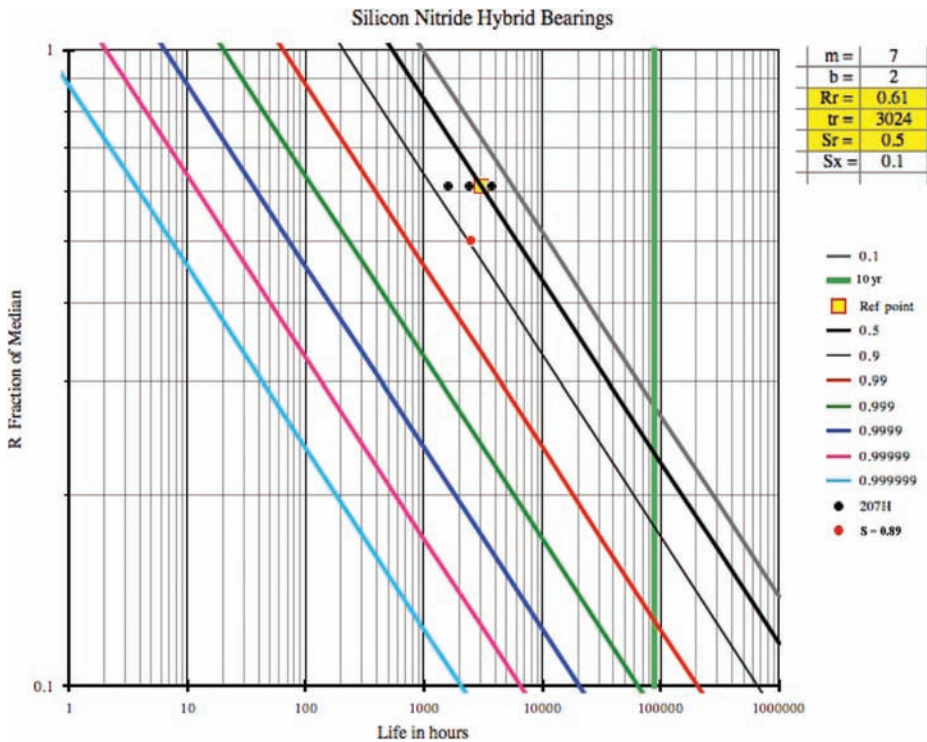


FIG. 9—Reliability forecast chart for 207H SNHBs.

The life values at 10 % probability are as follows:

At 0.80 msi (hertz stress), life = 2.7 (million cycles); $S(L10) = 0.9$.

At 0.62 msi (hertz stress), life = 180 (million cycles); $S(L10) = 0.9$.

With the two points plotted and $b \approx 1$, Fig. 11 shows that a good fit to both points is provided by $m = 16$, giving a load-life exponent $m/b = 16$.

These last two efforts at forecast charts for SNHBs exhibit very different parameters. One suffers from a small database (four fatigue failures), and the other draws on old silicon nitride data from balls in a ball test machine (see Figs. 10 and 12).

The 19 data points for the test group at 800 ksi hertz stress are plotted in Fig. 11 in order to estimate the shape factor. The graph of Weibull distributions with a shape factor $b = 1$ shows a good enough fit, at least below the median.

Modeling Progressive Damage in Rolling Contact Fatigue

We address here the static fatigue data obtained with Kevlar composite strands (see Fig. 13). These data, at first glance similar to the glass data in Fig. 1, are in fact dramatically different. The display of the individual Weibull shape factors (shown in Fig. 14) exhibits a progressive and sharply increasing shape factor

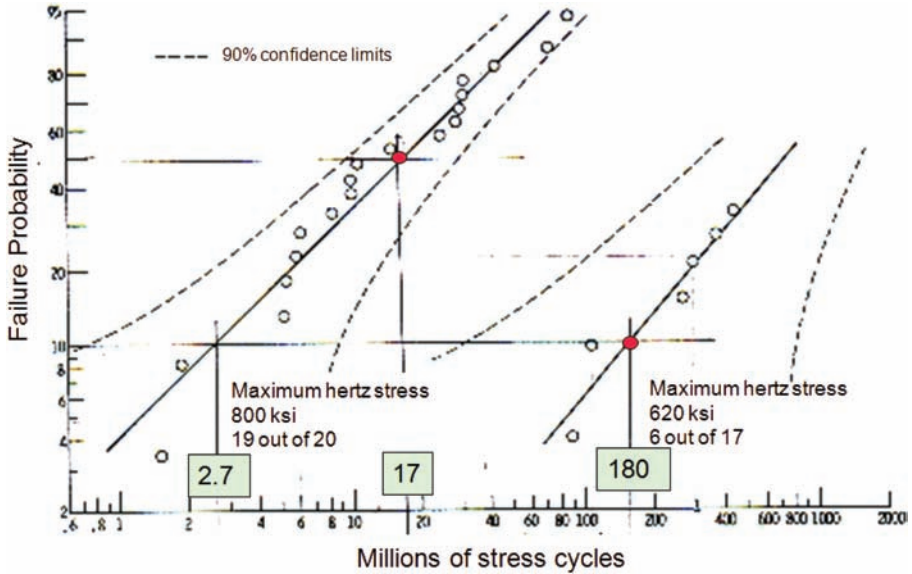


FIG. 10—Silicon nitride balls in five-ball test machine, tested at hertz stresses of 800 ksi and 620 ksi, showing $b = 1$ (after Zaretsky).

(reduced scatter): with increasing lifetimes, the shape factor is increasing. This is understood in terms of a falling off of the probability contours that causes them to bunch together and reflect a lower scatter. For the case of static fatigue, as well as for RCF, it is presumed that corrosion degrades the strength of a susceptible portion of the bearing surfaces and, consequently, the fatigue life. The corrosion is taken as a first-order degradation process with a time constant t_c . In terms of the equation for the forecast chart, this leads to the following plausible model

$$R = \{1 - f[1 - \exp(t/t_c)]\} * R_r(t_r/t)^{(e/m)} [\ln(1/S)/\ln 2]^{(1/m)} \quad (14)$$

where:

f = fraction of surface material affected by degradation, and
 t_c = chemical corrosion time constant.

Invoking these two additional parameters leads to the forecast chart of F , which fits the data quite well and calls for similar observations with RCF life.

Such a model might be useful in addressing cases in which bearings are progressively degraded, the probability contours bend down, and the lifetimes are cut short.

Assessment and Conclusions

- There is large variability in RCF life shape factor, and therefore there is no simple model that can be used to rationalize all of the observed test results.

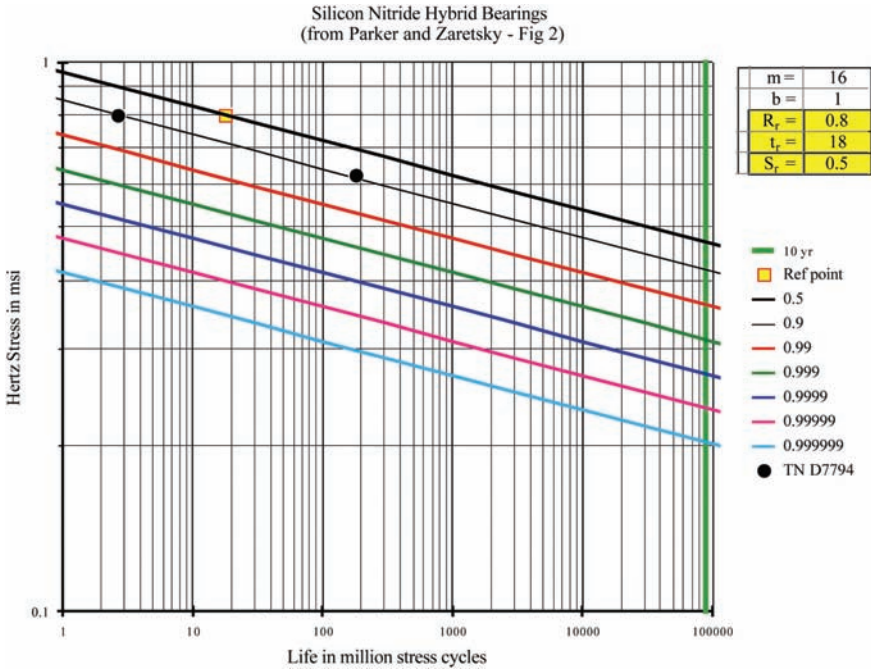


FIG. 11—Reliability forecast chart for silicon nitride balls.

- There is no large database of SNHB RCF that identifies consistent quantitative characteristics, and the experimental data (fatigue life, Weibull modulus, and load-life exponent) vary over a significant range. Consequently, bearing fatigue test conditions need to emulate the operational conditions, and even if this condition is met, high margins and redundancy will be called upon to minimize the risk of fatigue failure in operation [11–19].
 - The static fatigue model below can be usefully applied to RCF data to produce rational and comprehensive reliability estimates.
- In the double exponential form (bivariate Weibull distribution) of Eq 11

$$S = 0.5^{(R/R_{50})^m (t/t_{50})^b}$$

In the form for constructing the RFC, Eq 13

$$\log\left(\frac{R}{R_r}\right) = -\left(\frac{b}{m}\right) \log\left(\frac{t}{t_r}\right) + \left(\frac{1}{m}\right) \left[\log\left[\frac{\ln\left(\frac{1}{S}\right)}{\ln\left(\frac{1}{S_r}\right)} \right] \right]$$

- In order to construct the RFC, fatigue data are needed at two or more loads to establish the slope of the probability contours (b/m), or m is

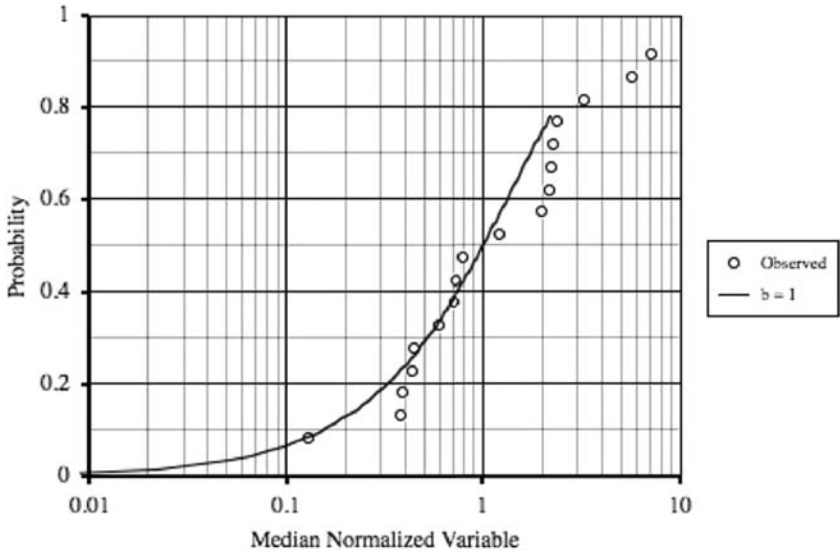


FIG. 12—Fatigue life distribution of silicon nitride balls.

derived from strength statistics of the silicon nitride material. The two points allow definition of the forecast chart and inference of the load-life exponent. The life shape factor is estimated by fitting a Weibull distribution to the lifetime data.

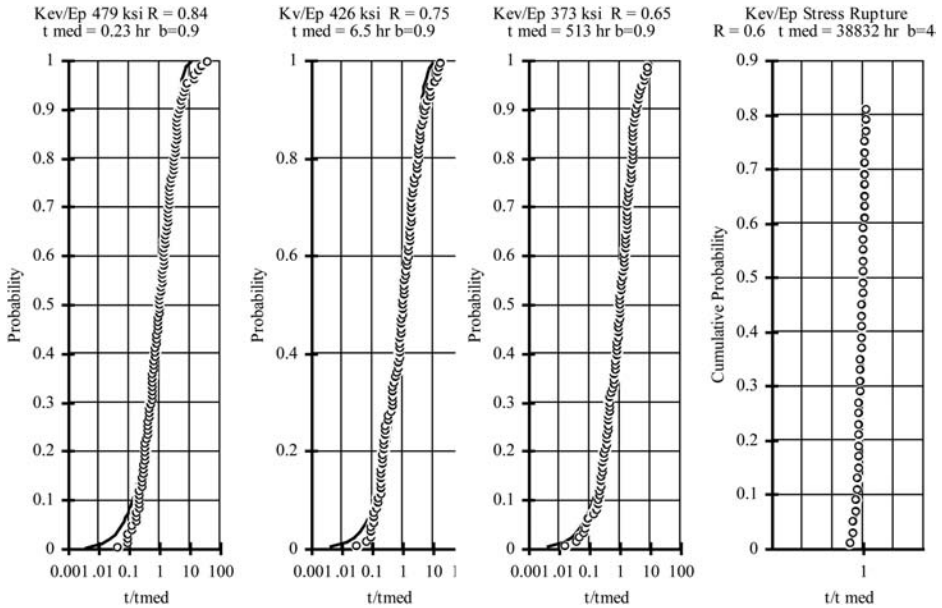


FIG. 13—Kevlar static fatigue data. Kev/Ep, Kevlar/epoxy composite.

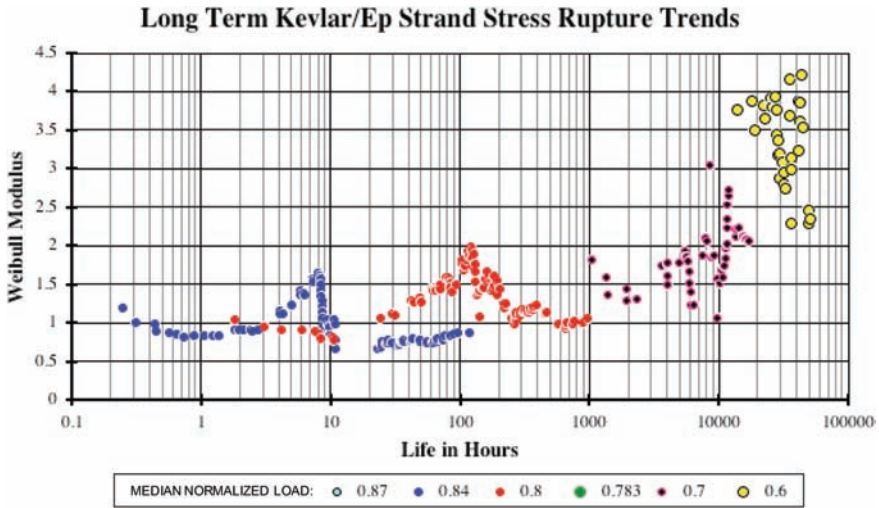


FIG. 14—Shape factor trends for Kevlar/epoxy composite.

- The model for progressive damage in static fatigue can be a useful guide for defining the statistical effect of progressive damage in RCF, i.e., a turndown of the probability contours. A hypothetical model that works for static fatigue might be adaptable to such a progressive overarching damage process

$$R = \{1 - f[1 - \exp(t/t_c)]\} * R_r(t_r/t)(e/m)[\ln(1/S)/\ln 2](1/m)$$

- The forecast chart is recommended as a potentially useful tool giving a comprehensive overview of the data, their reliability, and their service life in design.
- The plots of the individual life shape factors (b) versus life provide a useful early indication of a trend in endurance limit or extrapolated time invariance of the failure statistic, or the onset of an overarching damage mechanism such as corrosion.

Appendix: Useful Weibull Formulations

The Weibull modulus is denoted in this paper by e (Weibull slope or modulus) or b , as used in an early paper by the author [7]. The designation for the brittle material strength shape factor is m , as used by Weibull. Some useful formulations presented here were used in analyzing the reference static fatigue database and the bivariate model for some dynamic fatigue data.

Estimating Weibull Modulus from Life Ratios

Figure 4 provides another basis for estimating shape factors based on life ratios and illustrates the method used by Lieblein and Zellen [9].

Design Reliability Factors

Given a desired design reliability S_d , the corresponding fraction of median life is given as

$$\frac{X_{50}}{X_d} = \left[\frac{\ln(2)}{\ln\left(\frac{1}{S_d}\right)} \right]^{1/e} \quad (\text{A1})$$

Estimating Shape Factor from Each Observation

Each scaled and ranked observation X_i may be used to estimate the Weibull modulus

$$m_i = \ln \left[\frac{\ln\left(\frac{1}{S_i}\right)}{\ln(2)} \right] \quad (\text{A2})$$

where:

S_i = survival probability of the observed value, and

X_i = observed life of the i th ranked specimen.

$S_i = (I - 0.3)/(N + 0.4)$, the median rank plotting position [10], in which i = rank and N = total number.

Figure 2 provides a plot of these shape factor trends with life for the LRL glass strand static fatigue data including only the glass composite data. Three possible trends are expected, in principle:

1. horizontal (invariant) trend if the invariant life shape continues to hold to longer life (e.g., Fig. 2),
2. a decreasing magnitude of the shape factor trend (progressively higher scatter) if an asymptote is approached, causing longer lives (higher scatter and lower shape factor) to be observed, and
3. an increasing shape factor at long life, indicating a bunching together of the probability contours, causing increased shape factor and indicative of a turndown caused by accelerating damage and shortened life (e.g., Fig. 14).

Such analysis and observed trends might be expected with RCF, although causes of failure under dynamic fatigue are different and much more complex than with static fatigue.

Figure 15 shows the forecast chart with the contours falling off at long times accompanied by the increased Weibull modulus. This accelerating damage process is discussed and rationalized by a modified forecast model. For these data, the parametric assumption in the damage model was an increasing susceptibility to damage, based on a first-order chemical damage process with a time constant of 150 000 h (see Fig. 15).

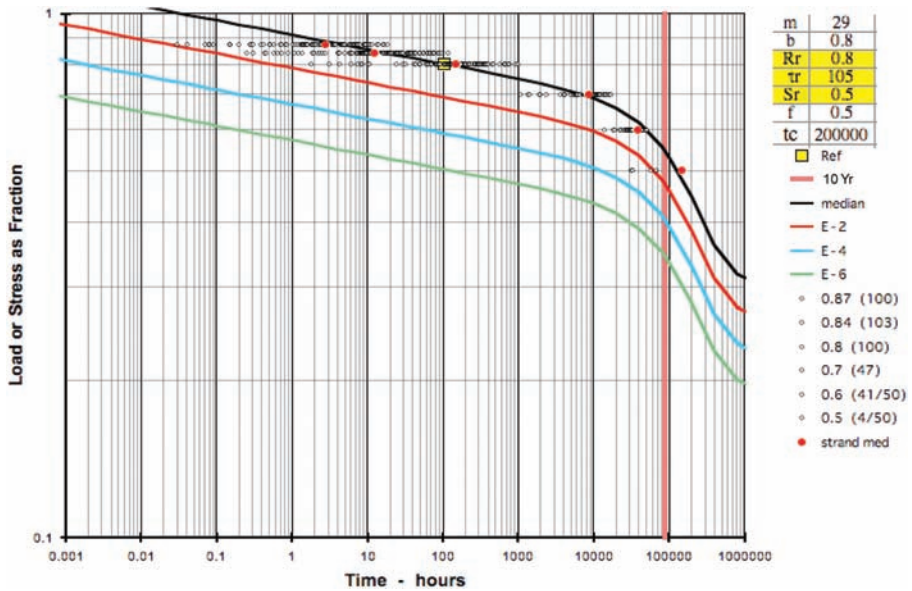


FIG. 15—Kevlar/epoxy composite stress rupture showing contour downtrend and decreased scatter (see Fig. 14).

Acknowledgments

The writer acknowledges the technical and administrative support provided by Dr. Michael Hilton, Head of Mechanical Systems Department, and the support of Dr. Yoshimi Takeuchi, who was the conference chair of the 2011 ASTM International Symposium on rolling component bearing fatigue, where a part of this paper was presented.

References

- [1] Weibull, W., "A Statistical Theory of the Strength of Materials," Report No. 151, *Ingeniörs Vetenskaps Akademiens Handligar [Proceedings of the Royal Swedish Academy of Engineering]*, Royal Swedish Academy of Engineering, Stockholm, Sweden, 1939.
- [2] Weibull, W., "The Phenomenon of Rupture in Solids," Report No. 153, *Ingeniörs Vetenskaps Akademiens Handligar [Proceedings of the Royal Swedish Academy of Engineering]*, Royal Swedish Academy of Engineering, Stockholm, Sweden, 1939.
- [3] Weibull, W., "A Statistical Distribution Function of Wide Applicability," *J. Appl. Mech.*, Vol. 18, 1951, p. 293.
- [4] Lundberg, A. G. and Palmgren, A., "Dynamic Capacity of Rolling Bearings," Report No. 196, *Ingeniörs Vetenskaps Akademiens Handligar*

- [Proceedings of the Royal Swedish Academy of Engineering], Royal Swedish Academy of Engineering, Stockholm, Sweden, 1947.
- [5] Zaretsky, E. V., "A. Palmgren Revisited. A Basis for Bearing Life Prediction," *NASA TM 107440*, National Aeronautics and Space Administration, Lewis Research Center, Cleveland, Ohio, 1997.
 - [6] Zaretsky, E. V., Poplawski, J. V., and Peters S. M., "Rolling Bearing Life Prediction: Past, Present, and Future," *NASA TM-2000-210529*, National Aeronautics and Space Administration, Lewis Research Center, Cleveland, Ohio, 2000.
 - [7] Robinson, E. Y., "A Statistical Model for Static Fatigue of Multi-filament Strands," *Fibre Sci. Technol.*, Vol. 2, 1970, pp. 177–192.
 - [8] Parker, R. J., and Zaretsky, E. V., "Rolling-element Fatigue Life of Silicon Nitride Balls," *NASA TN D-7794*, National Aeronautics and Space Administration, Lewis Research Center, Cleveland, Ohio, 1974.
 - [9] Lieblein, J. and Zellen, M., "Statistical Investigation of the Fatigue Life of Deep-groove Ball Bearings," *J. Res. Natl. Bur. Stand.*, Vol. 57, No. 5, 1956.
 - [10] Johnson, L. G., *The Median Ranks of Sample Values in Engineering Experiments*, McGraw-Hill, New York, 1973.
 - [11] Lundberg, A. G., and Palmgren, A., "Dynamic Capacity of Roller Bearings," Report No. 210, *Ingenjörers Vetenskaps Akademiens Handligar [Proceedings of the Royal Swedish Academy of Engineering]*, Royal Swedish Academy of Engineering, Stockholm, Sweden, 1952.
 - [12] Tallian, T., "Weibull Distribution of RCF Life and Deviations Therefrom," *ASLE Transactions*, Vol. 5, 1962, pp. 183–196.
 - [13] Chiao, T. T. and Moore, R. L., "Stress-rupture of S-glass Epoxy Multifilament Strands," *J. Compos. Mater.*, Vol. 5, 1971, p. 2.
 - [14] Robinson, E. Y., "Stress Rupture Model for Design with Brittle Materials," AIAA 97-2723, *33rd AIAA/ASME/ASEE Joint Propulsion Conference*, Seattle, WA, July 6-9, 1997.
 - [15] O'Brien, M. J., Presser, N., and Robinson, E. Y., "Failure Analysis of Three Si₃N₄ Balls Used in Hybrid Bearings," *Eng. Failure Anal.*, Vol. 10, 2003, pp. 453–473.
 - [16] Robinson, E. Y., "Reliability Analysis and Life Extension Forecast of STS-Orbiter Kevlar COPV. An Independent Assessment," *ATR-2005(5146-88)-1*, The Aerospace Corporation, El Segundo, CA, 2005.
 - [17] Harris, T. A., and Kotzalas, M. N., *Rolling Bearing Analysis: Essential Concepts of Bearing Technology*, 5th ed., CRC Press, Boca Raton, FL, 2006.
 - [18] Robinson, E. Y., "Fatigue Life Forecast of Silicon Nitride Hybrid Bearings," *ATR-2008(8257)-1*, The Aerospace Corporation, El Segundo, CA, 2007.
 - [19] Robinson, E. Y., and Kohli, R., "Preliminary Risk of Rupture Assessment for Gr/Ep Composite Overwrapped Vessels on the International Space Station," *ATR-2009(5298)-6*, The Aerospace Corporation, El Segundo, CA, 2009.

J. H. Cowles, Jr.¹ and C. A. Houle²

Roller Profile Development for an Axially Loaded, Single Row Spherical Roller Bearing in an Oscillating Application

REFERENCE: Cowles, J. H., Jr. and Houle, C. A., "Roller Profile Development for an Axially Loaded, Single Row Spherical Roller Bearing in an Oscillating Application," *Rolling Element Bearings* on April 13–15, 2011 in Anaheim, CA; STP 1542, Yoshimi R. Takeuchi and William F. Mandler, Editors, pp. 47–70, doi:10.1520/STP103891, ASTM International, West Conshohocken, PA 2012.

ABSTRACT: Typically it is not recommended to use a single row spherical roller bearing in an application in which there are significant axial loads. The use of these bearings in a less than optimal customer application can be driven by the need for a high load capacity within a limited geometric envelope. Specialized roller profiles can be developed that eliminate high contact stresses on the roller ends that can lead to the premature failure of the bearing. In this paper, analytical results are presented for the roller to raceway contact stress of four different roller profiles with a given load applied to the bearing. The profiles considered are single transverse radius, double transverse radii, and optimized symmetric and asymmetric. The piecewise logarithmic profiles developed are applied to a barrel roller in this paper, but a generalized method is developed such that the profile could be applied to any type of roller—cylindrical, barrel, hourglass, or tapered. The analysis results for two of the profiles are validated by dynamic testing performed within RBC's Corporate Test Laboratory on a three axis, hydraulic, computer controlled test rig. The bearing inner ring is oscillated while the bearing is subjected to simultaneous radial and axial loads. The control software allows the monitoring of radial displacement and bearing torque during testing. The test results correlate very well with the analytical results. The displacement and torque readouts are shown to predict the onset of raceway fatigue and spalling for one of the roller configurations before testing is complete and the bearing is disassembled.

Manuscript received April 4, 2011; accepted for publication November 14, 2011; published online December 2011.

Corporate Headquarters, RBC Bearings, 1 Tribology Center, Oxford, CT 06478

Aircraft Products Div., RBC Bearings, 2788 Winsted Rd., Torrington, CT 06790

Copyright © 2012 by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

KEYWORDS: spherical roller bearing, roller profile, logarithmic profile, edge loading, numerical analysis, optimization, pressure distribution, contact analysis, test, testing, oscillation, oscillating

Nomenclature

- A = constant of a generalized Lundberg logarithmic profile
 A_1 = constant of a first generalized logarithmic profile
 A_2 = constant of a second generalized logarithmic profile
 d = roller diameter where roller drop is zero
 E = Young's modulus
 γ = angular misalignment of roller axis to raceway
 L_1 = axial length of major transverse roller radius
 L_2 = axial length to secondary transverse roller radius and corner radius intersection
 L_3 = total axial length of roller
 L_b = total length of the contact patch (or blanket)
 L_{eff} = effective length of roller
 ν = Poisson's ratio
 Q = roller load
 r_i = inner raceway radius
 r_o = outer raceway radius
 R_1 = major transverse roller radius
 R_2 = secondary transverse roller radius
 R_3 = roller corner radius
 x = distance along roller axial length from roller or contact patch center
 x' = distance defined along a roller profile
 x'_f = distance along a roller profile where the profile drops are zero
 x'_{2f} = distance along a roller profile from the beginning of the second piecewise roller profile (where $y' = 0$) to the at rest point of profile tangency to the raceway
 y = roller drop as function of x
 y' = roller profile drop as a function of x'
 y'_{2f} = calculated profile drop of the at rest point of second piecewise roller profile tangency to the raceway
 $y_i(x)$ = calculated portion of the roller drop that would cause the roller to conform perfectly to the contacting raceway
 z = effective length of a generalized logarithmic profile
 z_1 = effective length of a first generalized logarithmic profile
 z_2 = effective length of a second generalized logarithmic profile

Introduction

This paper was born out of the practical need to solve a problem encountered while performing the initial testing of a single row spherical roller bearing subject to a constant radial and axial load within an oscillating application. The production order for this part depended upon the successful completion of the

required testing without evidence of raceway spalling. Initial testing of the bearing with a traditional barrel roller design using a single radius transverse roller profile resulted in an end stress condition and premature failure of the part due to outer ring spalling.

The papers of Hertz [1,2] are certainly the most well known, studied, and quoted treatment on the subject of contact stress between elastic bodies. His contribution to the bearing industry as a whole cannot be underestimated. Hertz used a half-space assumption to solve the elastic field equations in terms of potential functions, assuming an elliptical contact area. Lundberg [3] further built upon the Hertzian analysis and, again using a half-space assumption, developed a logarithmic profile that results in a perfectly uniform state of stress across a roller that has no misalignment with the raceway.

The next significant advance in the analysis of rolling element bearings began with the advent of generally available computers, which enabled solution of the half-space problem numerically. A numerical solution enabled the calculation of contact stresses between arbitrarily shaped bodies without simple mathematical descriptions. The works of Hartnett [4,5] are extremely important papers written at a time when numerical computation was coming of age. He developed a numerical procedure incorporating the flexibility method [6] of analysis in conjunction with the Boussinesq [7] half-space force-deflection relationships. The development of a linear system of equations is described, and those equations are then efficiently solved using the Gauss–Siedel [8] method. The method of Hartnett, 30 years later, is still held in high regard, being referenced in works such as Ref 9, ISO 281 [10], and ISO/TS 16281 [11], and being implemented as an option within the COBRA [12] bearing analysis code.

Much work has been done to clarify the effect of a roller profile on the bearing life. A very good summary is presented by Poplawski et al. [13]. A first order improvement in life can come from minimizing the maximum contact stress between roller and raceway, especially in applications like the current one that have relatively short life cycle requirements of tens of thousands of oscillations, rather than millions of revolutions. Minimization of the maximum contact stress via the use of variants of the Lundberg profile for cylindrical rollers was demonstrated by Losche [14] and, more recently, Fujiwara et al. [15,16], who developed a profile optimization procedure using the method of Rosenbrock.

In the work presented here, the analysis procedure of Hartnett was implemented and used to predict that the initial single radius barrel roller design, which failed prematurely, did in fact exhibit an end stress condition. The practical solution to the end stress condition was to develop a double radius profile that is simple to manufacture and measure. Analysis and testing were done in order to verify the successful implementation of this new profile, using the same bearing envelope dimensions and major roller diameter. Further, a numerical optimization procedure was developed in order to allow both a

symmetric and an asymmetric optimized roller profile for this roller and outer ring geometry.

The symmetric and asymmetric optimized roller profiles are characterized by roller drops that are defined relative to the maximum roller diameter. A roller drop is defined as half the reduction in roller diameter measured at any axial location along the roller axis (see Fig. 4) relative to the maximum diameter of the roller. The drop profiles consist of two, potentially different logarithmic profile functions connected by a variable length of zero drops from the perfect radius at the center of contact pressure. The form of the logarithmic profile chosen forces the roller profile to be smoothly varying; the connection points of the three sections of the profile are forced to have continuous derivatives. When the logarithmic profiles are used, there are only four variables to be optimized, so the minimization algorithm is computationally efficient.

The optimization procedure developed utilizes the conjugate gradient method coupled with a line search algorithm outlined in Ref 17.

Problem Description and Test Setup

A cross-section of the spherical roller bearing is shown in Fig. 1. It has a spherical outer raceway and a flanged inner raceway. The outer raceway has a radius r_o ; the inner raceway has a radius r_i . The bearing envelope dimensions, required loading, and oscillation cycles are listed in Table 1.

The test rig used to validate the bearing design is an MTS 322 frame with a 490 kN vertical actuator, a 49 kN horizontal actuator, and a rotary actuator capable of $\pm 45^\circ$ of oscillation at 2300 N-m. A cross-section of the test hardware is shown in Fig. 2.

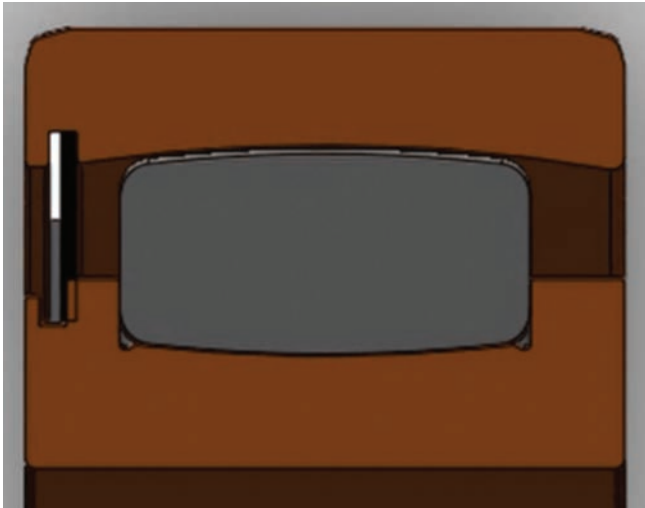


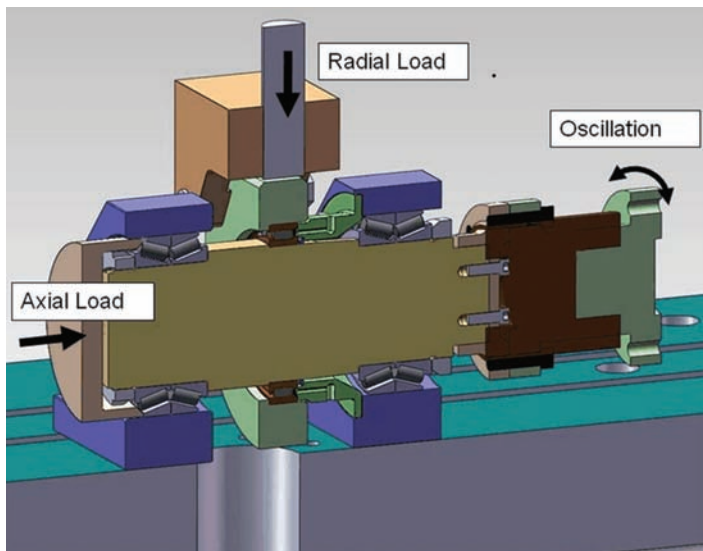
FIG. 1—*Spherical roller bearing cross-section.*

TABLE 1—*Bearing size and test parameters.*

Bearing Parameter	Value
Bore (inner diameter), mm	159
Outer diameter, mm	210
Width, mm	35
Radial load, N	133 450
Axial load, N	13 345
Oscillation angle, deg	± 45
Oscillation speed, cycles/min	25
Required oscillations	60 000
Grease interval, cycles	4000

In order to ensure that the axial load was fairly applied, the following protocol was used at test start-up. First, the shaft oscillation was started under no load. The axial load was slowly ramped up to the maximum load of 13 340 N, and finally the radial load was linearly ramped to a value of 133 400 N. Each slave bearing set, consisting of two tapered roller bearings, has a slight slip fit to the respective housing. The axial load is applied to the outer ring of the left slave bearing via the use of an annulus. The axial load gets transferred to the test shaft; the test shaft, in turn, transfers the load to the test bearing inner ring. The force gets reacted out to the right slave bearing housing through another annulus.

Figure 3 shows the physical test setup.

FIG. 2—*Solid model of test rig.*

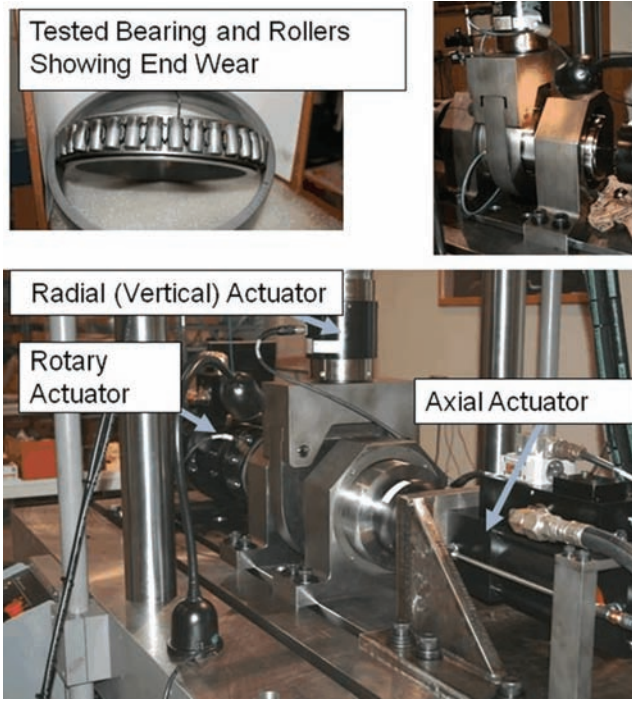


FIG. 3—Physical test setup.

Roller Profiles Considered

Single and Double Radius Profile Definition

The basic single and double radius profile used for the barrel roller is shown in Fig. 4. For the case of a single radius, $R_1 = R_2$. The transverse radius R_1 extends the length of L_1 , and the second radius R_2 extends beyond L_1 on each side to the intersection of the corner radius R_3 .

Piecewise Logarithmic Profile Development

Lundberg [3] defined a symmetric logarithmic roller profile as in Eq 1.

$$y(x) = \frac{2Q(1 - \nu^2)}{\pi E L_{\text{eff}}} \ln \left[\frac{1}{1 - \left(\frac{2x}{L_{\text{eff}}} \right)^2} \right] \quad (1)$$

This profile yields a nearly uniform contact stress distribution along the length of a roller when the roller contacts a transversely flat raceway with no misalignment. When there is misalignment, edge stresses can still occur. If the

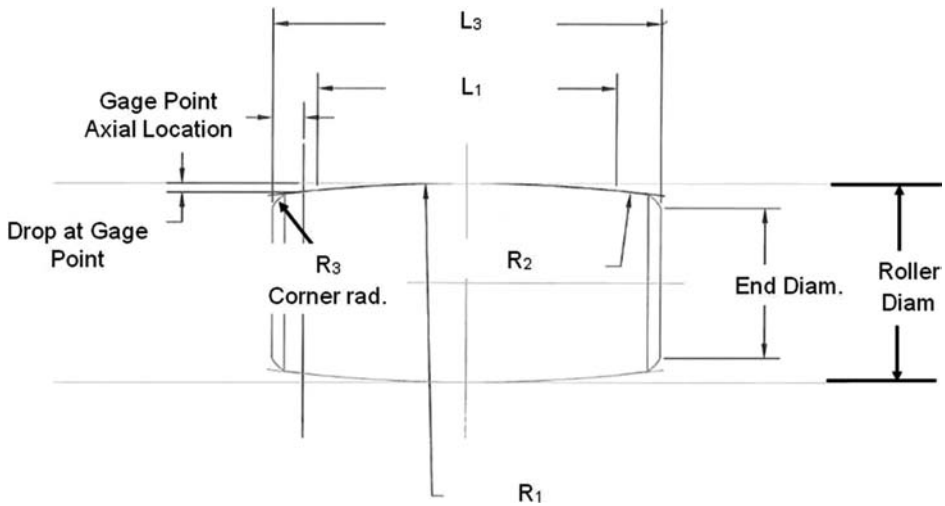


FIG. 4—Basic single and double radius roller profile.

profile is chosen such that edge stresses are avoided upon misalignment, there remains the problem that the stress will not be uniform at any location on the roller.

To counteract the problem of using one profile for a roller with misalignment, a piecewise profile is developed that allows an optimized contact stress prediction for the two conditions of a symmetric roller profile or an asymmetric profile. The symmetric profile is easier to produce and measure, but it allows stress uniformity only over the half of the roller where the contact is centered. The asymmetric profile allows for a uniform contact stress distribution over essentially the entire length of the roller.

The method is applicable to any roller type, including cylindrical, barrel, hourglass, and tapered. For the case presented in this paper, optimized symmetric and asymmetric roller profiles are developed between a misaligned barrel roller and the outer raceway of the bearing described previously.

In order to discuss the development of the profile to be used under misalignment, a few terms and general assumptions are discussed. The roller drops (y) are the drops of the roller normal to the roller axis from a cylinder of diameter d . γ is the angular misalignment of the roller axis to the raceway. A positive misalignment angle moves the center of contact in the $-x$ direction. The roller profile must be developed such that when the roller is misaligned, the profile is aligned with the raceway. Therefore, an x', y' coordinate system is introduced (see Figs. 5 and 6) that is aligned with the profile being developed along the roller surface. In this coordinate system, y' is considered as the profile drop. The profile drop is in the direction normal to the profile axis x' . Once developed, the profile drop $y'(x')$ must be transformed to a roller drop $y(x)$.

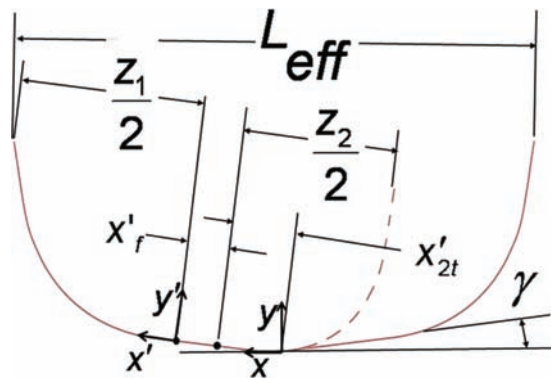


FIG. 5—Parameters for optimized symmetric profile (roller profile exaggerated).

The terms “contact patch” or “grid” are used in this paper to represent a rectangular grid that contains the roller to raceway contact area for analysis. This grid is split up into small rectangular areas of constant pressure which are utilized to calculate deformation and forces. This is demonstrated very clearly in Ref 4.

The piecewise drop profile, in both the symmetric and the asymmetric case, has a section in the center of the roller to raceway contact that has a drop of zero relative to the contacting raceway. The length of this section can be either a direct or an indirect variable controlled during the minimization of the maximum contact pressure. A modified logarithmic profile is connected to either side of this zero drop section and has the following general form:

$$y'(x') = \frac{A}{z} \ln \left[\frac{1}{1 - \left(\frac{2x'}{z} \right)^2} \right] \quad (2)$$

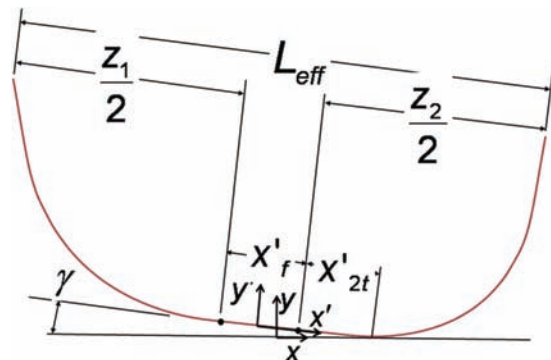


FIG. 6—Parameters for optimized asymmetric profile (profile drops exaggerated).

The variable A in Eq 2 replaces the function of roller load, Young's modulus, and Poisson's ratio in Eq 1. The variable z replaces the roller effective length.

Finally, in the derivations shown, the angular misalignment of the roller is considered small. This approximation has been examined and shown to be true for roller misalignment slopes of up to 0.010 mm/mm with roller lengths of up to 60 mm. This assumption should be reexamined for larger misalignment slopes and roller lengths.

The next sections describe the derivation of the piecewise symmetric and asymmetric profiles.

Piecewise Symmetric Profile—The origin of the x,y coordinate system is centered on the roller and roller contact patch as shown in Fig. 5. The x', y' axis origin is to the left of the section of zero profile drop. The zero profile drop section has a length x'_f . The piecewise equations for the drop profile illustrated in Fig. 5 are shown in Eqs 3–5.

$$y'(x') = \frac{A_1}{z_1} \ln \left[\frac{1}{1 - \left(\frac{2x'}{z_1} \right)^2} \right], x' \geq 0 \quad (3)$$

$$y'(x') = 0.0, -x'_f \leq x' \leq 0 \quad (4)$$

$$y'(x') = \frac{A_2}{z_2} \ln \left[\frac{1}{1 - \left(\frac{2(x' + x'_f)}{z_2} \right)^2} \right], -(x'_{2t} + x'_f) \leq x' \leq -x'_f \quad (5)$$

The second profile ends abruptly at the x,y origin, the center of the symmetric roller. The coordinates of this point in the second profile are defined in Eqs 6 and 7.

$$x'_{2t} = \frac{(L_{\text{eff}} - z_1)}{2} - x'_f \quad (6)$$

$$y'_{2t} = \frac{A_2}{z_2} \ln \left[\frac{1}{1 - \left(\frac{2x'_{2t}}{z_2} \right)^2} \right] \quad (7)$$

In order to enforce the continuity of the slope of the symmetric roller profile across the x,y origin (the center of both the roller and the contact patch in this

case), the slope of the drop profile in x', y' must be equal to γ , the roller angle of misalignment. Differentiating Eq 5 with respect to x' and setting the slope at $x' = -(x'_{2t} + x'_f)$ equal to γ yields an equation for A_2 as shown in Eq 8.

$$A_2 = \frac{z_2^3 \gamma \left[1 - \left(\frac{2x'_{2t}}{z_2} \right)^2 \right]}{8x'_{2t}} \quad (8)$$

After the elimination of A_2 as a profile variable, the symmetric profile has four variables remaining that can be used for profile optimization: A_1 , A_2 , z_1 , and x'_f . It should be noted that the equations are set up so that x'_f can be zero; x'_f cannot, however, be negative.

Finally, roller drops are defined from a perfect cylinder of diameter d . In Lundberg's work, the roller drop profile was developed relative to an axially straight raceway. For the generalized calculation of drops, the drop profile developed needs to be relative to a perfectly conforming roller and raceway. This outer raceway shape can be defined for whichever bearing type is being analyzed. For the current case of a spherical outer raceway, the initial transverse drops have the shape of a circle centered at $x = 0.0$ with radius r_o .

$$y_i(x) = r_o \left\{ 1 - \left[1 - \left(\frac{x}{r_o} \right)^2 \right]^{1/2} \right\} \quad (9)$$

With the profile drops defined, the x', y' coordinates are transformed back to the roller coordinates as follows:

$$x' = x - x'_{2t} - x'_f \quad (10)$$

$$y(x) = (y'(x') - y'_{2t}) + (x' + x'_f + x'_{2t}) \sin(\gamma) + y_i(x) \quad (11)$$

Piecewise Asymmetric Profile—The x, y coordinate system origin is centered on the roller contact patch as shown in Fig. 6. The x', y' axis origin is centered on the section of zero profile drop that, similar to the symmetric profile, has a length x'_f . In the symmetric profile, the length x'_f was considered a variable. For the case of the asymmetric drop profile, it is easier to use the variables A_1 , A_2 , z_1 , and z_2 as variables. Based on this, x'_f is written in terms of the other profile parameters in Eq 12.

$$x'_f = L_{\text{eff}} - \frac{z_1}{2} - \frac{z_2}{2} \quad (12)$$

The piecewise equations for the drop profile illustrated in Fig. 6 are shown in Eqs 13–15.

$$y'(x') = \frac{A_1}{z_1} \ln \left[\frac{1}{1 - \left(\frac{2 \left(x' + \frac{x'_f}{2} \right)}{z_1} \right)^2} \right], x' \leq -\frac{x'_f}{2} \quad (13)$$

$$y'(x') = 0, -\frac{x'_f}{2} \leq x' \leq \frac{x'_f}{2} \quad (14)$$

$$y'(x') = \frac{A_2}{z_2} \ln \left[\frac{1}{1 - \left(\frac{2 \left(x' - \frac{x'_f}{2} \right)}{z_2} \right)^2} \right], x' \geq \frac{x'_f}{2} \quad (15)$$

Similar to the symmetric profile derivation, the tangency point between the roller and raceway is defined in Eqs 16 and 17.

$$x'_{2t} = \frac{\left[-A_2 + \left(\frac{A_2^2 + z_2^4 \gamma^2}{4} \right)^{1/2} \right]}{z_2 \gamma} \quad (16)$$

$$y'_{2t} = \frac{A_2}{z_2} \ln \left[\frac{1}{1 - \left(\frac{2x'_{2t}}{z_2} \right)^2} \right] \quad (17)$$

Then the profile drops are transformed back to the roller coordinates (Eqs 18 and 19).

$$x' = x - \frac{(L_b - z_2 - x'_f)}{2} \quad (18)$$

$$y(x) = (y'(x') - y'_{2t}) + \left(x'_{2t} + \frac{x'_f}{2} - x' \right) \sin(\gamma) + y_i(x) \quad (19)$$

Analytical Method

Single Radius and Double Radii Roller Analysis

After the first bearing was tested with a single radius profile, it was apparent that a high end stress was occurring due to axial movement of the roller relative to the outer race. It was also noted that the contact location on the inner raceway flange was very evenly worn near the load zone of the bearing. Based on the initial test results, outer raceway spalling, and inner raceway flange wear, it was assumed for the purposes of analysis that the roller was essentially forced straight along its axis until contact was made with the outer raceway. This assumption is shown in Fig. 7.

Based on this assumption, a series of roller drops were determined geometrically between the roller and outer race as shown in Fig. 7(b).

The method of Hartnett was used to calculate the contact stress between the roller and the raceway using the assumption detailed above. The analysis predicted the observed end stress condition evident in the testing. Further geometric changes made to the roller, modified single radius, and double radii designs were analyzed using the same geometric assumption and numerical calculation method. Iterations were completed by manually changing the roller profile, using a spreadsheet to calculate the correct roller drops between roller and raceway, and inputting these calculated roller drops into the contact stress calculation code.

Optimization using Piecewise Logarithmic Roller Profiles

For the single transverse radius and dual transverse radii roller contact analyses, roller to raceway drops were generated outside the contact stress calculation code because of the unique characteristic of the bearing geometry coupled



FIG. 7—(a) Initial centered location with diametrical clearance. (b) Roller to raceway contact when roller is pushed along its axis until contacting the outer race.

with the applied load. The optimization code must automatically generate roller to raceway drops. The code developed is based on roller misalignment rather than an axial shift of roller to raceway locations. In order for the use of optimization to be an option, a roller misalignment between roller and raceway has to be found that yields the same stress results as the shift of the roller relative to the raceway for the same bearing geometry. Knowing the contact stress results due to the roller to raceway shift, it was a simple task to keep the roller centered on the raceway and vary the roller misalignment until the equivalent stress state was achieved.

The optimization routine developed is discussed in the Introduction. The purpose of this work was not to develop the best or most efficient solution method, but to develop a robust procedure that gives very good results. It is certain that the results derived are as close to optimal as can be practically manufactured and measured on a real-world roller. To that end, the conjugate gradient method was used in conjunction with a line-search algorithm because it is a method familiar to this author that he has used in several other optimization tasks. Each roller profile is defined based on 4 independent variables.

Symmetric profile: A_1, z_1, z_2, x'_f

Asymmetric profile: A_1, A_2, z_1, z_2

Initial guesses are made for each variable based on the loads and roller length. The initial guesses are transformed into roller drops along the roller length. A function is defined with a result that is the maximum pressure of the contact. Functional derivatives are calculated numerically by perturbing each variable slightly in the positive and negative directions one at a time. The function and its derivatives are used to generate the next iterative solution. The code continues until the maximum contact stress is minimized.

The code limits the total excursion of the variables such that at all times the profile can be derived over the entire length of the roller. For example, in the asymmetric profile, $z_1 + z_2$ must always be less than or equal to L_{eff} . It is beyond the scope of this paper to describe every possible fatal condition and how to avoid those conditions numerically. In any event, there are likely more sophisticated ways than the method described here to ensure that the roller profile is always within bounds and that the best optimization is found.

Analysis Results

Analysis results for all the roller profiles developed, obtained using the analytical methods described, are shown in this section. All contact stress profiles are for the highest loaded roller in the bearing subject to either the required radial load or a combination of radial and axial loads that have been outlined previously. The equivalent roller misalignment for the geometric shift of the roller shown in Fig. 7 was found to be 0.0038 mm/mm. This is the roller misalignment used for the symmetric and asymmetric roller optimizations.

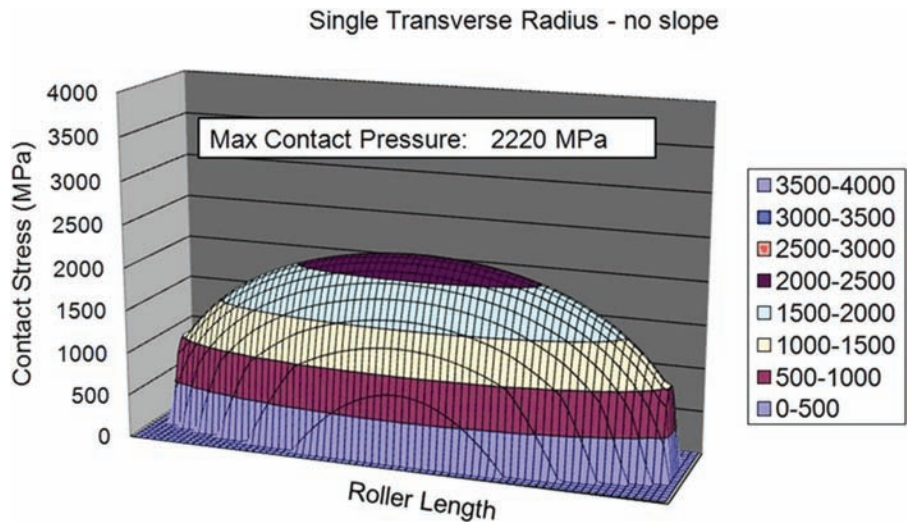


FIG. 8—Single transverse radius roller, radial load only.

Firstly, Fig. 8 shows the contact stress of the highest loaded roller of the bearing subject to the required radial load only. The stress profile shows no end stress, and the magnitude of the maximum stress is very good, considering the bearing's limited number of required oscillation cycles. Figure 9 shows the roller contact stress when the combined axial and radial loads are applied. In this case, a definite end stress is predicted that exceeds the 4000 MPa generally

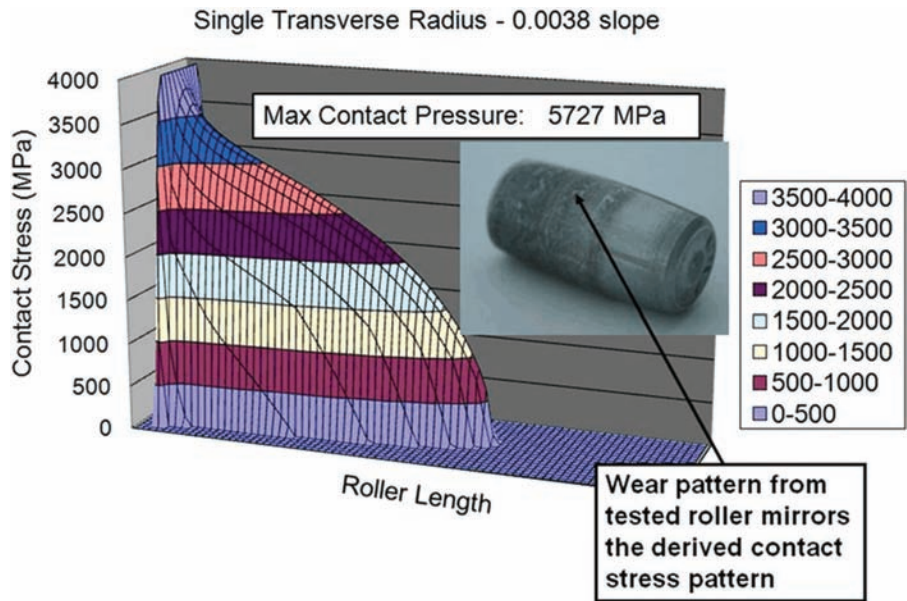


FIG. 9—Single transverse radius roller, radial and axial loads applied.

considered to cause permanent deformation or Brinelling of the bearing raceway. The picture of the tested roller in the figure demonstrates that the actual contact wear pattern matches very well with the predicted area of contact predicted by the analysis method.

Figures 10 and 11 show the dual radii profile subject to radial loading alone and to combined axial and radial loading, respectively. The contact stress for the pure radial load is slightly higher in the dual radii case than for the single radius roller due to the decreased contact area of the roller. Figure 11 highlights that the use of the dual radii profile is predicted to eliminate the end stress condition under combined axial and radial loading. With the end stress eliminated, the peak stress is below 3300 MPa, which is reasonable for a bearing with low required oscillations over its lifetime.

Figure 12 shows the contact stress of the highest loaded roller under combined axial and radial loading if the optimized symmetric profile roller is used. The maximum contact stress between roller and raceway is 2841 MPa—a 13.5% decrease from the dual radii profile roller. Under pure radial load, the roller contact stress shows a high peak value in the center of the roller, as shown in Fig. 13. This highlights the fact that a roller profile optimized for one loading condition might not be optimal for all load cases. Care must be used when trying to optimize a roller profile; if several different loads are applied to the bearing within its duty cycle, the roller profile should be optimized over all expected load conditions, not just for a single one.

Figure 14 shows the contact stress for the optimized asymmetric profile with the bearing subject to combined radial and axial loading. The maximum stress of the roller contact is 2073 MPa, which represents a 37% decrease over

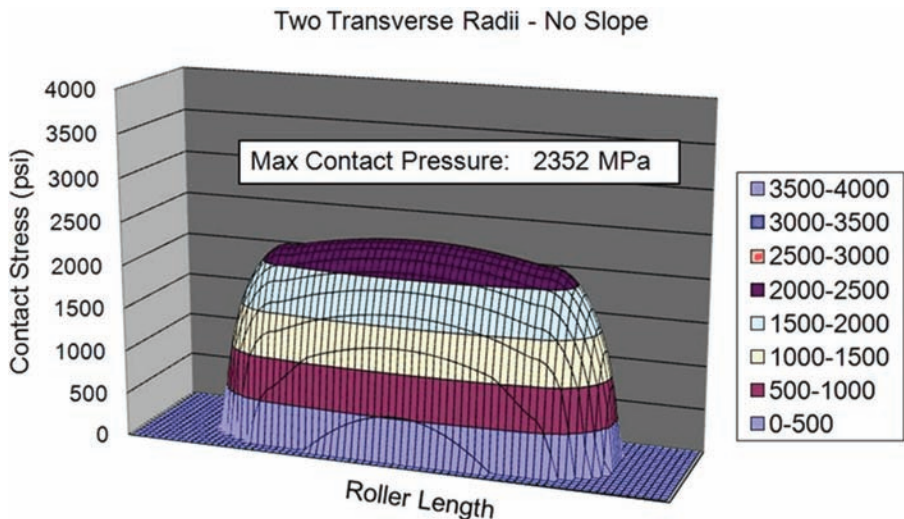


FIG. 10—Dual transverse radii roller, radial load only.

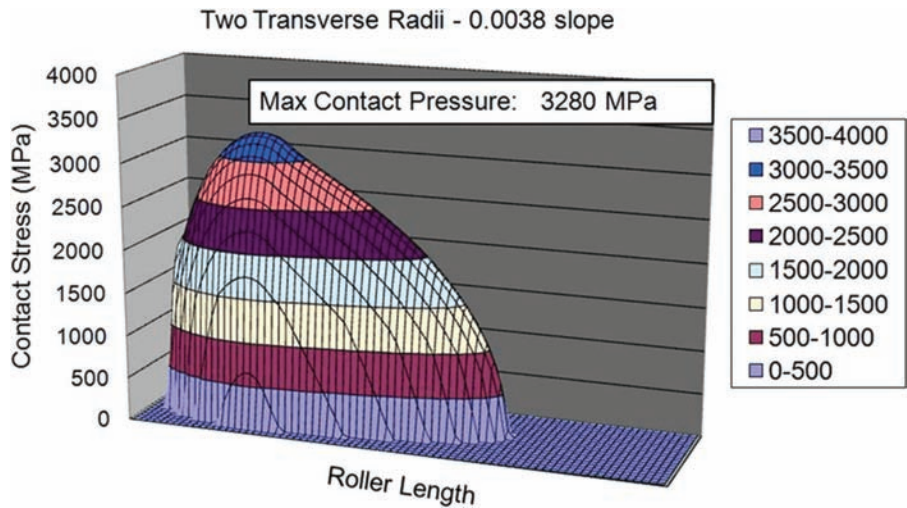


FIG. 11—Dual transverse radii roller, radial and axial loads applied.

the dual radius profile roller. The entire roller surface is efficiently utilized to transfer the load from the roller to the raceway.

Finally, Fig. 15 shows a summary of the contact gaps derived for all roller profiles. The contact gap is the distance between the roller and the raceway while the roller is at rest without applied load or misalignment. The differences between the single transverse and dual transverse radius profiles can be seen clearly. The comparison of the symmetric and asymmetric profiles is also notable. The nearly linear section of the contact gap profile can be seen to have an

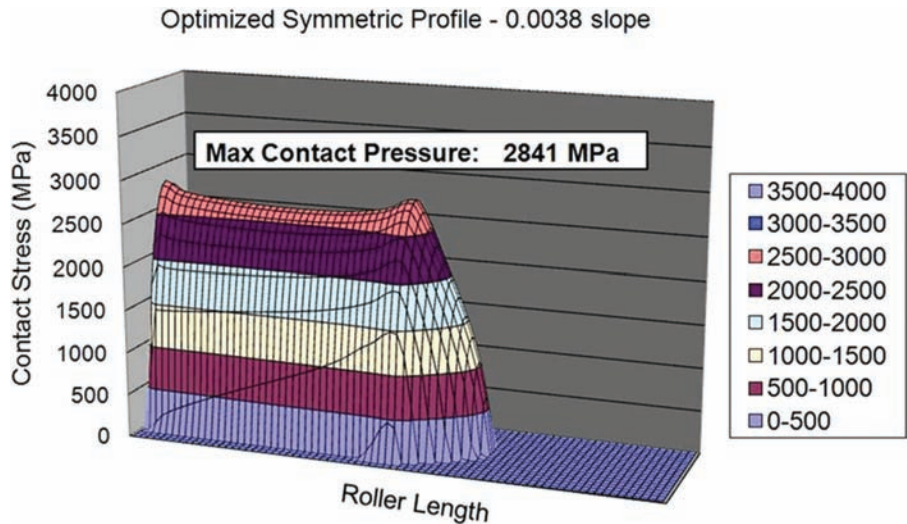


FIG. 12—Optimized symmetric profile roller, radial and axial loads applied.

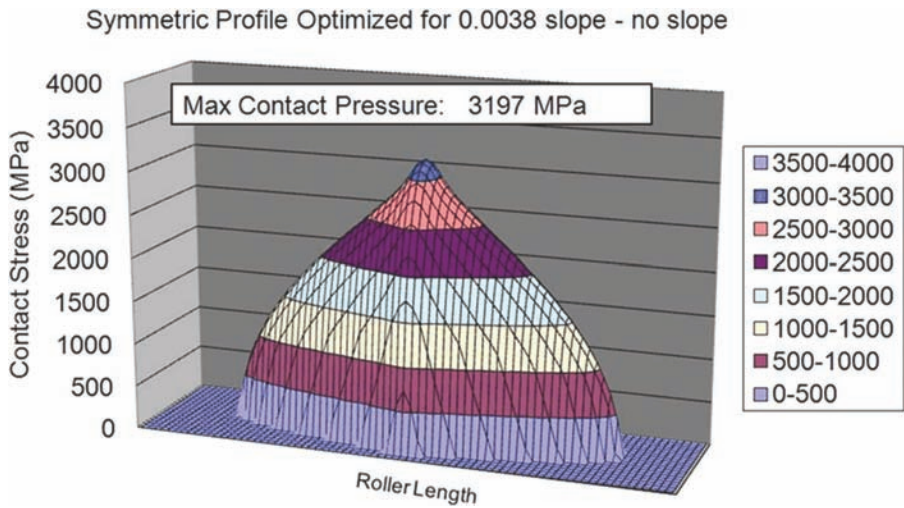


FIG. 13—Optimized symmetric profile roller, radial load only.

angle equal to the roller misalignment. The asymmetric profile shows a much longer section parallel to the symmetric profile; this is why the asymmetric profile yields lower contact stresses overall.

Test Results

The test setup used to test the bearing and roller design has been described previously. Several bearings needed to be tested to 60 000 oscillation cycles

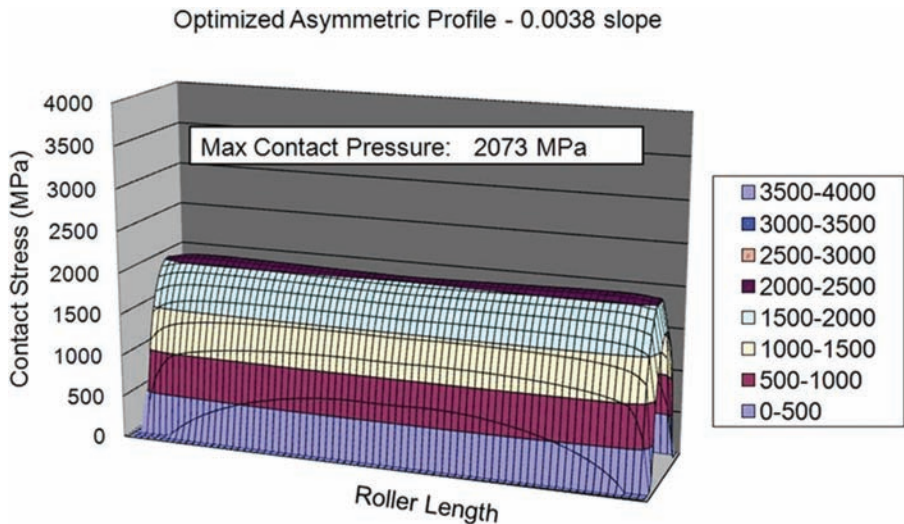


FIG. 14—Optimized asymmetric profile roller, radial and axial loads applied.

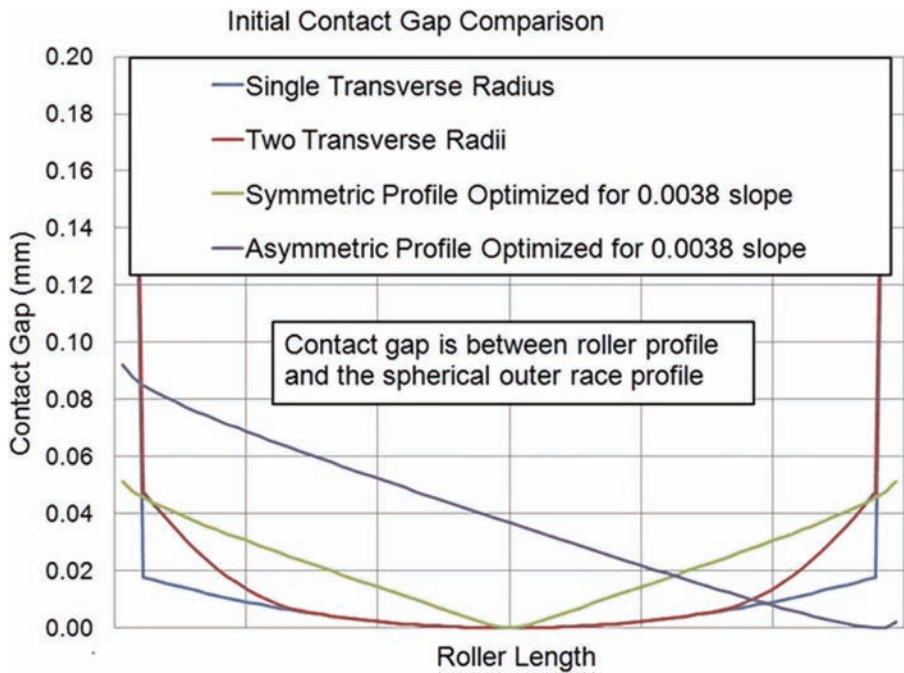


FIG. 15—Summary of resulting initial contact gaps between all rollers and the spherical outer raceway.

without evidence of spalling. The recorded output from the test rig includes the radial and axial loads, system torque, radial displacement, and oscillation angle versus time. It is convenient to plot important rig outputs over individual cycles at various times during the testing. The resulting text data files generated by the MTS control software can be read into Excel and plotted in sequences that show any significant changes occurring during the testing. Figures 16 and 17 reflect displacement and torque data recorded during the testing of a bearing made with the single radius profile rollers. The torque measured includes that for the single test bearing and the four tapered roller slave bearings; therefore, the torque is system torque rather than specifically the torque of the test bearing. However, changes in the torque level during testing are likely a result of changes to the test bearing as the slave bearings are oversized.

The test results are very instructive. Both the displacement and the torque start out very smoothly, as can be seen by the results up to 36 000 cycles. The next several cyclic plots show higher peak-to-peak noise levels in both displacement and torque signatures. Near the end of the testing, the local peak-to-peak displacements exceed 0.01 mm, which is much higher than in the early cyclic data, where the average peak-to-peak displacements are estimated as 0.002 mm.

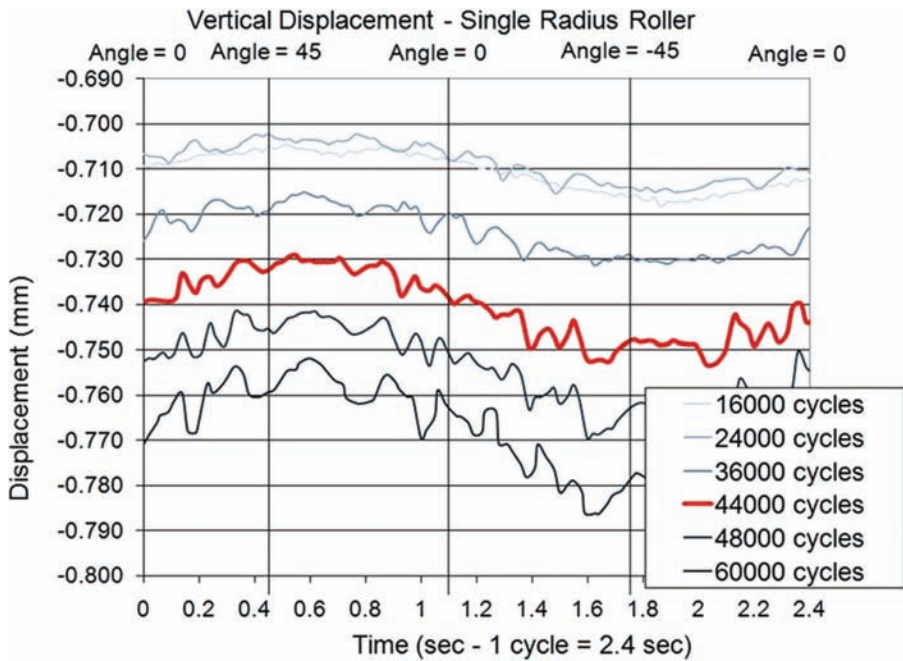


FIG. 16—Radial displacement (mm) of the bearing housing over one cycle at various cycle counts throughout the testing. Combined radial and axial load on bearing made with rollers of single transverse radius profile.

The torque data show a similar trend. At 16 000 cycles, the peak-to-peak torque (i.e., not related to the change in rotation direction) is approximately ± 10 N-m. Toward the end of testing, that peak-to-peak torque increases to ± 30 N-m.

After the completion of the first bearing test, the bearing was disassembled, with significant evidence of spalling due to roller end stress. The rollers, although not spalled, showed significant discoloration on the loaded side (refer again to Fig. 9). The second bearing test was performed with an effort to observe the progression of spalling. The test was stopped at 44 000, 52 000, and 60 000 cycles, at which points the position of the inner and outer rings on the test shaft were carefully marked and the bearing was taken out of the test housing. The inner raceway and rollers could be rotated 90° to the outer race while keeping the rollers and raceway together. The grease was cleaned out of the bearing, and pictures were taken of the outer ring load zone. Once the pictures were taken, the bearing was carefully put back together and mounted on the test shaft and housing, making sure the inner ring and outer ring load zones were in the same relative orientation. The bearing was then re-greased, and the testing continued. Using this method, the progression of spalling of the outer

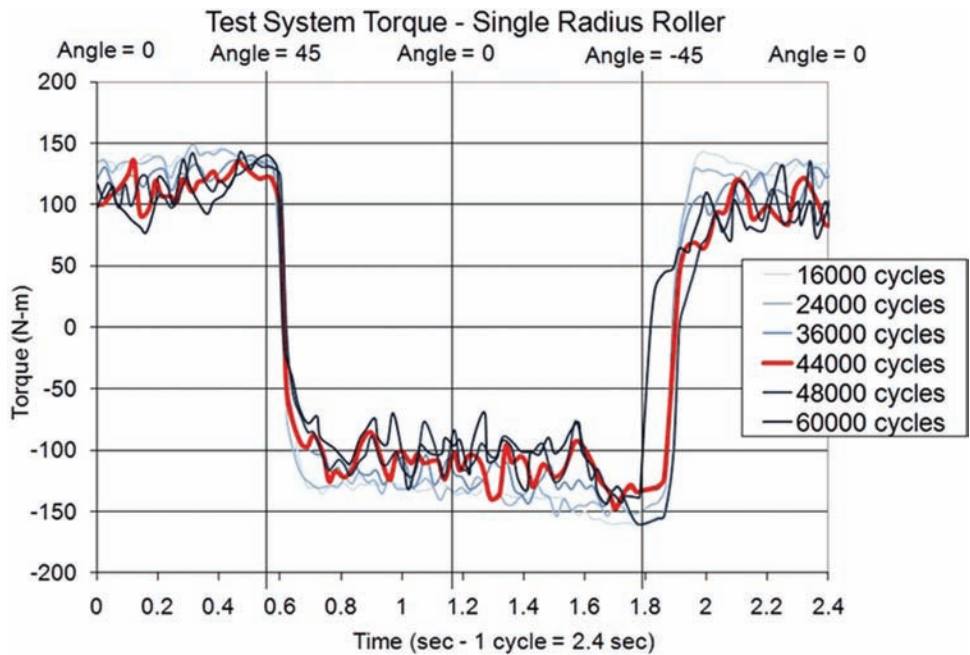


FIG. 17—System torque (N-m) over one cycle at various cycle counts throughout the testing. Combined radial and axial load on bearing made with rollers of single transverse radius profile.

raceway could be clearly seen. Figure 18 illustrates the early onset of spalling during the test.

The analysis work presented in the preceding section was completed in order to determine the reason for the early failure of the bearing. The analysis showed the roller end-stress condition, and this closely correlated with the test results. The dual radius profile was chosen for its simplicity in terms of manufacturing and measuring. As the detailed load history of the bearing was not known to us, it was felt that good contact stress under purely radial loads was important.

Tests were conducted on several bearings containing rollers with the dual radii profile, and all bearings passed the test with no evidence of spalling. Figures 19 and 20 show the bearing displacement and system torque during testing. It can be seen that neither of the sets of data shows significant changes in the peak-to-peak character of the output over the length of the test.

Conclusions

This paper has shown how testing and analytical capabilities go hand-in-hand. By testing bearings and observing the resulting failures, one can derive

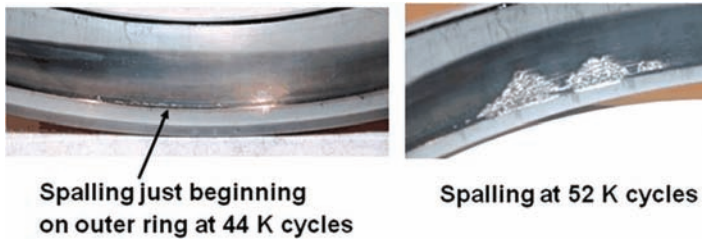


FIG. 18—*Spalling fatigue due to roller end stress condition at 44 000 and 52 000 cycles. Combined radial and axial loads on bearing made with rollers of single transverse radius profile.*

appropriate analytical assumptions that lead to accurate contact stress analysis. A bearing using rollers with a single transverse radius failed a test when combined radial and axial loading was applied. Based on the test data, a set of assumptions were developed that could be used to properly analyze the bearing's roller complement. The analysis procedure predicted failure for the original roller and was then used to design a roller profile that was predicted to pass. The bearing made with the analytically successful roller design successfully passed the testing.

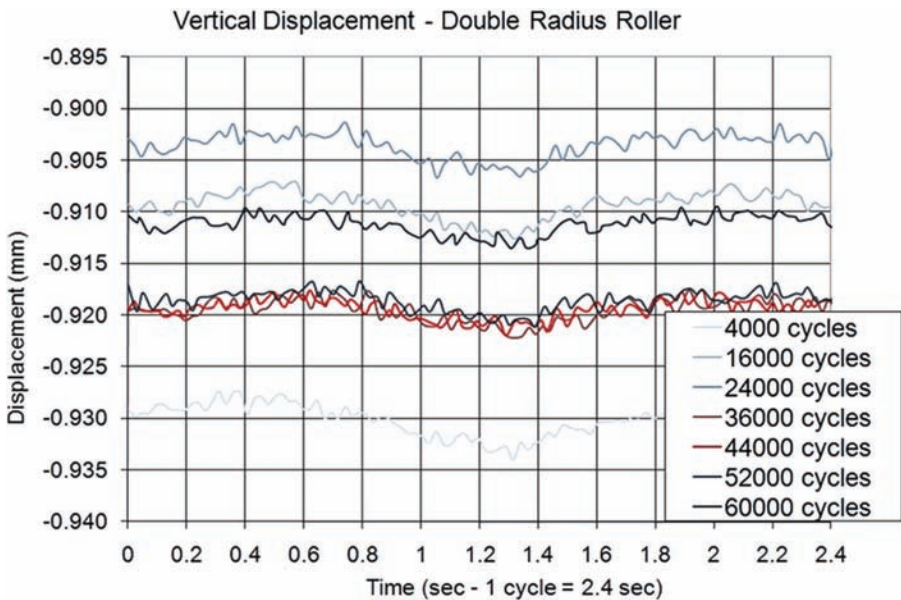


FIG. 19—*Radial displacement (mm) of the bearing housing over one cycle at various cycle counts throughout the testing. Combined radial and axial load on bearing made with rollers using the dual transverse radii profile.*

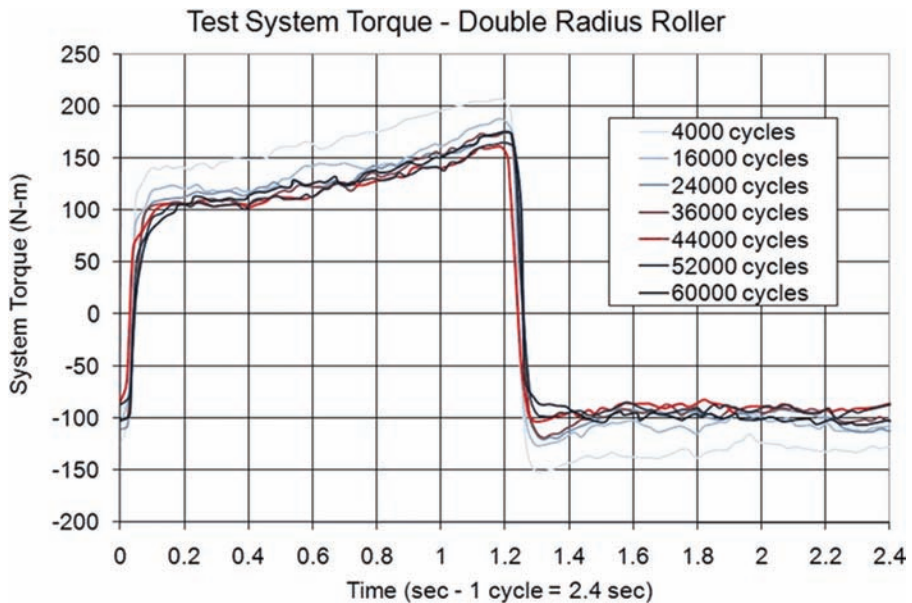


FIG. 20—System torque (N-m) over one cycle at various cycle counts throughout the testing. Combined radial and axial load on bearing made with rollers using the dual transverse radii profile.

Beyond the development of a roller profile that would work in the given application, optimized rollers were developed using piecewise logarithmic profiles. Symmetric and asymmetric profiles were developed using the methods defined in this paper and resulted in 13.5% and 37% reductions in contact stress, respectively, over the dual radii roller profile for the given combined radial and axial loads applied. As shown, when a purely radial load is applied to the symmetric profile optimized for both radial and axial loading, the contact stress in the center of the roller is unacceptably high. It becomes the analyst's job to optimize a roller profile over the full spectrum of expected loads for the bearing. This is a challenge that is not addressed within this paper.

Finally, it is mentioned that the optimizations reported in this paper relate to the interaction of the roller with the outer ring. Once the roller profile is developed relative to the outer ring, a new inner ring raceway profile must be developed. The same techniques developed in this paper can address the profile of the other raceway. This is left for a later study.

Acknowledgments

The writers would like to thank RBC Bearings for permission to publish this paper. The writers would like to thank Mr. K. Smith and Mr. G. Hamel for proofreading the manuscript and making helpful suggestions about the content,

and Mr. M. Wegrzyniak for his hard work during the testing phase of the project.

References

- [1] Hertz, H., "On the Contact of Elastic Bodies," *Miscellaneous Papers*, Macmillan, London, 1896.
- [2] Hertz, H., "On the Contact of Rigid Elastic Solids and on Hardness," *Miscellaneous Papers*, Macmillan, London, 1896.
- [3] Lundberg, G., "Elastic Contact between Two Semi-Infinite Bodies," *Forsch. Geb. Ingenieurwes.*, Vol. 5, 1939, pp. 201–211.
- [4] Hartnett, M. J., "The Analysis of Contact Stresses in Rolling Element Bearings," *ASME J. Lubr. Technol.*, Vol. 101, 1979, pp. 101–105.
- [5] Hartnett, M. J., "A General Numerical Solution for Elastic Body Contact Problems," *Presented at the Winter Annual Meeting of the American Society of Mechanical Engineers*, Chicago, IL, Nov. 16–21, 1980, ASME, NY.
- [6] Ghali, A., and Neville, A. M., *Structural Analysis: A Unified and Matrix Approach*, Intext Educational Publishers, Scranton, PA, 1972.
- [7] Sokolnikoff, I. S., *Mathematical Theory of Elasticity*, Krieger, Malabar, FL, 1983.
- [8] Carnahan, B., Luther, H. A., and Wilkes, J. O., *Applied Numerical Methods*, Wiley & Sons, New York, 1969.
- [9] Harris, T. A., and Kotzalas, M. N., *Rolling Bearing Analysis, Advanced Concepts of Bearing Technology*, 5th ed., CRC Taylor & Francis, New York, 2007.
- [10] ISO 281, 2007, *Rolling Bearings—Dynamic Load Ratings and Rating Life*, 2nd ed., International Organization for Standardization, Geneva, Switzerland.
- [11] ISO/TS 16281, 2008, "Rolling Bearings—Methods for Calculating the Modified Reference Rating Life for Universally Loaded Bearings," International Organization for Standardization, Geneva, Switzerland.
- [12] Poplawski, J. V., "COBRA AHS: Advanced High-Speed Computer Optimized Bearing Analysis for High Speed Multi-Bearing/Shaft Systems with ElastoHydrodynamic Lubrication," Poplawski & Associates, Bethlehem, PA, 2011.
- [13] Poplawski, J. V., Saretsky, E. V., and Peters, S. M., "Effect of Roller Profile on Cylindrical Roller Bearing Life Prediction," *NASA TM-2000-210368*, National Aeronautics and Space Administration, Hanover, MD, 2000.
- [14] Losche, T., "Capacity Increase Due to Modified Line Contact in Cylindrical and Tapered Roller Bearings," *FAG 1987-1*, Schaeffler Technologies FAG, Schweinfurt, Germany, 1987.

- [15] Fujiwara, H., Kawase, T., Kobayashi, T., and Yamauchi, K., “Optimized Logarithmic Roller Crowning Design of Cylindrical Roller Bearings and Its Experimental Demonstration,” *Proceedings of the ASME/STLE 2009 International Joint Tribology Conference IJTC2009*, Memphis, TN, Oct. 19–21, 2009, Memphis, TN, ASME, NY.
- [16] Fujiwara, H., and Kawase, T., “Logarithmic Profiles of Rollers in Roller Bearings and Optimization of the Profiles,” *Trans. Jpn. Soc. Mech. Eng., Ser. C*, Vol. 72, 2006, pp. 3022–3029.
- [17] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., *Numerical Recipes in Fortran 77, The Art of Scientific Computing*, 2nd ed., Cambridge University Press, New York, 1992.

Alan Leveille,¹ Peter Frantz,² and Garry Rosene²

A Model to Estimate Separator Forces during Ball Speed Variations

REFERENCE: Leveille, Alan, Frantz, Peter, and Rosene, Garry, "A Model to Estimate Separator Forces during Ball Speed Variations," *Rolling Element Bearings* on April 13–15, 2011 in Anaheim, CA; STP 1542, Yoshimi R. Takeuchi and William F. Mandler, Editors, pp. 71–91, doi:10.1520/STP104209, ASTM International, West Conshohocken, PA 2012.

ABSTRACT: When the combined loads on ball bearings include forces in the radial or moment directions, the balls will not orbit the bearing at one common speed. This ball speed variation (BSV) might result in additional bearing friction forces if the variation in speed is large enough to allow the balls to spread out from their normal spacing by an amount that exceeds the cage pocket clearances. In this paper we present a model used to estimate the cage forces and friction torques caused by BSV and compare the model predictions with measured friction torque test data. The model first analyzes the forces on a single ball by determining the distance over which a ball must slip during a single orbit of the bearing center. The ball to race force is determined by equating the energy lost during this slip with the energy input through traction at the inner race interface. The component of the ball to cage force normal to the ball–pocket interface is then determined by balancing this force with the ball to race force. The cage to land force is similarly determined by balancing the collective forces at all the ball to race and ball to cage interfaces. Finally, the BSV drag torque is written as the sum of contributions from these three drag sources. In order to validate the model, test data were obtained using single 204 size bearings operating under applied thrust and radial loading. The bearing drag torque was found to depend on the degree of misalignment and cage pocket clearance, as predicted by the model.

KEYWORDS: ball bearing, ball speed variation, bearing misalignment, tribology

Manuscript received July 15, 2011; accepted for publication January 17, 2012; published online April 2012.

¹Independent Bearing Consultant, Las Vegas, NV.

²The Aerospace Corporation, El Segundo, CA.

Copyright © 2012 by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

Introduction

Ball speed variation (BSV) occurs when individual balls in a ball bearing are driven to orbit the bearing center at different angular velocities. BSV occurs to some degree in nearly every operating bearing, but it is not always large enough to significantly affect the bearing's function. When the variation in speed does become large, the faster balls and slower balls might influence each other by interacting through the separator [1]. This interaction produces additional forces that contribute to excess bearing drag. As the BSV increases, the drag due to ball skidding and cage binding can become considerable. Furthermore, component wear rates might also increase dramatically and cause rapid failure at the ball–race and ball–separator contacts.

This situation can be caused by a number of factors, including asymmetric loading, raceway distortions, ball size variations, etc. The effects can be amplified by many other factors, such as small pocket clearance, poor lubrication, tight raceway conformity, etc. In spite of these causes for speed variations, the angular velocity of any given ball, when averaged over many revolutions, must equal that of every other ball. If it didn't, then the balls would eventually pass each other, which is not possible. Ultimately, the angular velocity differences are eliminated by interaction with the separator (the retainer, or cage), which forces balls to speed up or slow down by skidding at the ball to race interfaces. This interaction with the cage introduces several modes of energy loss due to loaded sliding at the following interfaces: ball to inner race, ball to outer race, ball to separator, and separator to control land. The sum of these energy loss mechanisms can add a significant amount to the drag torque of an operating ball bearing and might result in bearing failure.

In this paper, we present a model for estimating the magnitude of forces on the separator and the amount of drag torque caused by BSV. The derivation of this model is divided into three sections. In the first section, we study the motion of a single ball with an angular velocity that differs from that of the separator in order to determine the amount of skidding of a given ball through one orbit about the bearing center. The second section shows how the normal force at a single ball to separator interface can be determined by equating it with the traction force driving the ball skid in the first section. The third section demonstrates how the forces on the separator from each contributing ball are balanced by the separator to land forces. This provides a means for calculating the separator to land force. Once all the forces are determined, they are used to estimate bearing drag torques by equating the energy input from traction at the inner ring with the energy lost via dissipation from friction at the various sliding interfaces.

This BSV model requires, as inputs, detailed knowledge of the bearing geometry, materials properties, and mounting configuration. Some of these values are used directly in the model, and others must be used in conjunction with a computational program that provides the loads and deflections of a bearing in

its mounted and operational condition [2,3]. These loads and deflections are then used as inputs for this model. Table 1 shows the notation used for the inputs necessary for this model. Many other factors that are required as inputs for a general bearing model are not shown here because they are not explicitly used in our BSV model.

Ball Skidding Distance

In this section, we determine the total range of motion of a ball with respect to the separator Δ . If the ball range exceeds the ball-pocket clearance C , then the ball will interact with the cage pocket wall. The difference between the range of motion of the ball with respect to the separator and the pocket clearance must be the total sliding distance (assuming rigid cage behavior), which will occur at the ball interface with the inner and outer races. The energy that is dissipated due to sliding friction at these interfaces must be supplied to the bearing and will be a primary component of the ball speed variation drag torque.

The angular velocity of the ball center is

$$\frac{\partial \phi}{\partial t} = \omega_B \quad (1)$$

The generalized equilibrium diagram for a ball operating without consideration of the pocket forces is given in Fig. 1. It is known from the basic mechanics of angular contact bearings [4] that the orbital angular velocity of a ball at low speed ($\beta_i = \beta_o = \beta$) is given by

$$\omega_B = \frac{\omega_i}{2}(1 - \gamma \cos \beta) \quad (2)$$

TABLE 1—Notation used in the presented model.

Bearing geometry and material properties used directly in this model	
E	Pitch diameter
d	Ball diameter
μ, μ_s, μ_l	Friction coefficients: ball/race, ball/separator, separator/land
C	Separator ball/pocket clearance
Bearing properties that must be determined with a bearing computational tool	
P, P_i, P_o	Ball loads: low speed, high-speed inner race, high-speed outer race
β, β_i, β_o	Operating contact angles: low speed, high-speed inner race, high-speed outer race
Definitions of other factors used in this derivation	
ω_B, ω_s	Orbital angular velocity: ball, separator
ω_b, ω_i	Rotational angular velocity: ball, inner ring
ϕ, ϕ_l	Angular position: ball, separator/land contact

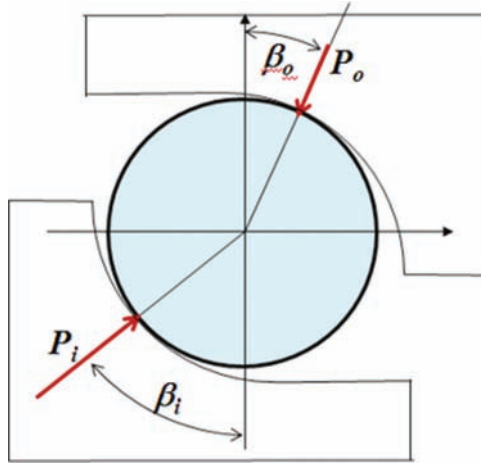


FIG. 1—Schematic cross section of a ball between an inner race (bottom) and outer race (top) showing the high speed ball load vectors and contact angles.

The orbital angular velocity at high speed with inner race ball control is given by

$$\omega_B = \frac{\omega_i}{(1 + \cos \delta)} [\cos \delta - \gamma \cos \beta_o] \quad (3)$$

The orbital angular velocity at high speed with outer race ball control is given by

$$\omega_B = \frac{\omega_i}{(1 + \cos \delta)} [1 - \gamma \cos \beta_i] \quad (4)$$

where:

$$\gamma = d/E, \text{ and}$$

$$\delta = \beta_i - \beta_o$$

Assuming that the ball center remains at radius $E/2$, the differential circumferential distance travelled by the ball center over time dt is given by

$$dC_B = \frac{E}{2} \omega_B dt \quad (5)$$

A similar formulation provides the circumferential distance travelled by the separator (dC_s) at an angular velocity ω_s . During the time interval dt , a given ball will travel a differential distance that is different than that of the separator.

$$d\Delta = \frac{E}{2} (\omega_B - \omega_s) dt \quad (6)$$

By substituting Eq 1 into Eq 6, we obtain

$$d\Delta = \frac{E}{2} \left[1 - \frac{\omega_s}{\omega_B} \right] d\varphi \quad (7)$$

The value of $[1 - \omega_s/\omega_B]$ is found from Eqs 2, 3, and 4.

Case 1: Slow Speed

$$\left[1 - \frac{\omega_s}{\omega_B} \right] = 1 - \left(\frac{1 - \gamma \cos \beta_s}{1 - \gamma \cos \beta} \right) \quad (8)$$

Case 2: High Speed and Inner Race Control

$$\left[1 - \frac{\omega_s}{\omega_B} \right] = 1 - \frac{1 + \cos \delta}{1 + \cos \delta_s} \left[\frac{\cos \delta_s - \gamma \cos \beta_{os}}{\cos \delta - \gamma \cos \beta_o} \right] \quad (9)$$

Case 3: High Speed and Outer Race Control

$$\left[1 - \frac{\omega_s}{\omega_B} \right] = 1 - \frac{1 + \cos \delta}{1 + \cos \delta_s} \left[\frac{1 - \gamma \cos \beta_{is}}{1 - \gamma \cos \beta_i} \right] \quad (10)$$

The terms with the subscript s refer to the contact angle of the ball that determines the separator speed. For simplicity, in the remainder of this presentation we assume that the separator speed is equal to that of the heaviest loaded ball. This ball is referred to as the “controlling ball,” because it controls the angular velocity of the separator. This assumption is physically unrealistic; it is more likely that the separator speed is closer to the load-weighted average speed of all the balls. However, the assumption of one controlling ball is necessary in order to allow a tractable mathematical analysis. Despite its being unrealistic, we suggest that it still provides a good estimate for the magnitude of the forces. A reduction in the magnitude of forces at a slow ball due to an error in the cage speed would be balanced by an increase in the magnitude of forces at a fast ball. Ultimately, the fidelity will be tested in comparison with experimental data that are described below.

With a positive $d\Delta$ ($\omega_B > \omega_s$), the ball is trying to move faster than the separator. The opposite is true for ($\omega_B < \omega_s$).

Figure 2 shows an example of how Eq 7 might look over one complete revolution of the separator. This is a case such as a radially loaded bearing, in

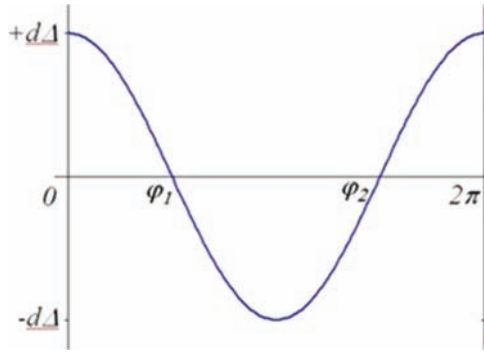


FIG. 2—Example of ball excursions over an element of time dt through a full orbit of a ball.

which a given ball speed would have a sinusoidal profile and the separator speed would remain constant.

In Fig. 2,

$$\int_{\varphi=0}^{\varphi_1} d\Delta + \int_{\varphi_2}^{2\pi} d\Delta = \Delta_1 = \text{positive distance}$$

$$\int_{\varphi_1}^{\varphi_2} d\Delta = \Delta_2 = \text{negative distance}$$

For an ideal case, if all the balls were located in the ball pockets with a clearance of greater than Δ_1 between them and the forward edge of the pocket and a clearance of greater than Δ_2 between them and the trailing edge of the pocket, then no ball to separator interaction would occur. Thus, $\Delta_1 + |\Delta_2| = \Delta = \text{the minimum desired pocket clearance}$. Therefore, the total displacement of the ball with respect to the separator is the integral of the absolute value of $d\Delta$,

$$\Delta = \int_{\varphi=0}^{2\pi} |d\Delta| \quad (11)$$

Now let

$$\xi(\varphi) = \left| \left[1 - \frac{\omega_s}{\omega_B} \right] \right| \quad (12)$$

Then Eqs 11 and 12 become

$$\Delta = \frac{E}{2} \int_{\varphi=0}^{2\pi} \xi(\varphi) d\varphi \quad (13)$$

Equation 13 may be numerically integrated between any two limits φ_1 and $\varphi_1 + 2\Delta\varphi$ according to Simpson's Rule,

$$\Delta = \frac{E}{2} \int_{\varphi_1}^{\varphi_1 + 2\Delta\varphi} \xi d\varphi = \frac{E}{2} \left\{ \frac{\Delta\varphi}{3} [\xi(\varphi_1) + 4\xi(\varphi_1 + \Delta\varphi) + \xi(\varphi_1 + 2\Delta\varphi)] \right\} \quad (14)$$

For our case, $\Delta\varphi = 2\pi/2n$ so that $\varphi_1 + 2\Delta\varphi$ will take the analysis from a ball located at φ_1 to the next ball located at $\varphi_1 + 2\pi/n$.

The physical basis of the model is illustrated by the specific example of a bearing under a combined thrust and radial load, as in Fig. 3. The ball at $\varphi = 0$ is the heaviest loaded (the controlling ball) and has minimum ω_B . The lightest loaded ball is at $\varphi = \pi$ and has maximum ω_B . Starting at $\varphi = 0$, a ball is not in contact with the leading edge of the retainer until either $\int_{\varphi=0}^{\varphi=\varphi_x} d\Delta = C$ or $\varphi_x = \pi$, whichever occurs first.

Analysis of Forces due to Ball Skidding

The magnitude of ball pocket forces generated is estimated by an energy balance method in which the energy lost at the interfaces between the ball and the inner and outer races is equal to the energy input into the bearing by a traction force at the inner ring that is necessary in order to overcome the drag force due to skidding. When the ball and separator are in contact, the ball will slip an amount $d\Delta$ in a time increment dt . This slippage will cause a traction-induced energy loss of $d\Delta \times \mu P$ at each sliding ball–race interface, which in turn shows

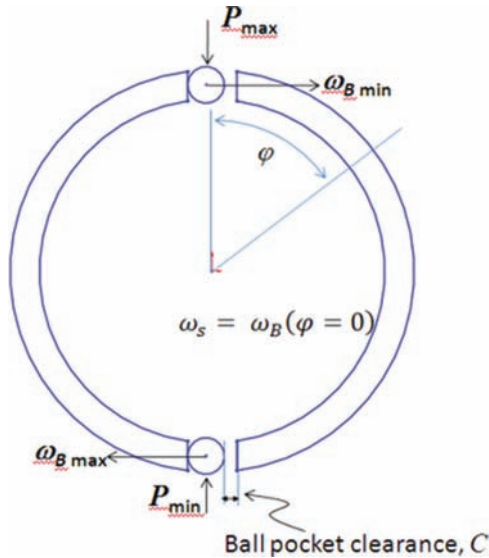


FIG. 3—Example of a possible ball and separator arrangement with a slow ball at $\phi = 0$ and a fast ball at $\phi = \pi$.

up as added bearing torque. Additional energy is lost by the ball pocket force against the rotating ball and by the net retainer force sliding against the control land. Torque drag due to these energy loss mechanisms is treated in a later section.

For the condition in which a ball is in contact with the ball pocket leading or trailing edge (other than the control ball), $d\Delta$ is given by Eq 6. The meaning of a positive $d\Delta$ is that the ball is trying to move faster than the separator and will push the separator once $\Delta > C$. For negative $d\Delta$, the separator is pushing the ball. In either case, we stipulate that $d\Delta$ represents the distance that the ball must slip with respect to the inner and outer races after the contacting conditions have been met.

The incremental energy loss due to this sliding action is

$$de = \mu P_i d\Delta + \mu P_o d\Delta \quad (15)$$

For slow speeds, where $P_i = P_o = P$,

$$de = 2\mu P d\Delta \quad (16)$$

Equations 15 and 16 assume rigid behavior between the ball and the separator pocket. The real amount of slippage will be less due to the deflection that occurs in the separator. For the sake of simplicity, we continue to assume rigid behavior as an approximation.

Energy input into the bearing in order to overcome the energy loss of Eqs 15 and 16 must come from a traction force F_i located at the ball contact with the inner ring. That is,

$$de = F_i \frac{1}{2} (E - d \cos \beta_i) \omega_i dt = \mu d\Delta (P_i + P_o) \quad (17)$$

Substituting Eq 6 for $d\Delta$ in Eq 17 gives

$$F_i \frac{1}{2} (E - d \cos \beta_i) \omega_i dt = \mu (P_i + P_o) \frac{E}{2} (\omega_B - \omega_s) dt \quad (18)$$

In solving for the traction force, we find that

$$F_i = \frac{\mu (P_i + P_o)}{(1 - \gamma \cos \beta_i)} \left(\frac{\omega_B}{\omega_i} - \frac{\omega_s}{\omega_i} \right) \quad (19)$$

At low speeds, the traction force at the outer ring F_o is equal to that of the inner ring (Eq 19). However, in order to determine this value at high speed, we must also consider the separator forces acting on the ball, as shown in the next section.

Analysis of Forces due to Sliding at the Ball–Pocket Interface

Consider the force diagram in Fig. 4, showing a ball at $\phi = 0$. At equilibrium, the magnitude of the normal force at the ball–pocket interface must be balanced by the traction forces due to sliding at the inner and outer races. In other words, the sum of the forces in the x -direction and the sum of moments about the inner race contact (point A in Fig. 4) must both equal zero. Thus,

$$\sum F_x = F_o + F_i - N = 0 \quad (20)$$

$$\sum M_{zz} = F_o \frac{d}{2} [\cos \beta_o + \cos \beta_i] - N \frac{d}{2} \cos \beta_i = 0 \quad (21)$$

Solving for F_o in Eq 21, we get

$$F_o = \frac{N \cos \beta_i}{\cos \beta_o + \cos \beta_i} \quad (22)$$

With the onset of ball skidding, the ball also slides at the ball–pocket contact, producing a traction force $\mu_s N$. We neglect the influence of $\mu_s N$ on the equilibrium forces P_i, P_o based on the assumption that $\mu_s N \ll P_i$ and P_o , but the energy loss due to this additional traction force is accounted for in our total energy and torque generation equations. This is accomplished by our assuming an additional element of traction force at the inner ring due to the ball to separator interaction F_s ; we match energies via

$$F_s \frac{1}{2} (E - d \cos \beta_i) \omega_i dt = \mu_s N \frac{d}{2} \omega_b dt \quad (23)$$

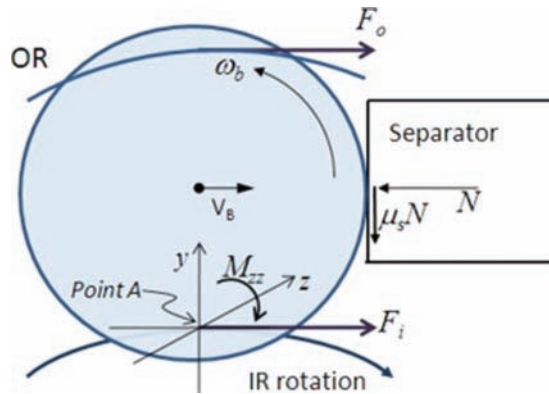


FIG. 4—Schematic of ball, races, and separator contacts showing the traction force vectors. IR, inner ring; OR, outer ring.

$$F_s = \frac{\gamma \mu_s N}{(1 - \gamma \cos \beta_i)} \left(\frac{\omega_b}{\omega_i} \right) \quad (24)$$

Now we solve for all of the forces acting at the interfaces between the ball and races and the ball and separator pocket.

Equation 19 gives us

$$F_i = \frac{\mu(P_{i+o})}{(1 - \gamma \cos \beta_i)} \left(\frac{\omega_B}{\omega_i} - \frac{\omega_s}{\omega_i} \right) \quad (25)$$

Combining Eq 22 with Eq 20 gives

$$N = F_i \left[\frac{\cos \beta_o + \cos \beta_i}{\cos \beta_o} \right] \quad (26)$$

By substituting Eq 26 into Eq 22, we get

$$F_o = F_i \frac{\cos \beta_i}{\cos \beta_o} \quad (27)$$

Finally, we substitute Eq 26 into Eq 24 to get

$$F_s = \left(\frac{\gamma \mu_s}{1 - \gamma \cos \beta_i} \right) \left(\frac{\omega_b}{\omega_i} \right) F_i \left[\frac{\cos \beta_o + \cos \beta_i}{\cos \beta_i} \right] \quad (28)$$

For any given ball that has met the criteria established in the preceding section, Eqs 25, 27, and 28 show the amount of force required as an additional input into the bearing in order to overcome the opposing drag forces at the inner ring, outer ring, and ball–pocket interface, respectively.

Analysis of Forces due to the Controlling Ball and Sliding at the Separator–Land Interface

We note that at equilibrium, the net force on the retainer due to the sum of all ball pocket forces must be balanced by a net reaction force between the separator and the control land. That force is shown generally in Fig. 5(a) for outer land control and in Fig. 5(b) for inner land control. The final task is to sum the contribution of the separator forces from each ball and balance these with the force at the interface between the separator and the control land surface. This control land can occur at either the inner or outer ring, depending on the bearing design, as shown in Fig. 5. Figure 6 shows the force diagram for this situation. Here, the ball pocket forces discussed above are balanced by the reaction force on the separator by the controlling ball S and the control land reaction

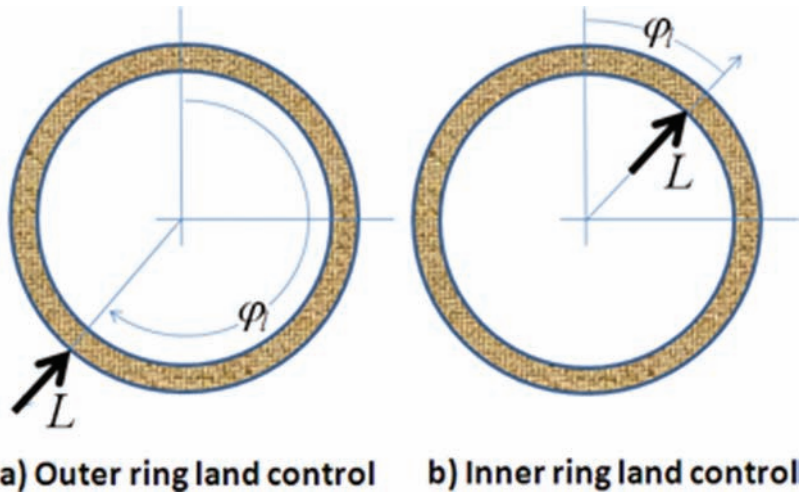


FIG. 5—Two possible orientations of the land to separator force vector for an outer ring piloted separator (a) and an inner ring piloted separator (b).

force L , as illustrated for the case of outer ring control. The coefficient of sliding friction between the separator and control land is represented by μ_l .

In order to balance these forces, we first resolve the forces on each ball into their components in the x - and y -directions and the moment about the bearing center as shown in Fig. 7. Referring to Fig. 7, the summation of forces and moments is given by

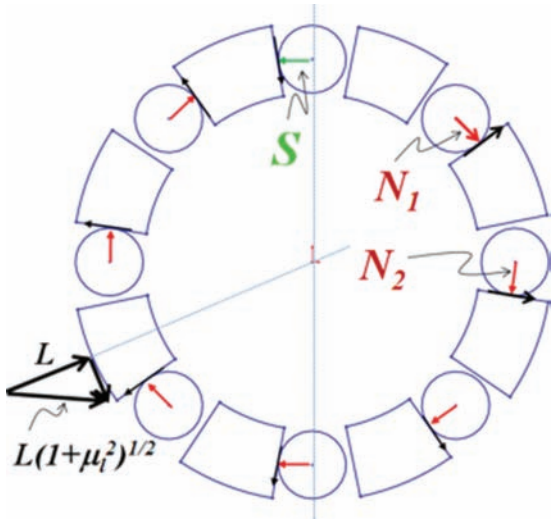


FIG. 6—This figure shows an example of a possible arrangement of normal load and traction force vectors.

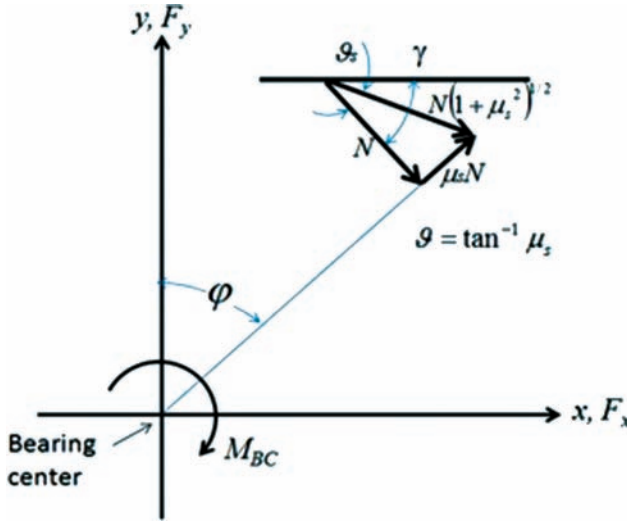


FIG. 7—This figure shows a possible orientation of the ball to separator normal and traction force vector and the vector sum.

$$\sum N_x = \sum_{\varphi=0}^{2\pi} N(1 + \mu_s^2)^{1/2} \cos(\varphi - \vartheta_s) \quad (29)$$

$$\sum N_y = \sum_{\varphi=0}^{2\pi} N(1 + \mu_s^2)^{1/2} \sin(\varphi - \vartheta_s) \quad (30)$$

$$\tan \vartheta_s = \mu_s \quad (31)$$

$$\sum M_{N,B,C} = \frac{E}{2} \sum_{\varphi=0}^{2\pi} N \quad (32)$$

We now need to satisfy the equilibrium component forces $\sum N_x$, $\sum N_y$ and the moments $\sum M_{N,B,C}$ with S and L . L is located at some angular position φ_l around the bearing center. Figure 8 illustrates the force system for a clockwise inner race rotation ω_i . In this figure, the controlling ball is positioned at the top, so that $\varphi_c = 0$. In this case, the reaction forces that the separator pocket experiences in the x - and y -directions are given by

$$S_x = -S \quad (33)$$

$$S_y = -\mu_s S \quad (34)$$

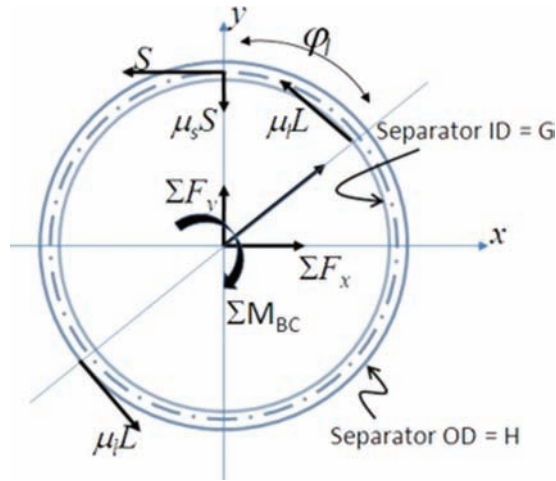


FIG. 8—The normal load and traction forces on the separator from the controlling ball and piloting land are shown here. ID, inner diameter; OD, outer diameter.

For the more general case of $\varphi_c \neq 0$,

$$S_x = -S \cos \varphi_c + \mu_s S \cos \left(\varphi_c + \frac{\pi}{2} \right) \quad (35)$$

$$S_y = -S \sin \varphi_c + \mu_s S \sin \left(\varphi_c + \frac{\pi}{2} \right) \quad (36)$$

$$\sum M_{s,BC} = \frac{E}{2} S \quad (37)$$

But

$$\cos \left(\varphi_c + \frac{\pi}{2} \right) = \sin \varphi_c \quad (38)$$

$$\sin \left(\varphi_c + \frac{\pi}{2} \right) = \cos \varphi_c \quad (39)$$

Thus,

$$S_x = -S(\cos \varphi_c + \mu_s \sin \varphi_c) \quad (40)$$

$$S_y = -S(\sin \varphi_c + \mu_s \cos \varphi_c) \quad (41)$$

Returning to Fig. 9, the land force L and its friction component $\mu_l L$ combine to produce a vector sum

$$L_v = (L^2 + \mu_l^2 L^2)^{1/2} = L(1 + \mu_l^2)^{1/2} \quad (42)$$

Depending upon whether we have an inner or outer race control situation, L_v will produce a force with x and y components as follows:

Inner Race Control

$$L_x = L_v \sin(\varphi_l - \vartheta_l) \quad (43)$$

$$L_y = L_v \cos(\varphi_l - \vartheta_l) \quad (44)$$

Outer Race Control

$$L_x = L_v \sin(\varphi_l + \vartheta_l) \quad (45)$$

$$L_y = L_v \cos(\varphi_l + \vartheta_l) \quad (46)$$

$$\sum M_{l,BC} = \mu_l LR \quad (47)$$

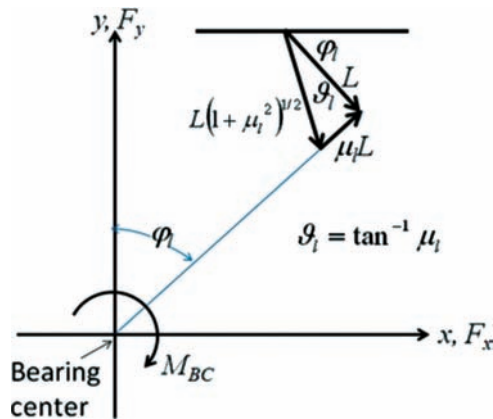


FIG. 9—This figure shows a possible orientation of the inner ring piloting land to separator normal and traction force vectors and the vector sum.

where:

$R = G/2$ for outer race control,

$R = H/2$ for outer race control, and

G and H are the inner diameter and outer diameter of the separator, as shown in Fig. 8.

Let $\lambda = \varphi_l \mp \vartheta_l$, where the minus sign applies for inner ring control and the plus sign applies for outer ring control. We can now write the equilibrium equations for the separator as

$$\sum F_x = \sum N_x + S_x + L_v \sin \lambda = 0 \quad (48)$$

$$\sum F_y = \sum N_y + S_y + L_v \cos \lambda = 0 \quad (49)$$

$$\sum M_{BC} = \sum M_{N,BC} - \frac{E}{2} S - \mu_l LR = 0 \quad (50)$$

Next we rewrite the equilibrium equations, combining them with Eqs 32, 39, and 40 and making the following substitutions:

$$A = -(\cos \varphi_c + \mu_s \sin \varphi_c) \quad (51)$$

$$B = -(\sin \varphi_c + \mu_s \cos \varphi_c) \quad (52)$$

$$C = \frac{\mu_l}{(1 + \mu_l^2)^{1/2}} R \quad (53)$$

The final form of the equilibrium equations becomes

$$F_1(S, L_v, \lambda) = \sum N_x + AS + L_v \sin \lambda = 0 \quad (54)$$

$$F_2(S, L_v, \lambda) = \sum N_y + BS + L_v \cos \lambda = 0 \quad (55)$$

$$F_3(S, L_v) = \sum N - S - \frac{2C}{E} L_v = 0 \quad (56)$$

We now have a system of three equations and three unknowns: S , L_v , and λ . Because λ is a nonlinear term, we can solve the equations through an iterative process. The solution values that satisfy $F_1 = F_2 = F_3 = 0$ are derived through initial estimates for the values S_0 , L_{v0} , and λ_0 and the determination of changes via the Newton-Raphson method.

$$\Delta S_0 = \frac{\begin{vmatrix} -F_{1,0} & \frac{\partial F_1}{\partial L_v} & \frac{\partial F_1}{\partial \lambda} \\ -F_{2,0} & \frac{\partial F_2}{\partial L_v} & \frac{\partial F_2}{\partial \lambda} \\ -F_{3,0} & \frac{\partial F_3}{\partial L_v} & \frac{\partial F_3}{\partial \lambda} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F_1}{\partial S} & \frac{\partial F_1}{\partial L_v} & \frac{\partial F_1}{\partial \lambda} \\ \frac{\partial F_2}{\partial S} & \frac{\partial F_2}{\partial L_v} & \frac{\partial F_2}{\partial \lambda} \\ \frac{\partial F_3}{\partial S} & \frac{\partial F_3}{\partial L_v} & \frac{\partial F_3}{\partial \lambda} \end{vmatrix}} \quad (57)$$

$$\Delta L_{v0} = \frac{\begin{vmatrix} \frac{\partial F_1}{\partial S} & -F_{1,0} & \frac{\partial F_1}{\partial \lambda} \\ \frac{\partial F_2}{\partial S} & -F_{2,0} & \frac{\partial F_2}{\partial \lambda} \\ \frac{\partial F_3}{\partial S} & -F_{3,0} & \frac{\partial F_3}{\partial \lambda} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F_1}{\partial S} & \frac{\partial F_1}{\partial L_v} & \frac{\partial F_1}{\partial \lambda} \\ \frac{\partial F_2}{\partial S} & \frac{\partial F_2}{\partial L_v} & \frac{\partial F_2}{\partial \lambda} \\ \frac{\partial F_3}{\partial S} & \frac{\partial F_3}{\partial L_v} & \frac{\partial F_3}{\partial \lambda} \end{vmatrix}} \quad (58)$$

$$\Delta \lambda_0 = \frac{\begin{vmatrix} \frac{\partial F_1}{\partial S} & \frac{\partial F_1}{\partial L_v} & -F_{1,0} \\ \frac{\partial F_2}{\partial S} & \frac{\partial F_2}{\partial L_v} & -F_{2,0} \\ \frac{\partial F_3}{\partial S} & \frac{\partial F_3}{\partial L_v} & -F_{3,0} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F_1}{\partial S} & \frac{\partial F_1}{\partial L_v} & \frac{\partial F_1}{\partial \lambda} \\ \frac{\partial F_2}{\partial S} & \frac{\partial F_2}{\partial L_v} & \frac{\partial F_2}{\partial \lambda} \\ \frac{\partial F_3}{\partial S} & \frac{\partial F_3}{\partial L_v} & \frac{\partial F_3}{\partial \lambda} \end{vmatrix}} \quad (59)$$

$$\frac{\partial F_1}{\partial S} = A, \quad \frac{\partial F_2}{\partial S} = B, \quad \frac{\partial F_3}{\partial S} = -1$$

$$\frac{\partial F_1}{\partial L_v} = \sin \lambda, \quad \frac{\partial F_2}{\partial L_v} = \cos \lambda, \quad \frac{\partial F_3}{\partial L_v} = -\frac{2C}{E}$$

$$\frac{\partial F_1}{\partial \lambda} = L_v \cos \lambda, \quad \frac{\partial F_2}{\partial \lambda} = -L_v \sin \lambda, \quad \frac{\partial F_3}{\partial \lambda} = 0$$

Computation of Total Ball Speed Variation Drag Torque

We have now solved for the forces at the ball–race interfaces F_i , the ball–pocket interface of the controlling ball S and $\mu_s S$, the ball–pocket interfaces of the remaining balls N and $\mu_s N$, and the separator–land forces L and $\mu_i L$. We have demonstrated how an energy balance method is used to determine the traction input force that is necessary in order to overcome the force due to energy loss at the ball–race interface. A similar method is used to determine the force required in order to overcome the energy loss due to other drag forces shown here. As these input forces are conveyed through the rotating inner ring at the contact between the ball and race, the drag torque component is simply the product of the force and the radius, $1/2(E - d \cos \beta_i)$.

Validation of Model

The BSV mechanical force model described above has been built into a bearing computational analysis program called DYBA using Fortran. The program requires detailed knowledge of the bearing geometry, material properties, and operating conditions. The program generates inputs for the bearing model, such as the operating contact angles and individual ball loads, and then receives drag torque values for each of the drag sources described above. This calculated BSV is then added to the nominal bearing drag, which is calculated separately within DYBA, to obtain the total bearing drag. The nominal bearing drag is segregated into estimates for the “load component,” due to Coulombic friction and material hysteresis at the ball to race interfaces, and for the “lube component,” due to viscous losses caused by shearing of the lubricant.

Over the past several years, this program has been used frequently in the design of new aerospace bearing applications and in the analysis of anomalies in existing applications. Although much of the data from those programs are not publicly available, the data described here appeared in a report to the U.S. Air Force in 1967 [5]. This paper reports a study in which cage design features and the external loads placed on the bearing were varied in order to allow observation of their effects on drag torque. A 204 sized bearing was operated at 5000 rpm using a light oil (MIL-L-7808). Specific details of the bearing design are listed in Table 2. The configuration of the experimental test fixture is shown

TABLE 2—Details of the bearing design used to validate the mathematical model.

Bore	0.7874 in.
Outside diameter	1.8504 in.
Ball diameter	0.3125 in.
Number of balls	8
Pitch diameter	1.330 in.
Outer ring curvature	0.55
Inner ring curvature	0.53
Bearing materials	52100 steel

schematically in Fig. 10. In all the cases we discuss, the land control was at the inner land, and the diametric land clearance was 0.020 in.

Eight bearing configurations were tested. The bearing configuration variables are summarized in Table 3, and testing included all eight permutations of these variables. The external bearing loads were set at 70 lb pure thrust load (L1) and 70 lb thrust load plus 10 lb radial load. The 10 lb radial load in the second loading arrangement (L2) was hung off center by 1 in. The circumferential ball pocket clearance was set at 0.005 in. (C1) and 0.020 in. (C2), and the transverse pocket clearance was set at 0.005 in. (T1) and 0.020 in. (T2). The variability in transverse clearance was not tested using the DYBA BSV model because it considers separator forces only in the circumferential

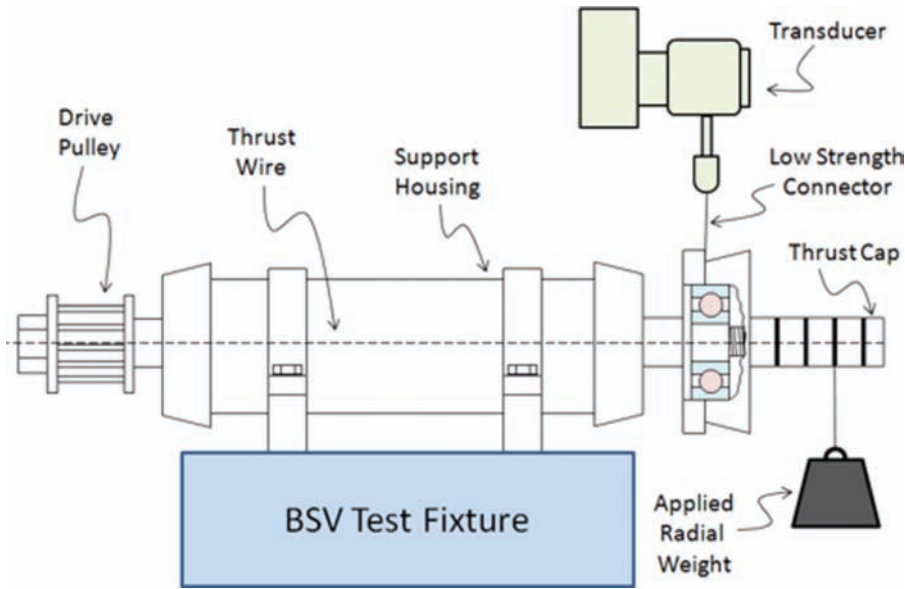


FIG. 10—Schematic diagram of the test fixture used to collect the data shown in Table 4 and Fig. 11 (not drawn to scale).

TABLE 3—Variables used in the bearing tests.

	1	2
Load (<i>L</i>)	70 lb thrust + 0 lb radial	70 lb thrust + 10 lb radial
Ball/pocket circumferential clearance (<i>C</i>)	0.005 in.	0.020 in.
Ball/pocket transverse clearance (<i>T</i>)	0.005 in.	0.020 in.

direction. Each test combination was run three times, and the eight combinations were randomized. Each of the eight test combinations was run with a bearing different from those in the general population for that design. A total of 12 different bearings were used in four different retainer geometry configurations, with each configuration tested three times at two different loads. Each replication represented a different test setup; the same bearing was not used in the same setup three times in a row.

The test results and analytical predictions are both listed in Table 4. The first load condition has no radial load and should therefore produce no BSV forces. Because there is no sustained contact with the separator, the pocket dimensions should not affect the forces. Thus, the theoretical predictions are consistent through the first four columns, representing all data with no radial load. The observed variation in the test drag torque is due to uncontrolled experimental variability, which is described in the source article.

With the application of a 10 lb radial load (see the last four columns of Table 4), we observe the onset of BSV drag in the calculation and an increase in measured drag in the bearing tests. The prediction of 0.398 oz in. drag torque with more open circumferential pocket clearance increased to 0.656 oz in. when this clearance was reduced. This reduction was expected, because the ball–pocket

TABLE 4—Drag torque predictions and measured values, in ounce-inches, for the eight different experimental configurations. *L1* and *L2* are two different loading conditions, *C1* and *C2* are two different circumferential ball pocket clearances, and *T1* and *T2* are transverse ball pocket clearances of 0.005 in and 0.020 in., respectively

	L1: 70 lb thrust + 0 lb radial				L2: 70 lb thrust + 10 lb radial			
	C1: 0.005 in.		C2: 0.020 in.		C1: 0.005 in.		C2: 0.020 in.	
	T1	T2	T1	T2	T1	T2	T1	T2
Test 1	1.08	0.77	1.17	1.23	1.52	1.11	1.50	1.38
Test 2	1.23	1.40	1.57	1.12	1.86	1.70	2.12	1.56
Test 3	1.05	0.95	1.38	1.03	1.31	1.21	1.18	1.38
DYBA lube torque	0.292	0.292	0.292	0.292	0.266	0.266	0.266	0.266
DYBA load torque	0.944	0.944	0.944	0.944	0.979	0.979	0.979	0.979
DYBA BSV torque	0.000	0.000	0.000	0.000	0.656	0.656	0.398	0.398
DYBA total torque	1.236	1.236	1.236	1.236	1.890	1.890	1.643	1.643
Average test data	1.15	1.09	1.34	1.15	1.65	1.48	1.61	1.49

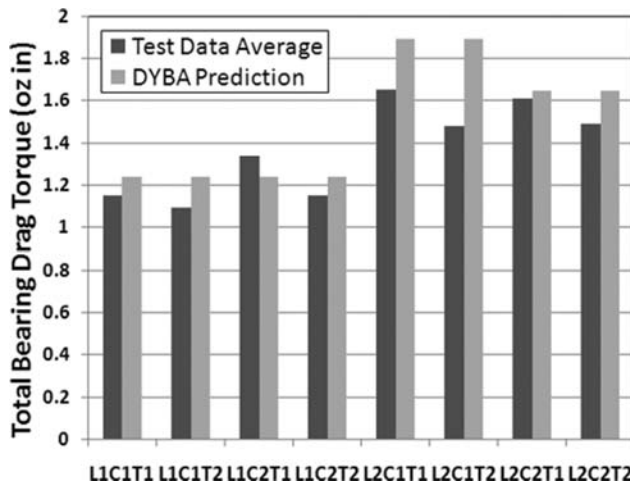


FIG. 11—Comparison of the total bearing drag torque including experimental test data and the DYBA prediction using the model presented here.

interaction is reduced with greater clearance. The experimental values show a comparable increase in drag torque with the application of the radial load. However, there was no discernable effect of the circumferential ball/pocket clearance. We suggest that this is because the expected change (0.246 oz in., from the DYBA BSV model) is within the range of experimental variability (0.25 oz in., from the pure thrust load testing). Note that the calculated values for DYBA load torque and DYBA lube torque are also dependent on the loading configuration and cage geometry. This is because the applied radial load results in higher losses at the ball to race interface, and opening the pocket clearances results in more retainer centrifugal forces at the land interface, with both effects resulting in a higher load torque. These effects also cause higher calculated temperatures and reduced viscous losses, resulting in a reduction of the lube torque component.

Figure 11 shows a graphical representation of the data listed in Table 4. We find that the DYBA prediction agrees well with the observed increase in drag torque upon application of the radial load in the case with greater circumferential clearance (L1C2 to L2C2). The DYBA prediction shows a further increase in BSV force with a decrease in circumferential clearance (L2C2 to L2C1), as expected. However, the experimental data show only a very small increase, probably due to uncontrolled experimental variables.

Conclusion

A mathematical model of the primary mechanical forces involved in ball speed variation (BSV) was derived. This model includes forces at the ball to race, ball to separator, and separator to piloting land interfaces. These forces can be

converted to individual sources of bearing drag torque. When added to the nominal bearing drag, the results of this model provide a quantitative estimate of the total bearing drag under conditions of misaligned or radial loading. This analytic tool has been applied to the design of new bearing systems and in the study of anomalous behavior with existing systems. The results have been favorable, contributing to several successful designs.

In the study presented here, we compared the model predictions with the results of bearing tests conducted in a laboratory. We found excellent agreement with the observed increase in total bearing drag when the loading arrangement was changed from a pure thrust load to a thrust load plus a radial load. The model also provided reasonable predictions for growth in the BSV drag when the ball pocket clearance was reduced, but the experimental data showed only a very small increase.

Acknowledgments

The writers gratefully acknowledge support from The Aerospace Corporation's Independent Research and Development Program.

References

- [1] Barish, T., "Ball Speed Variation in Ball Bearings and Its Effect on Cage Design," *Lubr. Eng.*, Vol. 25, 1969, pp. 110–116.
- [2] Jones, A. B., "A General Theory for Elastically Constrained Ball and Radial Roller Bearings under Arbitrary Load and Speed Conditions," *J. Basic Eng.*, Vol. 82, 1960, pp. 309–320.
- [3] Jones, A. B., "Ball Motion and Sliding Friction in Ball Bearings," *J. Basic Eng.*, Vol. 81, 1959, pp. 1–12.
- [4] Harris, T. A., *Rolling Bearing Analysis*, 2nd ed., Wiley, New York, 1984.
- [5] Leveille, A. R., Pelletier, R. R., and Gianotti, R. J., "Aerospace Bearing—Analytical Investigation," *AFFDL-TR-66-134*, Wright-Patterson Air Force Base, Dayton, OH, 1967.

*Yoshimi R. Takeuchi, Sean E. Davis, Matthew A. Eby,
Jerry K. Fuller, Donald L. Taylor, and Michael J. Rosado*

Bearing Thermal Conductance Measurement Test Method and Experimental Design

REFERENCE: Takeuchi, Yoshimi R., Davis, Sean E., Eby, Matthew A., Fuller, Jerry K., Taylor, Donald L., and Rosado, Michael J., "Bearing Thermal Conductance Measurement Test Method and Experimental Design," *Rolling Element Bearings* on April 13–15, 2011 in Anaheim, CA; STP 1542, Yoshimi R. Takeuchi and William F. Mandler, Editors, pp. 92–117, doi:10.1520/STP104233, ASTM International, West Conshohocken, PA 2012.

ABSTRACT: Bearings are used in a number of spacecraft applications, ranging from minimal motion devices, such as pointing mechanisms, to high-speed components, such as control moment gyroscopes and reaction and momentum wheels. Terrestrial applications include pumps, axles, and tooling. Heat-transfer modes for rotational systems in a vacuum environment differ from their terrestrial counterpart. In space, with the absence of air, heat transfer consists of radiation from the rotating system and conductance through the bearings themselves. Depending on the application, conductance could dominate the effects on bearing temperatures and thermal gradients. Accurate thermal predictions are important because they can drive life and performance requirements. To accurately predict bearing temperatures, basic bearing thermal conductance data was needed. However, bearing thermal conductance tends to be the most significant unknown in a rotational system in the space environment. To address this shortcoming, this paper explores a new vacuum test rig designed to measure bearing conductance under simulated operational conditions. Experimental variables include control of the bearing rotational speed, applied axial load, and average bearing temperature and temperature gradient via an applied heat source/heat sink mechanism. All tests are conducted in vacuum. The experimental variables studied herein allowed parametric studies to be conducted

Manuscript received July 25, 2011; accepted for publication January 10, 2012; published online April 2012.

Engineering and Technology Group, The Aerospace Corporation, 2310 E. El Segundo Blvd., El Segundo, CA 90245.

All trademarks, service marks, and tradenames are the property of their respective owners.

Copyright © 2012 by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

under controlled thermal and mechanical conditions, permitting the exploration of the influences of those operational variables on bearing thermal conductance. This paper will describe the test method, the use of uncertainty analysis to design the experiment, and a verification study.

KEYWORDS: bearing thermal conductance, bearing thermal resistance, experiment, uncertainty analysis, design

Nomenclature

- A = shaft heat meter cross-sectional area
- a = major axis of the Hertzian contact ellipse
- b = minor axis of the Hertzian contact ellipse
- d = shaft heat meter diameter
- G_b = bulk thermal conductance across the test bearing
- h = flange heat meter flange thickness
- k_b = thermal conductivity of the ball material
- k_{fhm} = thermal conductivity of the flange heat meter
- k_i = thermal conductivity of the inner race material
- k_o = thermal conductivity of the outer race material
- k_s = thermal conductivity of the shaft material
- L_s = shaft pyrometer separation distance
- Q_{in} = input heat entering the test bearing inner race
- Q_{out} = output heat exiting the test bearing outer race
- R = calculated variable based on measured values x_1 through x_n
- R_b = total thermal resistance across a single bearing ball
- R_i = thermal resistance from the inner race to the ball in a bearing
- R_o = thermal resistance from the outer race through the ball of a bearing
- r_1 = radial distance to the flange heat meter inner thermocouple
- r_2 = radial distance to the flange heat meter outer thermocouple
- T_{fhm1} = measured flange heat meter inner temperature
- T_{fhm2} = measured flange heat meter outer temperature
- T_{in} = measured test bearing inner race temperature
- T_{out} = measured test bearing outer race temperature
- T_{s1} = measured lower shaft pyrometer temperature
- T_{s2} = measured upper shaft pyrometer temperature
- x_n = measured value to which R is sensitive
- δx_n = uncertainty in measurement of x_n

Introduction

Accurate temperature predictions for a bearing in a vacuum environment call for knowledge of two independent thermal parameters: bearing thermal conductance and heat generation. Heat generation, which is related to torque and speed, tends to be the better-known quantity. Models and software programs have previously been designed to calculate and quantify torque (i.e., software codes BRGS10C, developed by Alan Leveille, and the Jones software

developed by A. B. Jones). Torque and speed tend to be a known for a given application; thus, the focus of this study was to develop a means of experimentally quantifying bearing thermal conductance.

The only well-established analysis used to calculate bearing thermal conductance, developed by Yovanovich, is for dry (non-lubricated), static (non-moving) bearings [1,2]. A need exists to expand this capability to lubricated, dynamic bearings. In the past, a small handful of test rigs were constructed with the purpose of obtaining bearing thermal conductance data [2–13]. The first successful dynamic test rig was designed and constructed by Delil et al. [5,6]. This test rig was used to measure thermal conductance across a moving bearing under controlled thermal and mechanical conditions by other researchers [7–9]; however, the maximum bearing speeds were limited to approximately 2500 rpm. Most data collected was within the range of hundreds of rpm. Over the years, demands on bearings have increased and speeds of 3000 rpm and above in a vacuum environment are common. Takeuchi et al. designed a test facility to measure thermal conductance of small 101-sized bearings at these higher speeds to address data needs [10–12], but industrial demands have motivated development of capabilities to accommodate larger bearings.

A new test rig, described herein, was designed and constructed to measure the bulk effective thermal conductance across a bearing in a vacuum environment. The test rig was designed to control several thermal and mechanical constraints including speed, axial load, and thermal profile for the larger bearings up to 305-size. This paper describes the test facility's characteristics, discusses the use of uncertainty analysis to optimize the experimental measurement accuracy, and verifies the experimental method by comparing pilot measurements with the best-known analytical model available—the Yovanovich model for calculating thermal conductance across a dry, static bearing [1,2]. A follow-up paper [13] will describe how the test facility was used to conduct parametric studies and measure bearing thermal conductance on lubricated, dynamic (constant speed) bearings.

Test Capabilities and Design

Design of the test rig was based on the bearing thermal conductance equation:

$$G_b = \frac{1}{2} \cdot \frac{Q_{in} + Q_{out}}{T_{in} - T_{out}} \quad (1)$$

where G_b is the test bearing conductance, Q_{in} and Q_{out} the heat entering and leaving the test bearing, respectively, and T_{in} and T_{out} the temperature of the test bearing at the inner and outer races, respectively. The derivation of this equation is shown in Appendix A. The heat and temperature variables represent

the basic measurements that need to be recorded to determine the test-bearing thermal conductance.

The variables that need to be controlled include both the thermal and mechanical conditions that are applied to the test bearing. Speed, axial loading, temperature, bearing materials, and presence/absence of lubricant are monitored and varied under controlled conditions for parametric studies. These parametric studies are needed to separate out the impact of each variable and understand how operational conditions, materials, and lubricant influence the bearing thermal conductance.

The test rig is shown in Fig. 1. The test fixture assembly is enclosed in a vacuum chamber to eliminate convective heat-transfer effects. A data acquisition system is used to control the experiment and collect test data.

Figure 2 shows a schematic representation of the test rig. The particular configuration used to date was designed to accommodate a single 305-size bearing. A separate configuration that accommodates a 305-bearing pair has been built and results will be reported in the future. A prescribed axial (thrust) load is applied to the test bearing using an in-house developed axial load device, which is essentially a stepper motor connected to the load frame. The

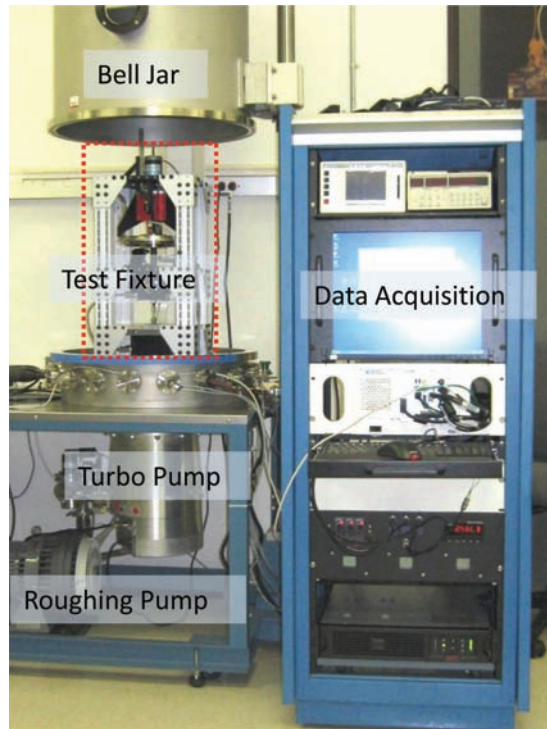


FIG. 1—*Photograph of the bearing thermal test chamber.*

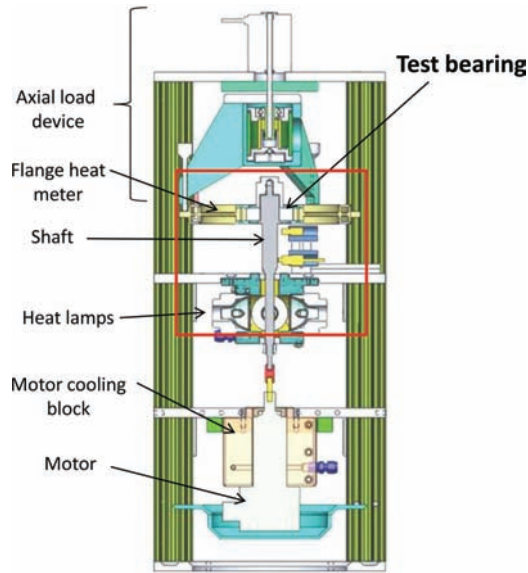


FIG. 2—Schematic diagram of the test fixture.

axial load device is used to apply a downward force on the outer race, whereas the inner race is held in place by a shoulder on the shaft.

The speed of the bearing inner race is controlled with a motor and spindle assembly, whereas the outer race is held stationary. An optical sensor is used to monitor the rotational speed of the shaft.

The thermal system was designed to be based on one-dimensional heat transfer, and is highlighted in a red box in Fig. 2. It is shown in more detail in Fig. 3. The heat sink consists of a cooling channel around the outer diameter of the flange heat meter. The thermal source consists of a heat lamp, which emits

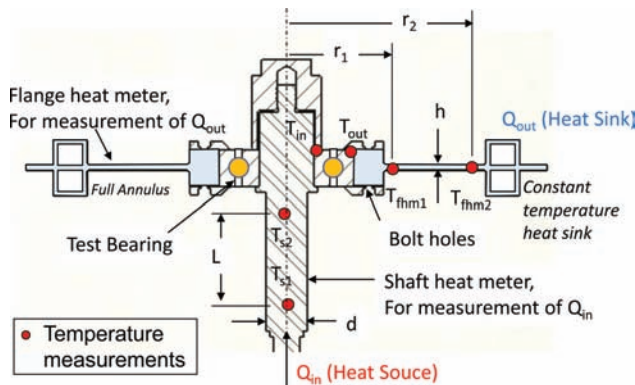


FIG. 3—Schematic diagram of the thermal system in the test fixture.

radiated energy. This thermal energy is absorbed into the lower portion of the shaft. The lamp intensity may be varied. In a vacuum, the heat traveling up the shaft may be assumed to enter the test bearing, as convection is absent and blankets and surface finishes have been designed to minimize radiative losses. The total heat entering the test bearing consists of the heat absorbed by the shaft from the lamps and, if the shaft is rotating, the heat that is generated by the support bearings.

The amount of heat entering the test bearing is measured using the shaft heat meter, which consist of two non-contacting infrared (IR) pyrometer sensors aimed at the shaft, measuring temperature at two discrete locations. These shaft temperature measurements, T_{s1} , and T_{s2} , at a known separation distance, L_s , made on a shaft with a known cross-sectional area, A , provide all the information needed to know the heat entering the test bearing, Q_{in} . The heat transferred to the test bearing is computed using the temperature difference measured via the pyrometers and knowledge of the shaft geometry; because one-dimensional (1D) heat transfer may be assumed, the heat calculation is straightforward.

Temperature of the inner race, T_{in} , is measured along a moving component, and was thus sampled using a non-contacting IR pyrometer sensor. The temperature of the outer race, T_{out} , measured at a stationary location, was sampled using a type E thermocouple. The choice of thermocouples for stationary components was preferred because of the ability to minimize temperature measurement and sensor location errors as compared to the pyrometers.

The heat transfer to and across the bearing and into the heat sink is also designed to adhere to the one-dimensional heat-transfer process. The heat travels up the shaft, through the test bearing, and then flows in a radial direction through the flange heat meter to the cooling channel that encompasses the flange heat meter outer diameter. The flange heat meter is based on the design of the heat flux meter originally conceived by Delil et al. [5]. The total heat leaving the bearing, Q_{out} , is determined by taking two temperature measurements, T_{fhm1} and T_{fhm2} , via E-type thermocouples at specified radial distances along the flanged part of the meter.

Experimental Design Optimization

Ultimately, the goal of this experiment is to obtain bearing thermal conductance data within reasonable error bounds. To accomplish this goal, an uncertainty analysis is used to (1) optimize the design of the experiment, (2) to study the influence of errors in each variable, and (3) to provide the experimentalist with an idea of the degree of accuracy required to maintain reasonable uncertainty bounds [14,15]. The following sections explain how the experimental optimization was performed.

Uncertainty Analysis

An uncertainty analysis is useful in optimizing the design of an experiment where the goal is to obtain an indirect measurement. An indirect measurement, e.g., bearing thermal conductance and heat measurements, is a quantity derived from multiple direct measurements, such as temperature measurements at different locations. The basic concept behind uncertainty analysis is that the total uncertainty bound for the indirect measurement is the root mean square sum of the partial derivative of each term that contributes to the total uncertainty multiplied by the uncertainty of that particular error term. The uncertainty analysis equation is summarized as follows:

$$\delta R = \left[\left(\frac{\partial R}{\partial x_1} \cdot \delta x_1 \right)^2 + \left(\frac{\partial R}{\partial x_2} \cdot \delta x_2 \right)^2 + \cdots + \left(\frac{\partial R}{\partial x_n} \cdot \delta x_n \right)^2 \right]^{1/2} \quad (2)$$

where R is the indirect measurement, $x_1, x_2, \dots, x_n$ are measured variables that could potentially contribute to errors, and $\delta x_1, \delta x_2, \dots, \delta x_n$ are the corresponding uncertainties of each of the measurements [14,15].

The advantage of deriving an equation for the calculation of the uncertainty bounds of a given measurement is that parametric studies can be conducted on the influence of each variable; thus providing the user with a means of selecting the best parameter value for an optimal experiment.

Derivation of Bearing Thermal Conductance Uncertainty

Equation 1 shows that bearing thermal conductance is dependent on the heat entering and exiting the bearing, as well as the temperatures of the inner and outer bearing races. Assuming that each variable could contribute to the total error in thermal conductance, and applying the uncertainty method to Eq 1, we have:

$$\delta G_b = \left[\left(\frac{\partial G_b}{\partial Q_{in}} \cdot \delta Q_{in} \right)^2 + \left(\frac{\partial G_b}{\partial Q_{out}} \cdot \delta Q_{out} \right)^2 + \left(\frac{\partial G_b}{\partial T_{in}} \cdot \delta T_{in} \right)^2 + \left(\frac{\partial G_b}{\partial T_{out}} \cdot \delta T_{out} \right)^2 \right]^{1/2} \quad (3)$$

Solving and rearranging the equation for the bearing thermal conductance % uncertainty we have:

$$\frac{\delta G_b}{G_b} \cdot 100\% = \left[\frac{\delta Q_{in}^2 + \delta Q_{out}^2}{(Q_{in} + Q_{out})^2} + \frac{\delta T_{in}^2 + \delta T_{out}^2}{(T_{in} + T_{out})^2} \right]^{1/2} \cdot 100\% \quad (4)$$

Note that the equation indicates that bearing uncertainty depends on both operational variables, such as heat and temperature, as well as the measurement errors for each of those operational variables. Temperature errors are restricted to the accuracies of the hardware. However, the heat terms are themselves indirect quantities that depend on temperature, hardware geometry, and knowledge of the material properties of the test rig components. The shaft and flange heat meter design geometries may be optimized to minimize the uncertainties in heat measurement, which ultimately impact the bearing thermal conductance uncertainty.

Optimization of Parameters

The next sections outline the optimization method for the shaft heat meter design parameters. A similar methodology implemented for the flange heat meter. In practice, many iterations were necessary given the multiple inter-related design dimensions requiring optimization, but for documentation purposes, the following parametric studies were conducted on one variable at a time.

Shaft Heat Meter

The shaft heat meter consists of two IR pyrometers focused at two locations along a shaft of a fixed length and diameter. This setup allows the determination of heat input to the test bearing.

Shaft Heat Meter Uncertainty—The equation to calculate the heat entering the bearing, Q_{in} , is

$$Q_{in} = \frac{k_s \cdot A}{L} \cdot (T_{s1} - T_{s2}) \quad (5)$$

where k_s is the shaft conductivity, A the shaft cross-sectional area, T_{s1} and T_{s2} the shaft temperature as measured at the lower and upper pyrometer spot locations, respectively, and L the pyrometer spot separation distance.

Applying uncertainty analysis to the shaft heat meter, we have:

$$\delta Q_{in} = \left[\left(\frac{\partial Q_{in}}{\partial d} \cdot \delta d \right)^2 + \left(\frac{\partial Q_{in}}{\partial L} \cdot \delta L \right)^2 + \left(\frac{\partial Q_{in}}{\partial k_s} \cdot \delta k_s \right)^2 + \left(\frac{\partial Q_{in}}{\partial T_{s1}} \cdot \delta T_{s1} \right)^2 + \left(\frac{\partial Q_{in}}{\partial T_{s2}} \cdot \delta T_{s2} \right)^2 \right]^{1/2} \quad (6)$$

Rearranging the equation and solving for % uncertainty, we have:

$$\frac{\delta Q_{\text{in}}}{Q_{\text{in}}} \cdot 100\% = \left[\left(\frac{2 \cdot \delta d}{d} \right)^2 + \left(\frac{\delta L}{L} \right)^2 + \left(\frac{\delta k_s}{k_s} \right)^2 + \frac{\delta T_{s1}^2 + \delta T_{s2}^2}{(T_{s1} - T_{s2})^2} \right]^{1/2} \cdot 100\% \quad (7)$$

There are four basic terms that govern the heat input measurement % uncertainty: shaft diameter, pyrometer separation distance, thermal conductivity, and shaft temperatures. These terms are not necessarily independent of each other. Shaft diameter and pyrometer separation distance, for instance, influence the shaft temperature terms. The geometric terms are design parameters, which can be selected for optimal test performance.

Shaft Heat Meter Optimization: Shaft Diameter—The shaft diameter influences the thermal gradient across the shaft heat meter, and, as such, impacts the uncertainty in input heat measurement. A parametric study was conducted to determine the effect the shaft diameter has on the heat input uncertainty bounds. The shaft material was chosen to be 410-stainless steel, because of mechanical and machining considerations. Analytical assumptions include:

1. Pyrometer temperature measurement errors were estimated to be approximately $\pm 1.0^\circ\text{C}$.
2. The thermal conductivity of 410-stainless steel (the shaft material) was known to within 10 %.
3. The pyrometer separation distance was to be measured to within ± 0.1 in.
4. The pyrometer separation distance was set at 1.6 in.

One purpose of the shaft heat meter is to allow quantification of the heat input into the test bearing. The associated uncertainty bounds for these measurements were calculated as a function of shaft diameter. The total heat input also influences the uncertainty in the results, so this value was also varied and the resulting effects are shown in Fig. 4. The shaft heat meter diameter, d , can be optimized, but the influence of the operational variable, Q_{in} , impacts the required test conditions as well; this effect that will be discussed in more detail in a later section.

The curves in Fig. 4 imply that a thinner shaft produces a larger temperature gradient and; thus, leads to the desired result: lower uncertainty bounds on the heat input measurement. The limit of how thin the shaft can be, however, was constrained by the need for mechanical rigidity, which limits the minimum shaft diameter.

The test rig went through several design iterations. Originally, the shaft was 1 in. in diameter, to ensure rigidity of the shaft. However, this left the heat input uncertainty unreasonably high. This result was magnified for lower heat inputs. A closer look showed that a slight decrease in shaft diameter to 0.8 in.

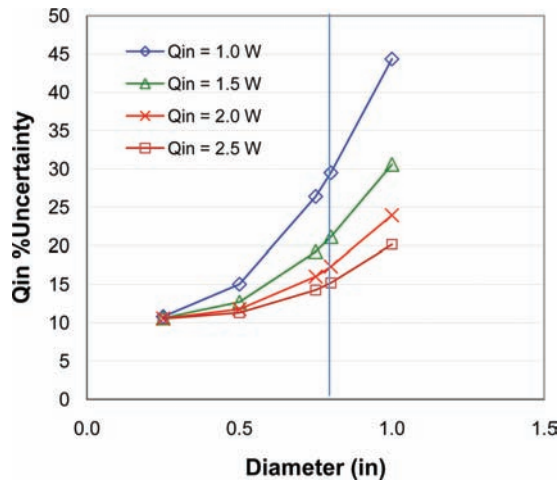


FIG. 4—Influence of shaft diameter on input heat measurement uncertainty.

dramatically improved the accuracy of the heat measurement. It was decided that the shaft rigidity would not be compromised with this slight decrease in diameter, and later tests confirmed that this assumption was valid.

Shaft Heat Meter Optimization: Pyrometer Separation Distance—Pyrometer separation distance also influences heat input measurement and its uncertainty. The assumptions for this analysis include:

1. Pyrometer temperature measurement errors were estimated to be approximately $\pm 1.0^{\circ}\text{C}$.
2. The thermal conductivity of 410-stainless steel (the shaft material) was known to within 10 %.
3. The pyrometer separation distance was to be measured to within ± 0.1 in.
4. The shaft diameter for this analysis was set to 0.8 in.

Figure 5 shows the influence of heat and pyrometer separation distance on heat input measurement and its uncertainty bounds. The figure shows that uncertainty decreases with increasing pyrometer separation distance. The increased amount of input heat will also decrease the measurement uncertainty bounds.

Note that a non-linear effect is evident where the degree of benefit diminishes at larger separation distances or for higher heat inputs. The reason a larger pyrometer separation distance results in lower uncertainties is that a greater distance between shaft temperature measurement locations results in higher measured temperature gradients, leading to smaller contributions to heat uncertainty from the temperature term. Eventually, with large thermal

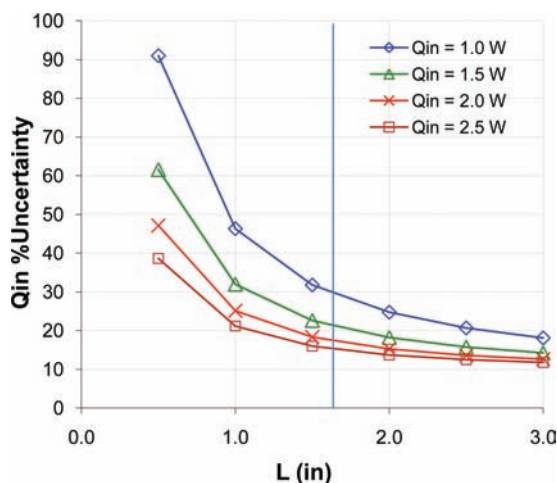


FIG. 5—Influence of pyrometer separation distance on input heat measurement uncertainty.

gradients, the temperature term becomes insignificant and errors from other terms, such as the shaft conductivity, dominate.

Whereas the above analysis suggests using as long a shaft length as possible, the need for mechanical rigidity limits the maximum length. Taking both constraints into consideration, the pyrometer separation distance was selected to be 1.6 in.

Shaft Heat Meter Optimization: Summary—There were two dimensions on the shaft heat meter that were optimized for thermal measurement accuracy: shaft diameter and pyrometer separation distance. Minimizing the diameter and maximizing the length would result in improved thermal measurement performance, i.e., smaller uncertainty bounds, but both of these dimensional trends also create mechanical and structural instability. Ultimately, the relationship between geometry and uncertainty is non-linear and diminishing returns are evident. Thus, the dimensions chosen for the shaft design include:

1. Shaft diameter $d = 0.8$ in.
2. Pyrometer separation distance $L_s = 1.6$ in.

Flange Heat Meter

The flange heat meter consists of a disc-shaped device of a prescribed web thickness with a cooling channel located around the outer circumference and two thermocouples located at two different fixed radial distances along the thin metal webbing.

Flange Heat Meter Uncertainty—The equation describing the heat transfer through the flange heat meter, Q_{out} , is:

$$Q_{\text{out}} = G_{\text{fhm}} \cdot (T_{\text{fhm1}} - T_{\text{fhm2}}) \quad (8)$$

G_{fhm} represents the flange heat meter thermal conductance, and T_{fhm1} and T_{fhm2} the temperatures measured at the inner and outer flange heat meter thermocouple locations, respectively.

Applying an uncertainty analysis to the output heat equation, we have:

$$\delta Q_{\text{out}} = \left[\left(\frac{\partial Q_{\text{out}}}{\partial G_{\text{fhm}}} \cdot \delta G_{\text{fhm}} \right)^2 + \left(\frac{\partial Q_{\text{out}}}{\partial T_{\text{fhm1}}} \cdot \delta T_{\text{fhm1}} \right)^2 + \left(\frac{\partial Q_{\text{out}}}{\partial T_{\text{fhm2}}} \cdot \delta T_{\text{fhm2}} \right)^2 \right]^{1/2} \quad (9)$$

Rearranging and solving for % uncertainty:

$$\frac{\delta Q_{\text{out}}}{Q_{\text{out}}} \cdot 100 = \left[\left(\frac{\delta G_{\text{fhm}}}{G_{\text{fhm}}} \right)^2 + \left(\frac{\delta T_{\text{fhm1}}^2 + \delta T_{\text{fhm2}}^2}{(T_{\text{fhm1}} - T_{\text{fhm2}})^2} \right) \right]^{1/2} \cdot 100 \quad (10)$$

The flange heat meter conductance is a calculated quantity. The equation for thermal conductance across a thin disk with a hole in the center is:

$$G_{\text{fhm}} = \frac{2\pi \cdot h \cdot k_{\text{fhm}}}{\ln(r_2/r_1)} \quad (11)$$

where h is the disk thickness, k_{fhm} the flange heat meter thermal conductivity, and r_1 and r_2 the inner and outer thermocouple locations, respectively, on the disk.

Applying the uncertainty analysis and solving for % uncertainty, we get:

$$\frac{\delta G_{\text{fhm}}}{G_{\text{fhm}}} \cdot 100\% = \left[\frac{(r_2 \cdot \delta r_1)^2 + (r_1 \cdot \delta r_2)^2}{(r_1 \cdot r_2 \cdot \ln(r_2/r_1))^2} + \left(\frac{\delta h}{h} \right)^2 + \left(\frac{\delta k_{\text{fhm}}}{k_{\text{fhm}}} \right)^2 \right]^{1/2} \cdot 100\% \quad (12)$$

Substituting into Eq 11, we have the heat output % uncertainty:

$$\begin{aligned} \frac{\delta Q_{\text{out}}}{Q_{\text{out}}} \cdot 100\% = & \left[\frac{(r_2 \cdot \delta r_1)^2 + (r_1 \cdot \delta r_2)^2}{(r_1 \cdot r_2 \cdot \ln(r_2/r_1))^2} + \left(\frac{\delta h}{h} \right)^2 + \left(\frac{\delta k_{\text{fhm}}}{k_{\text{fhm}}} \right)^2 \right. \\ & \left. + \left(\frac{\delta T_{\text{fhm1}}^2 + \delta T_{\text{fhm2}}^2}{(T_{\text{fhm1}} - T_{\text{fhm2}})^2} \right) \right]^{1/2} \cdot 100\% \end{aligned} \quad (13)$$

There are four basic terms that govern the heat output % uncertainty: thermocouple placement, webbing thickness, thermal conductivity, and flange temperatures. These parameters are not necessarily independent of one another, e.g., the webbing thickness and thermocouple placement influences the measured temperatures. Once again, the geometric terms will be optimized to minimize the experimental uncertainty in the heat output measurement.

Flange Heat Meter Optimization: Thermocouple Placement—Two thermocouples are affixed at two different radial locations to measure the heat leaving the bearing. The first thermocouple placement, r_1 , is fixed because of the size of the bearing. The second thermocouple placement is determined using uncertainty analysis.

Analysis assumptions include:

1. Inner thermocouple placement $r_1 = 0.4$ in.
2. Thermocouple placement accuracy was known within ± 0.05 in.
3. Errors in thermocouple temperature measurements were dictated by the hardware and our calibration techniques at $\pm 0.15^\circ\text{C}$.
4. The thermal conductivity of 305-stainless steel (flange heat meter material) was known to within 10 %.
5. Web thickness $h = 0.15$ in.

The results of the uncertainty analysis, shown in Fig. 6, illustrate that increasing thermocouple separation distance reduces the heat output uncertainty bounds, with gradually diminishing benefits seen for larger separation distances. Increased heat through the device also reduces the uncertainty bounds. A thermocouple separation distance of 1.5 in. was selected. Increasing

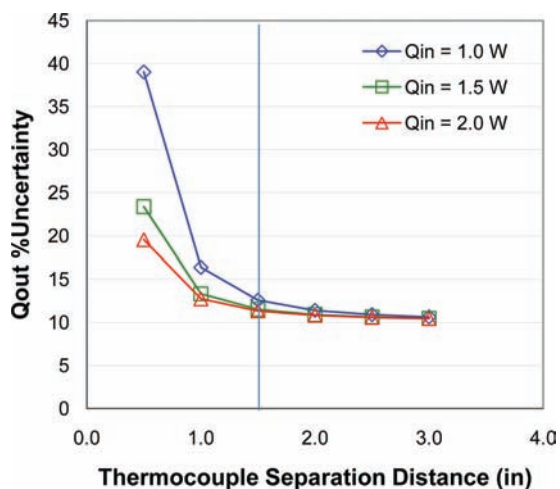


FIG. 6—Influence of thermocouple separation distance on output heat measurement uncertainty.

the radial distance further would increase risk of structural instability while providing insignificant gains in measurement accuracy.

Flange Heat Meter Optimization: Flange Thickness—The webbing thickness influences the temperature error term and, thus, the total heat output uncertainty. An uncertainty analysis was performed using the assumptions below:

Assumptions included:

1. Inner thermocouple placement, $r_1 = 0.4$ in.
2. Thermocouple separation distance $r_2 - r_1 = 1.5$ in.
3. Thermocouple placement accuracy was known within ± 0.05 in.
4. Errors in thermocouple temperature measurements were dictated by the hardware and our calibration techniques at $\pm 0.15^\circ\text{C}$.
5. The thermal conductivity of 305-stainless steel (flange heat meter material) was known to within 10 %.

The results are as shown in Fig. 7.

The figure shows that the uncertainty bounds for the measured heat output decrease with thinner webbings, but depicts gradually diminishing returns as the web thickness decreases. Results were sensitive to the amount of output heat. The webbing thickness was chosen to be 0.15 in. Thinner webbing resulted in machining challenges and did not generate a significantly reduced uncertainty.

Flange Heat Meter Optimization: Summary—There were two dimensions on the flange heat meter that were optimized for thermal measurement accuracy; thermocouple placement and flange thickness. Maximizing thermocouple separation distance and minimizing flange thickness would result in improved

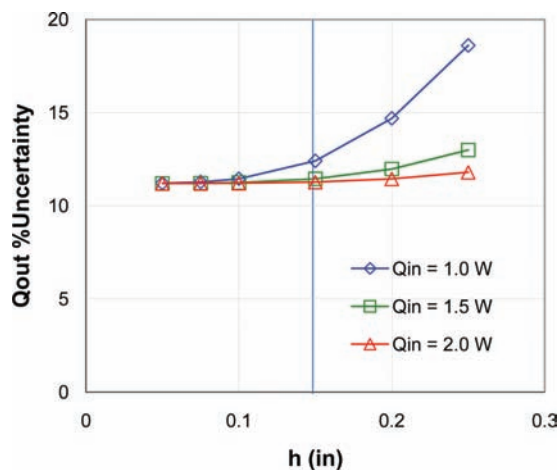


FIG. 7—Influence of flange thickness on output heat measurement uncertainty.

thermal measurement performance, i.e., smaller uncertainty bounds, but both of these dimensional trends also create structural instability and add to cost. As consistent with the shaft heat meter, the relationship between geometry and uncertainty in the flange heat meter is non-linear and diminishing returns are evident. Thus, the dimensions chosen for the flange heat meter design include:

1. Webbing thickness, $h = 0.15$ in.
2. T_c separation distance $r_2 - r_1 = 1.5$ in.

Bearing Thermal Conductance Uncertainty

As was expressed in Eq 1, bearing thermal conductance was based on measurements of heat entering and leaving the bearing as well as the temperatures of the inner and outer races. The uncertainty bounds are based on all of these quantities as well as the errors associated with each of these terms.

Magnitude of Error Terms

The equation to calculate the bearing thermal conductance % uncertainty is restated as follows:

$$\frac{\delta G_b}{G_b} \cdot 100\% = \left[\frac{\delta Q_{in}^2 + \delta Q_{out}^2}{(Q_{in} + Q_{out})^2} + \frac{\delta T_{in}^2 + \delta T_{out}^2}{(T_{in} + T_{out})^2} \right]^{1/2} \cdot 100\% \quad (14)$$

There are two types of terms: heat and temperature terms. The experimentalist is limited on what can be done to optimize the temperature terms, as this is constrained by the operational conditions, T_{in} and T_{out} , and the hardware limitations, δT_{in} and δT_{out} . The heat term, on the other hand, can be optimized. The focus of the optimization effort was to minimize the heat measurement uncertainties, δQ_{in} and δQ_{out} . The optimization of this term was worthwhile in that magnitudes of the heat term dominate the overall bearing thermal conductance term. To demonstrate, the magnitudes and influence of both of these terms will be plotted and compared to see the significance of each. The two terms that will be compared are:

$$Q_{term} = \frac{\delta Q_{in}^2 + \delta Q_{out}^2}{(Q_{in} + Q_{out})^2} \quad (15)$$

$$T_{term} = \frac{\delta T_{in}^2 + \delta T_{out}^2}{(T_{in} + T_{out})^2} \quad (16)$$

The assumptions for the analysis include the following:

1. The flange heat meter inner thermocouple radial location $r_1 = 0.4$ in.
2. The flange heat meter thermocouple separation distance $r_2 - r_1 = 1.5$ in.
3. The flange heat meter web thickness $h = 0.15$ in.
4. Thermocouple placement accuracy was known within ± 0.05 in.
5. Errors in thermocouple temperature measurements were dictated by the hardware and our calibration techniques at $\pm 0.15^\circ\text{C}$.
6. The thermal conductivity of each material was known to within 10 %.
7. The shaft heat meter diameter $d = 0.8$ in.
8. The pyrometer separation distance $L = 1.6$ in.
9. The pyrometer separation distance can be measured to within ± 0.1 in.
10. Pyrometer temperature measurement errors were estimated to be approximately $\pm 1.0^\circ\text{C}$.
11. Static (0 rpm) conditions in the bearing.

Figure 8 plots the magnitude of the Q and T terms as a function of heat input into the system. The graph shows that the magnitude of the heat term vastly dominates the temperature term, indicating that there is additional room for error in the race temperature measurements. This plot also indicates that the best way of minimizing bearing thermal conductance uncertainty bounds is to optimize the design of heat meters, which was done in the previous sections.

Bearing Thermal Conductance Error

Equation 14 showed that operational variables such as the heat flow through the system affects bearing thermal conductance % uncertainty. Figure 9 plots this influence.

The plot shows that an increase in heat input reduces bearing thermal conductance errors with gradually diminishing returns once the heat flow increases

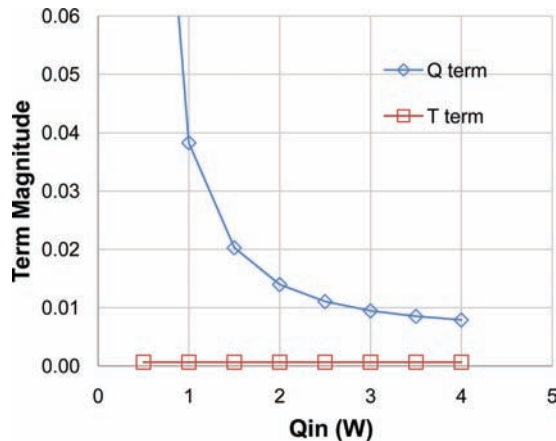


FIG. 8—Influence of input heat on error term magnitude.

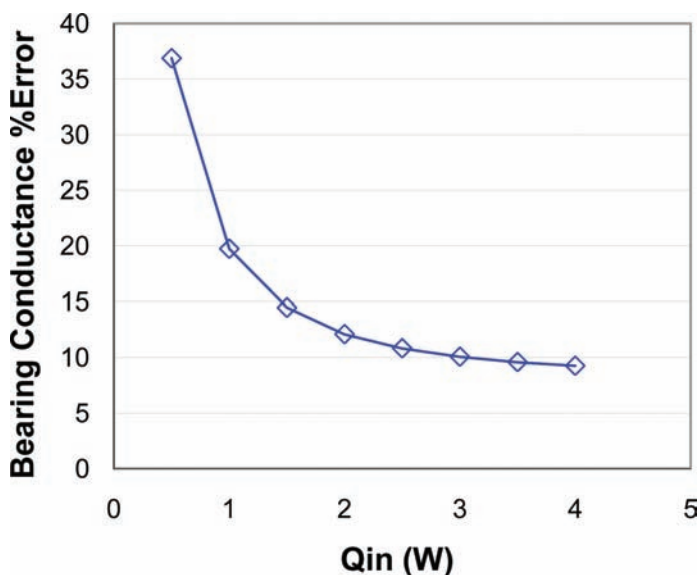


FIG. 9—Influence of input heat on total bearing thermal conductance error.

beyond 2 W. Overall, the bearing thermal conductance error magnitude is acceptable ($<20\%$) for a heat flow as low as ~ 1 W.

Validation of Testing Method

The method chosen to validate the experimental test apparatus was to compare experimental measurements with the Yovanovich bearing model, which is used to calculate thermal conductance across a dry (non-lubricated), static (non-moving) bearing. The next sections will explain the analysis and test method and compare the results.

Analytical Model

The analytical relationship between thermal conductance and axial load in a non-lubricated, non-rotating bearing is well known. Yovanovich [1,2] developed the best-known analytical method, deriving the equation for bearing thermal resistance. A comparison with this model was used to verify the experimental approach before proceeding to the investigation of the more complex bearing cases involving lubrication and motion.

Yovanovich's model assumes the mechanism of heat transfer through a non-lubricated, static bearing is pure conductance, from the hotter race, through the ball, and finally through the cooler race (Fig. 10). The driving sources of thermal resistance across the bearing arise from the thermal contact patches between the ball and the raceways.

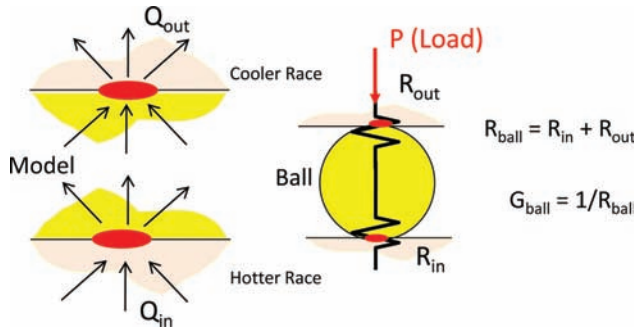


FIG. 10—Heat transfer in the Yovanovich model.

Yovanovich calculated the thermal resistance across a dry, static bearing by modeling the ball and races as semi-infinite half-planes with the elliptical Hertzian contact area serving as the thermal contact region. The basic equations to calculate thermal resistance across each of the ball-to-race contacts are:

$$R_i = \psi_i/4k_i a_i + \psi_i/4k_b a_i \quad \text{inner race to ball thermal resistance} \quad (17)$$

$$R_o = \psi_o/4k_b a_o + \psi_o/4k_o a_o \quad \text{ball to outer race thermal resistance} \quad (18)$$

$$R_b = (R_o + R_i) \quad \text{total thermal resistance across the ball} \quad (19)$$

where Ψ is a non-dimensional geometric factor defined as:

$$\Psi_n = \frac{2}{\pi} \int_0^{\pi/2} \frac{d\theta}{\left(1 - \frac{a_n^2 - b_n^2}{a_n^2} \sin^2 \theta\right)^{1/2}} \quad (20)$$

And n is designated either i or o for the inner or outer race.

The Hertzian contact ellipse and associated major, a , and minor, b , axes can be calculated by a number of existing programs based on classical Hertzian theory, such as BRGS10C. Ball-bearing Hertzian contact stress analysis indicates that the major and minor axes are related to the applied axial load to the $1/3$ power [16]:

$$a_n \propto b_n \propto P^{1/3} \quad (21)$$

Substituting the relation between semi-major and semi-minor axes and axial load (Eq 21) into Eq 20, we see that the non-dimensional geometric factor is

independent of axial load, as the influence of P divides out in the non-dimensional parameter Ψ . Knowing this, and substituting Eqs 20 and 21 into Eqs 17–19, we see that the relation between bearing thermal resistance (Eq 19) and axial load is the inverse of axial load to the 1/3 power.

Once the resistance across each ball is known, the conductance is determined by taking the inverse of the thermal resistance. Scaling this result by the number of balls (N) in the bearing yields the total bearing conductance:

$$G = N \cdot \frac{1}{R_b} \propto P^{1/3} \quad (22)$$

Equation 22 assumes that each ball provides an equivalent thermal pathway as a result of pure axial load, a condition consistent with our experimental setup.

The equation for Hertzian contact area is:

$$A_{\text{Hertzian}} = \pi \cdot \frac{a_n \cdot b_n}{4} \propto P^{2/3} \quad (23)$$

Both the semi-major and semi-minor axes are proportional to axial load to the 1/3 power, as stated earlier. Thus, the total Hertzian contact area is proportional to the applied axial load to the 2/3 power, as Eq 23 indicates.

Because we know that the Hertzian contact area is proportional to axial load to the 2/3 power, and we found that thermal conductance is related to axial load to the 1/3 power, we conclude that thermal conductance of a ball bearing geometry is proportional to the square root of area.

Both the magnitude and trend predicted by the analysis compared well with the experimental measurements, as shown in the following section.

Test Bearings

The test bearings consisted of, first, a standard, off-the-shelf, all-52100 steel ball bearing and, second, a geometrically similar hybrid bearing containing silicon nitride balls and steel races. A summary of the bearings are presented below:

1. Bearing make: barden.
2. Abec rating: abec 7.
3. Bearing size: 305.
4. External geometry.
 - (a) Bore diameter: 25 mm (0.9843 in.).
 - (b) Outer diameter: 62 mm (2.4409 in.).
 - (c) Width: 17 mm (0.6693 in.).
5. Internal geometry.
 - (a) Ball compliment: 10.
 - (b) Ball diameter: 11.0629 mm (15/32 in.).

6. Materials.

- (a) Inner race: 52|100 steel.
- (b) Outer race: 52|100 steel.
- (c) Ball:
 - i 52|100 Steel for the all-steel bearing.
 - ii. Bearing grade silicon nitride for the hybrid bearing.

Comparisons with Experimental Measurements

Two standard off-the-shelf Barden 305-size ABEC 7 bearings as described above were prepared for the series of tests.

Yovanovich models of both bearing types were created and compared with data obtained from the experimental test rig. For the Yovanovich calculations, material properties of both 52|100 steel and bearing grade silicon nitride were needed. Both the mechanical and thermal material properties influence the final thermal resistance across the bearing. The Hertzian contact dimensions a and b are related to material properties (Young's modulus, Poisson's ratio) and the axial load. These values are well known or easily determined. Thermal conductivities of the bearing-grade ball and race materials; on the other hand, are not as well-known a quantity.

Tables 1 and 2 provide material thermal conductivity information that is available in the literature. A wide scatter in thermal conductivities is evident for both bearing grade 52 100 steel (Table 1) and bearing grade silicon nitride (Table 2).

Surprisingly, the variation in measurements of 52 100 bearing steel thermal conductivity data is larger than the variation reported for bearing grade silicon

TABLE 1—52 100 Steel thermal conductivity data.

Material Name	Thermal Conductivity W/m-C (BTU-in/h- ft ² -°F)	References (as of 14 April 2011)
Alloy steel 52100	34.6 (240)	[17]
AISI E 52100 steel	46.6 (323)	[18]
52100 steel	46.6 (323)	[19]
High purity 52100 steel for centrifugal pump bearings	34.6 (240)	[20]
AISI 52100 alloy steel	43.25 (300)	[21]
Bearing steel	30–40 (208–277)	[22]
AISI 52100	43 (298)	[23]
AISI 52100	45 (312)	[24]
52100	47 (326)	[25]
AISI 52100	25 (173)	[26]
Minimum	25 (173)	[26]
Maximum	47 (326)	[25]

TABLE 2—*Bearing grade silicon nitride ceramic thermal conductivity data.*

Material Name	Thermal Conductivity W/m-C (BTU-in/h-ft ² -°F)	References (as of 14 April 2011)
HIP Si ₃ N ₄ ball grade 3	30.7 (213)	[23]
Bearing grade silicon nitride	30 (208)	[24]
Silicon nitride ball	21 (145)	[26]
Silicon nitride NBD200	29 (201)	[26]
Minimum	21 (145)	[26]
Maximum	30.7 (213)	[25]

nitride ceramic thermal conductivity data. This could partially be a result of improvements in accuracy for more recent thermal property experiments, the information for bearing grade silicon nitride being newer. The scatter in data for the silicon nitride conductivity is about 30 %, whereas it is about 50 % for the 52 100 steel.

To compare the Yovanovich model with the experimental results, the lower and upper analytical thermal conductance bounds were calculated based on either 305-size all-52100 steel, or hybrid bearings with silicon nitride balls and then compared with the experimentally derived bearing thermal conductance measurements. The results are shown in Figs. 11 and 12. Figure 11 shows the comparison between an all-52100 steel bearing analysis and experiment, whereas Fig. 12 shows the comparison for the hybrid bearing.

The figures show that the calculated analytical bearing thermal conductance values encompass the measured data, meaning that the measured data were within a reasonable range given the theoretical predictions.

Taking a closer look at the 52 100 steel conductivity data in Table 1, it may be noted that the upper bound information seems to be more consistent

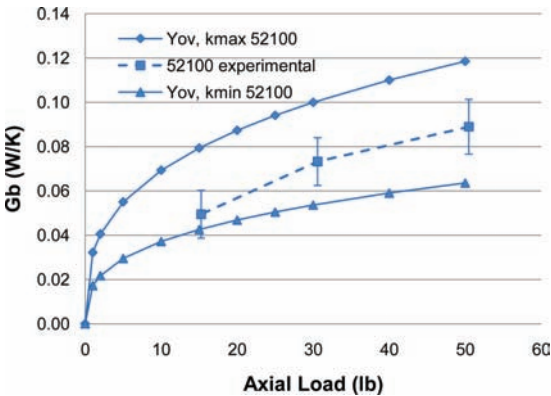


FIG. 11—*Influence of axial load on thermal conductance of a dry static 305-size 52 100 steel bearing.*

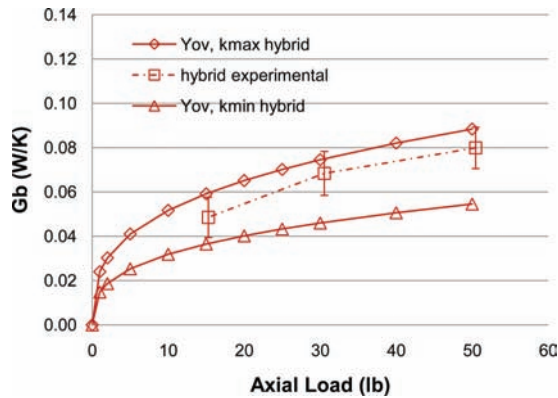


FIG. 12—Influence of axial load on thermal conductance of a dry static 305-size hybrid bearing.

throughout the literature. Taking that consistency into consideration, the upper bound is more likely to be reflective of the actual value. However, because the Yovanovich model used for this analysis assumes perfect contact, the theoretical values most likely overestimate the actual bearing thermal conductance. Thus, the measured values most likely should be lower than the calculated values. This would be true for both the 52 100 all-steel and hybrid bearings.

Summary

Little information is known on bearing thermal conductance. An experimental test rig was designed and constructed for the purpose of making bearing thermal conductance measurements under controlled conditions with the option of performing parametric studies on bearing temperatures, rotational speed, and axial load.

Equations were derived for the calculation of uncertainty for bearing heat input and output, as well as the overall bearing thermal gradient. Analysis indicated that uncertainty in heat measurements was the dominant source of error bounds. Thus, both the shaft and flange heat meters were optimized to obtain reasonable heat measurement error bounds.

The uncertainty analysis suggested an optimal shaft heat meter diameter of 0.8 in., and a minimum pyrometer separation distance of 1.6 in. For the flange heat meter, the webbing thickness of 0.15 in. and a thermocouple separation distance of 1.5 in. were chosen as optimal.

There was some difficulty encountered in comparing experimentally measured and analytically derived bearing thermal conductance values for dry, static bearings because the analytical values were limited to the accuracies of the thermal conductivity information known about the 52 100 and silicon nitride-

bearing materials. However, the test values do lie between the lower and upper limits of the analytical calculations, providing some validation of the experimental setup.

Conclusion

A test rig was developed to measure bearing thermal conductance. An uncertainty analysis was used to optimize the experimental design as well as to determine the dominant source of measurement errors. During the design of the test rig, an uncertainty analysis proved to be a powerful tool, as pilot data showed that the experimentally determined values of bearing thermal conductance for 305-size all-steel and hybrid bearings are well within the bounds predicted by analytical solutions.

Acknowledgments

This work is funded by The Aerospace Corporation's Independent Research and Development (IRAD) Program.

APPENDIX: A: DERIVATION OF THE BEARING THERMAL CONDUCTANCE EQUATION

The equation for thermal conductance across a bearing may be derived from the generalized Fourier heat conduction equation and an energy balance of the heat flow across the bearing, as shown in Fig. 13.

Bulk effective bearing thermal conductance is calculated in this example from the bearing inner race to the outer race; thus neglecting contact resistance to the supporting structures. Heat enters the bearing from the hotter race (in this case the inner race),

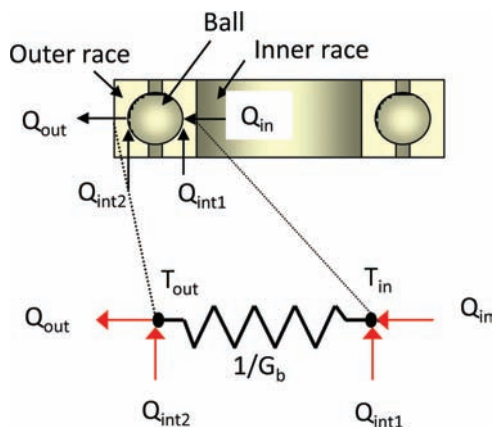


FIG. 13—*Thermal model of a bearing.*

represented as Q_{in} , and exits the bearing at the cooler (outer) race, represented as Q_{out} . Heat is also generated within the bearing and is represented as two variables, Q_{int1} (ball to inner race interface), and Q_{int2} (ball to outer race interface).

There are two basic equations that can be derived from the figure: a heat balance equation, and the bearing conductance equation.

The first relation, a heat balance equation, can be represented as follows:

$$Q_{out} = Q_{in} + Q_{int1} + Q_{int2} \quad (A1)$$

Also, we know that the heat through the bearing is:

$$Q = Q_{in} + Q_{int1} = G_b \cdot (T_{in} - T_{out}) \quad (A2)$$

For this derivation, both Q_{int1} and Q_{int2} are assumed to be equal, although the actual split between the two can be slightly asymmetrical:

$$Q_{int1} = Q_{int2} = Q_{int} \quad (A3)$$

Substituting Eqs. A1 and A3 into Eq. A2 and solving for G_b , we have:

$$G_b = \frac{1}{2} \cdot \frac{Q_{in} + Q_{out}}{T_{in} - T_{out}} \quad (A4)$$

Note that the bearing thermal conductance measurement is proportional to the heat flow through the system and inversely proportional to the temperature drop across the bearing.

References

- [1] Yovanovich, M. M., "Thermal Constriction Resistance between Contacting Metallic Paraboloids: Application to Instrument Bearings," *Proceedings of Heat Transfer and Spacecraft Thermal Control*, MA Institute of Technology, Cambridge, MA, 1971, pp. 337–358.
- [2] Yovanovich, M. M., "Analytical and Experimental Investigation on the Thermal Resistance of Angular Contact Instrument Bearings," *E-2215*, Instrumentation Laboratory, MA Institute of Technology, Cambridge, MA, Dec. 1967.
- [3] Nakajima, K., "Theoretical and Experimental Investigation of the Thermal Contact Resistance of a Space-Use Deep Groove Ball Bearing," *Proceedings of the 18th International Symposium on Space Technology and Science*, AGNE Publishing, Tokyo, Japan, 1992, pp. 879–888.
- [4] Nakajima, K., "Thermal Contact Resistance between Balls and Rings of a Bearing under Axial, Radial, and Combined Loads," *J. Thermophys. Heat Transfer*, Vol. 9, No. 1, 1995, pp. 88–95.

- [5] Delil, A. A. M., Heemskerk, J. F., and Vreeburg, J. P. B., "Design Report on the ESRO Test Rig to Measure the Thermal Conductance and Friction Torque of Rotating Bearings in Vacuum," *Report No. NLR-TR-74069-U*, ESRO CR(P) 05, May 1974.
- [6] Heemskerk, J. F., Delil, A. A. M., and Vreeburg, J. P. B., "A Test Rig to Measure the Thermal Conductance and Friction Torque of Bearings in Vacuum," *Presented at the European Space Tribology Symposium*, Frascati, Italy, 1975, 18 pp.
- [7] Stevens, K. T. and Todd, M. J., "Thermal Conductance across Ball Bearings in Vacuum," *Report No. ESA-ESTL-25*, National Centre of Tribology, Risley, U.K., Feb. 1977, pp. 1–51.
- [8] Rowntree, R. A. and Todd, M. J., "Thermal Conductance and Torque of Thin Section Four-Point Contact Ball Bearings in Vacuum," *Report No. ESA-ESTL-54*, National Centre of Tribology, Risley, U.K., Mar. 1983, pp. 1–57.
- [9] Rowntree, R. A., "Stiffness, Torque, and Thermal Conductance of Thin-Section Four-Point Contact Ball Bearings for Use in Spacecraft Mechanisms," *Proceedings of the 1st European Space Mechanisms & Tribology Symposium*, Noordwijk, Netherlands, 1983, pp. 91–100.
- [10] Takeuchi, Y. R., Dickey, J. T., Demsky, S. M., and Eby, M. A., "Heat Transfer across Ball Bearings in Vacuum," *Proceedings of the 2005 ASME Heat Transfer Conference*, San Francisco, CA, July 2005.
- [11] Demsky, S. M., Takeuchi, Y. R., and Dickey, J. T., "Thermal Conductance Measurement Techniques for High-Speed Bearings in Vacuum," *Proceedings of the 2005 ASME Heat Transfer Conference*, San Francisco, CA, July 2005.
- [12] Takeuchi, Y. R., Eby, M. A., Blake, B. A., Demsky, S. M., and Dickey, J. T., "Influence of Oil Lubrication on Spacecraft Bearing Thermal Conductance," *Proceedings of the 38th Aerospace Mechanism Symposium*, Williamsburg, VA, May 2006, pp. 291–302.
- [13] Davis, S. E., Takeuchi Y. R., and Eby, M. A., "Steel and Hybrid Spacecraft Ball Bearing Thermal Conductance Comparisons," *Presented at the 2011 ASTM International Symposium on Rolling Element Bearings*, Anaheim, CA, April 2011.
- [14] Moffat, R. J., "Using Uncertainty Analysis in the Planning of an Experiment," *J.Fluid Eng.*, Vol. 107, 1985, pp. 173–177.
- [15] Beckwith, T. G., Buck, N. L., and Marangoni, R. D., *Mechanical Measurements*, 3rd Ed., Addison-Wesley, Boston, 1982.
- [16] Harris, T. A., *Rolling Bearing Analysis*, John Wiley and Sons, New York, 1991.
- [17] www.suppliersonline.com (Last accessed April 14, 2011).
- [18] www.matweb.com (Last accessed April 14, 2011).

- [19] Harris, S. J., Krauss, G., Siniawski, M. T., Wang, Q., Shuangbiao, L., and Ao, Y., "Surface Feature Variations Observed in 52100 Steel Sliding Against a Thin Boron Carbide Coating," *Wear*, Vol. 249, No. 10, Nov. 2001, pp. 1004–1013.
- [20] www.pump-zone.com (Last accessed April 14, 2011).
- [21] Mudhivarthi, S., "Dry Sliding Tribological Characteristics of Hard, Flat Materials With Low Surface Roughness," M.S. Thesis, University of South Florida, Sept 2003.
- [22] www.ahrinternational.com (Last accessed April 14, 2011).
- [23] Paleu, V., Cretu, S., Dragan, B., and Balan, R., "Test Rig for Friction Torque Measurement in Rolling Bearings," *Machine Design*, 2004, pp. 85–91.
- [24] SKF Material Data Sheet, www.skf.com/ (Last accessed April 14, 2011).
- [25] Enceratec Material Data Sheet, <http://www.enceratec.com/> (Last accessed April 14, 2011).
- [26] Conley, P., *Space Vehicle Mechanisms*, Wiley, John & Sons, Incorporated, Feb. 1998.

Yoshimi R. Takeuchi,¹ Sean E. Davis,¹
and Matthew A. Eby²

Steel and Hybrid Spacecraft Ball-Bearing Thermal Conductance Comparisons

REFERENCE: Takeuchi, Yoshimi R., Davis, Sean E., and Eby, Matthew A., "Steel and Hybrid Spacecraft Ball-Bearing Thermal Conductance Comparisons," *Rolling Element Bearings* on April 13–15, 2011 in Anaheim, CA; STP 1542, Yoshimi R. Takeuchi and William F. Mandler, Editors, pp. 118–142, doi:10.1520/STP104234, ASTM International, West Conshohocken, PA 2012.

ABSTRACT: In vacuum, bearing temperatures are primarily governed by two properties: bearing thermal conductance, which governs the overall heat transfer through the bearing, and heat generation, which is caused by the frictional mechanical losses during bearing motion. Both quantities need to be known for accurate bearing thermal predictions in a vacuum environment. The heat generation term, which is related to torque multiplied by rotational speed, is the better known of the two quantities. Information on bearing thermal conductance; on the other hand, tends to be more poorly quantified, and thus will be the focus of this paper. A comparison will be made between the thermal conductance of all-steel and hybrid bearings containing silicon nitride ceramic balls inside steel races. Thermal conductance is a function of various operational conditions, such as rotational speed, temperature, mechanical loading and presence of lubricant. The exact effects the aforementioned variables have on bearing thermal conductance will be explored in more detail for both bearing types. Results indicate that, for a given bearing size and geometry, bearing thermal conductance is most strongly influenced by the presence and amount of liquid lubricant and state of motion. Both all-steel and hybrid bearings react to these variables in a comparable manner.

KEYWORDS: bearing thermal conductance, thermal resistance, experiment, hybrid bearings, steel bearings

Manuscript received July 25, 2011; accepted for publication March 20, 2012; published online August 2012.

¹Ph.D., Engineering and Technology Group, The Aerospace Corporation, 2310 E. El Segundo Blvd., El Segundo, CA 90245.

²Engineering and Technology Group, The Aerospace Corporation, 2310 E. El Segundo Blvd., El Segundo, CA 90245.

Copyright © 2012 by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

NOMENCLATURE

- A_{Hertzian} = Hertzian contact area between the ball and race
 a = major axis of the Hertzian contact ellipse
 b = minor axis of the Hertzian contact ellipse
 G_b = bulk thermal conductance across the test bearing
 P = axial loading
 Q_{in} = input heat entering the test bearing inner race
 Q_{out} = output heat exiting the test bearing outer race
 T_{ave} = average bulk bearing temperature
 T_{in} = measured test bearing inner race temperature
 T_{out} = measured test bearing outer race temperature
 ΔT = temperature difference between the bearing inner and outer races

Introduction

Accurate thermal predictions for a bearing in a vacuum environment call for estimation of two independent properties: bearing thermal conductance and heat generation. Heat generation, which is related to torque multiplied by speed, tends to be the better known quantity. Models and software programs have been designed to calculate and quantify torque, and speed tends to be a known for a given application. Thus, the focus of this study was to experimentally measure and quantify thermal conductance, the lesser known bearing variable, and in the process understand the influence of operational conditions such as axial load, temperature, effect of motion, presence of lubricant and materials.

A literature review reveals little information on bearing thermal conductance [1–19], and less is known about the influence of materials, lubricant, and influence of operational variables such as axial load, temperature, and state of motion.

Analytically, Yovanovich [1,2] and Nakajima [3,4] developed a closed form solution to calculate thermal conductance across a dry (non-lubricated), static (non-rotating) bearing. Experimental studies have been performed to validate these models [1–4].

For oil lubricated bearings under dynamic conditions, some experimental work has been performed. Stevens and Todd [5] measured thermal conductance of bearings typically for low speeds, such as a few hundred rpm, but in one case, up to speeds of 2500 rpm using a test rig developed by Delil et al. [6,7]. Additional studies involving thin cross-section bearings were also executed at their laboratory over the years [8,9].

As spacecraft bearings have been used at higher speeds, operational demands on bearings have grown and thermal concerns have similarly increased. A typical spacecraft control moment gyroscope or reaction wheel may operate at speeds of around 6000 rpm. The work conducted at The Aerospace Corporation by Takeuchi et al. [10–12] has addressed these higher speeds. Some bearing studies have been conducted under cryogenic temperatures by other researchers [13,14].

Other types of bearings have been studied as well, including tapered roller bearings [15], journal bearings [16], and radial bearings [17]. This paper will be focused on ball bearings for space or other vacuum applications.

The sporadic data in literature for various bearing sizes and geometries renders comparisons of what little information is known on bearing thermal conductance difficult. This paper addresses this concern by comparing the thermal conductance of two geometrically similar, 305-size hybrid and all-steel bearings; thus providing the means to investigate the influence of ball-bearing material. A comparison of hybrid bearings with silicon nitride balls and 52 100 steel races to its all-52 100-steel-bearing counterpart (assuming similar internal and external geometries) allows for an understanding of material influences. The effects of various operational conditions are studied, including axial load, temperature, and speed. Furthermore, the influences of liquid lubricant and run-in conditions are considered.

Experimental Hardware

The objective of the experiment is to measure the bulk effective conductance across a bearing, from the inner surface of the inner race to the outer surface of the outer race. Experiments are conducted while controlling variables such as speed, axial load, temperature, ball material, and the presence of lubricant for the purpose of allowing the researcher the capability of performing parametric studies.

A schematic diagram of the test fixture is shown in Fig. 1. The design of the test fixture is explained in more detail in a companion paper [20]. The test is conducted in a vacuum environment while the basic measurements needed to determine the bearing thermal conductance are taken while thermal and mechanical constraints are tightly controlled.

Temperature is controlled using a heat lamp and cooling channel as a thermal source/sink combination. The thermal components are shown in greater detail in Fig. 2. The heat lamp warms the shaft; this heat is conducted through the test bearing from the inner race to the outer race and then sinks radially through the Flange Heat Meter (based on the Heat Flux Meter as designed by Delil [6,7]) to the cooling channel that extends around the outer diameter of the meter.

The bearing thermal conductance, G_b , is calculated from experimental measurements using the following equation

$$G_b = \frac{1}{2} \cdot \frac{Q_{in} + Q_{out}}{T_{in} - T_{out}} \quad (1)$$

where T_{in} and T_{out} are the temperature of the inner and outer races, respectively and Q_{in} and Q_{out} the heat entering and exiting the bearing, respectively. The

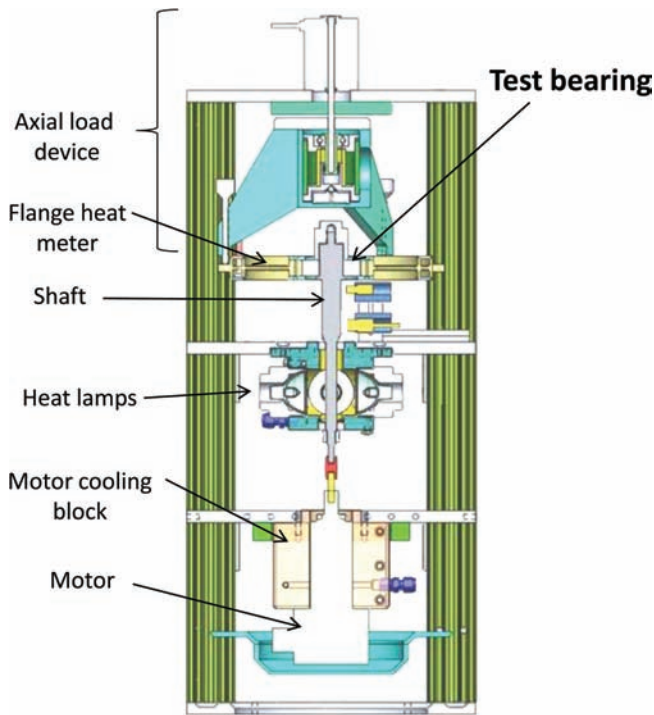


FIG. 1—Schematic diagram of the bearing thermal test fixture.

heat input to the test bearing, Q_{in} , is determined by taking two temperature measurements along the length of the shaft below the test bearing. The heat output from the test bearing, Q_{out} , is determined by taking two temperature measurements radially along the flange of the heat meter. Temperatures of the

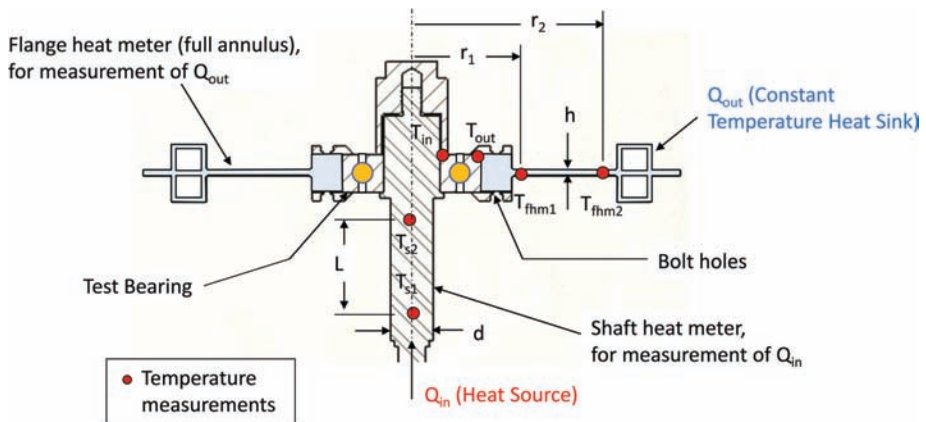


FIG. 2—Schematic diagram of the thermal system in the bearing thermal test fixture.

rotational components are measured using infrared pyrometers, while temperatures of stationary components are measured with type E thermocouples.

Mechanical controls include rotational speed and axial loading. The speed of the test bearing inner race is driven by a motor and spindle assembly, whereas the outer race is held stationary. The axial load on the test bearing is applied using a stepper motor loading device that applies a downward force to the outside rim of the flange heat meter; this ultimately places an axial load on the outer race of the test bearing, whereas the inner race opposes this force via a shoulder on the shaft.

The experiment was designed to minimize the measured errors using an uncertainty analysis as described in [20]. The analysis was also used to estimate the magnitude of uncertainty in the experimentally measured bearing thermal conductance.

Parametric Studies Overview

Parametric studies were used to determine how different operational conditions influence bearing conductance. Impact of operational conditions, such as temperature and preload, were studied by performing parametric tests on each of these variables. Material effects were studied by comparing the difference in thermal conductance between the 305-size all-steel bearing with hybrid bearings of similar geometry but containing silicon nitride balls. Experimental measurements determined if differences in conductance existed at all. Furthermore, experiments conducted on both static and dynamic bearings were designed to quantify the effect of motion on thermal conductance.

Static tests were performed on hybrid and all-steel bearings under both dry (non-lubricated) and oiled conditions to examine the influence of lubricant presence. Temperature and axial load studies were conducted in each case.

Dynamic (oil lubricated and running at a constant inner race speed) test cases were compared with the static cases; influence of temperature and axial load variations were studied for the dynamic cases as well.

In each case, a test matrix was necessary to separate out the influence of individual variables, such as axial load, speed, or temperature. The test results presented in this paper were based on a matrix reduction for the separation of each variable.

Test Bearing Details

The hybrid and all-steel bearings were geometrically similar with designed differences intentionally limited to the ball material. A summary of the all-steel and hybrid bearing material composition and geometries are shown in Table 1. As shown in that table, the sole difference between the hybrid and steel

TABLE 1—*Test bearing materials and geometry summary.*

	Steel bearing	Hybrid bearing
Bearing size	305-size	305-size
Bearing make	Barden	Barden
ABEC	7	7
Race material	52 100 steel	52 100 steel
Ball material	52 100 steel	Bearing grade silicon nitride
Ball diameter	0.46875 in. (11.9 mm)	0.46875 in. (11.9 mm)
Number of balls	10	10
Bore diameter	0.9843 in. (25 mm)	0.9843 in. (25 mm)
Outer diameter	2.441 in. (62 mm)	2.441 in. (62 mm)
Width	0.6693 in. (17 mm)	0.6693 in. (17 mm)

bearings is that the ball material for the hybrid bearing is made of bearing grade silicon nitride instead of the 52 100 steel. The race material in each bearing type, however, was 52 100 steel. Bearing grade silicon nitride is a highly dense form of ceramic well-suited for the long operational life expected for bearing applications.

The oil used for all tests requiring lubricated bearings was Nye Synthetic Oil 2001.

Static and Dry Bearings

The influences of average temperature and temperature difference on bearing thermal conductance were explored. Average temperature is defined as the bulk average temperature of both bearing races

$$T_{\text{ave}} = \frac{(T_{\text{in}} + T_{\text{out}})}{2} \quad (2)$$

Whereas temperature difference is defined as

$$\Delta T = (T_{\text{in}} - T_{\text{out}}) \quad (3)$$

Note that the average temperature through the bearing may remain constant, whereas the temperature difference across the bearing changes, and vice-versa. Adjustment of the heat lamp voltage and cooling channel temperature of the experimental test rig yielded the necessary thermal control. A test matrix was conducted for a series of heat lamp setting and heat sink temperatures. Then, the test matrix reduction was ultimately used to separate the influence of each variable.

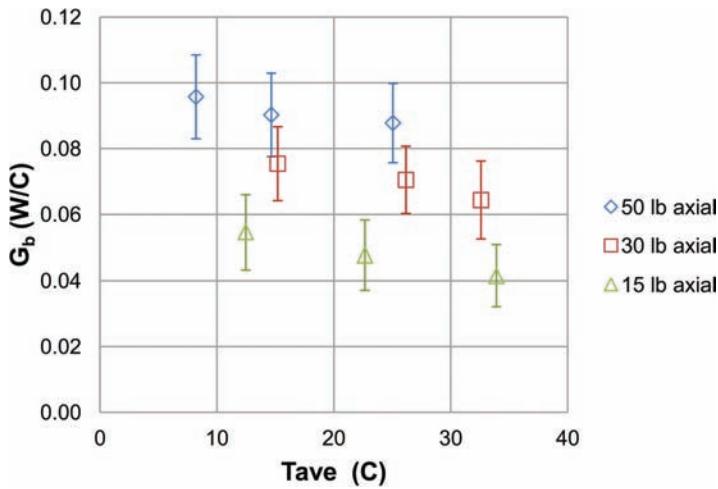


FIG. 3—Static dry steel bearing, G_b versus T_{ave} , for $\Delta T = 20^\circ\text{C}$.

Figures 3 and 4 illustrate the relation between bearing thermal conductance and average temperature (T_{ave}) for all-steel and hybrid bearings, respectively. Test data were taken under three different axial loads and the bearing race temperature difference (ΔT) was 20 C for all tests considered here. The uncertainty bounds on the experimentally measured thermal conductance data represent a 95 % confidence interval as derived from [20]. For both bearing types, the figures show that the average temperature does not exert a substantial influence on bearing thermal conductance; whereas a slightly

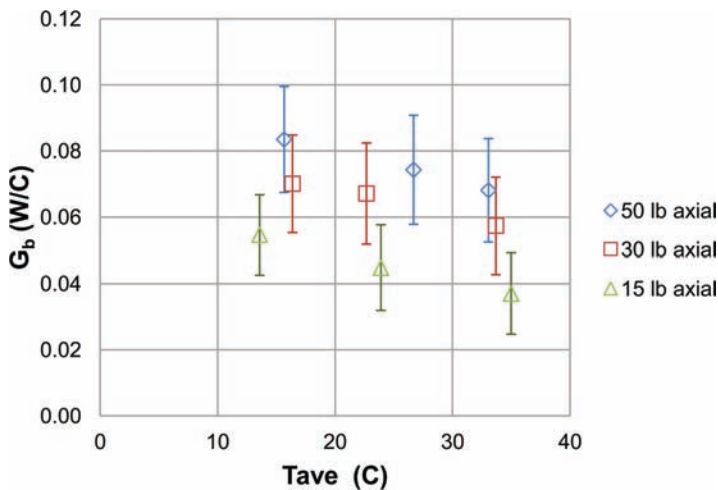


FIG. 4—Static dry hybrid bearing, G_b versus T_{ave} , for $\Delta T = 20^\circ\text{C}$.

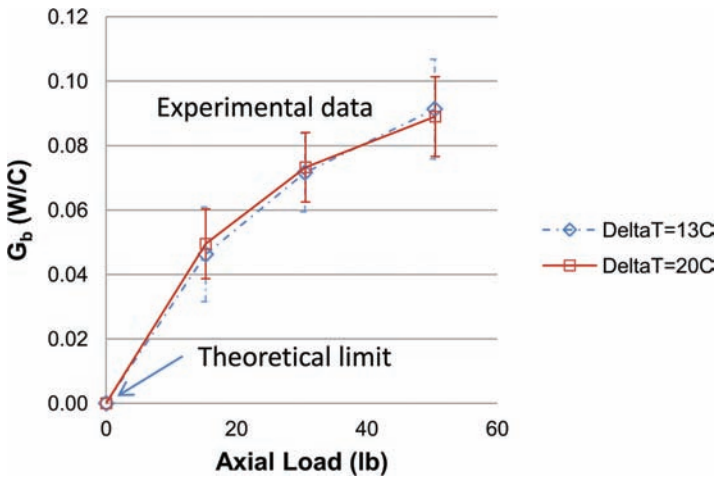


FIG. 5—Static dry steel bearing, G_b versus axial load, for $T_{ave} = 20^\circ\text{C}$.

negative trend may be discerned by eye, the uncertainty bounds encompass the trend, and as such, there remains doubt as to whether average temperature affects bearing conductance. Axial loading; on the other hand, had a clearly substantial impact on bearing thermal conductance. Thermal conductance increased approximately 30 %–50 % with a doubling of the axial load over the test range.

Figures 5 and 6 illustrate the relation between bearing thermal conductance and axial load for a constant average temperature (T_{ave}) of 20°C . The plot

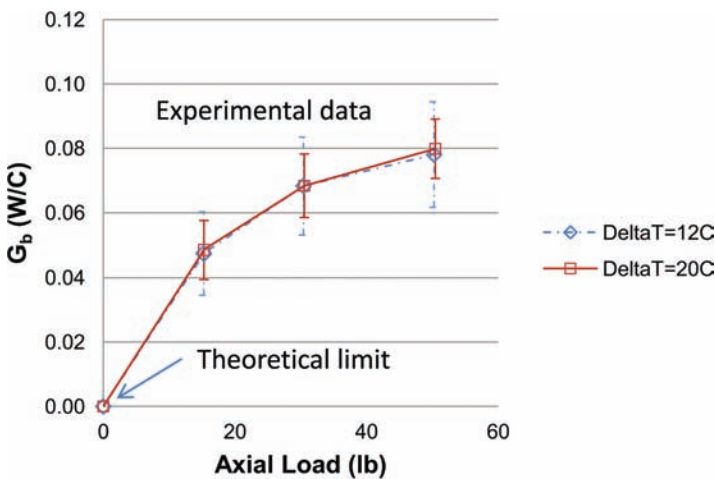


FIG. 6—Static dry hybrid bearing, G_b versus axial load, for $T_{ave} = 20^\circ\text{C}$.

shows curves for temperature differences (ΔT) of 13°C and 20°C. Figure 5 corresponds to the all-steel bearing case whereas Fig. 6 represents the hybrid bearing case. In both, it is shown that the bearing thermal conductance is independent of thermal difference, ΔT , with the conductance versus axial load curves for both temperature differences falling into close agreement over the entire test range considered. However, bearing thermal conductance did show sensitivity to axial load in a monotonically-increasing and non-linear manner for both the dry all-steel and hybrid bearings.

The experimental conductance measurements performed here were compared with the predictions of a well known analytical model to understand the trends observed in the experiments. Yovanovich [1,2], who, as previously mentioned, derived an analytical model to calculate bearing thermal conductance, studied the effects of loading on bearing thermal conductance for dry, static conditions. He determined the relation between axial load and bearing thermal conductance to be a 1/3 power law

$$G_b \propto P^{1/3} \quad (4)$$

The relation between axial load and Hertzian contact area is a 2/3 power law

$$A_{\text{Hertzian}} \propto P^{2/3} \quad (5)$$

Together, these results lead to the conclusion that the bearing thermal conductance is related to the square root of area.

The Hertzian contact area in turn is related the mechanical material properties such as Young's modulus and Poisson's ratio. The Young's modulus of bearing grade silicon nitride is higher than that of 52 100 steel, leading to a smaller Hertzian contact area for a proscribed axial load. Smaller Hertzian contact areas thus lead to higher thermal resistances, or a slightly lower bearing thermal conductance.

Figure 7 shows a comparison of measured thermal conductance for a hybrid and all-steel bearing as a function of axial loading for a constant average temperature and thermal gradient. The figure indicates that the measured hybrid bearing thermal conductance is slightly less than that of the all-steel bearing at higher loads (approximately 10 % less for a 50-lb axial load). Such behavior is expected, because a higher Young's modulus for the ceramic ball should yield smaller contact areas between the ball and raceway. However, the bearing thermal conductance difference between a dry hybrid and all-steel bearing is minimal. Whereas the hybrid data is consistently slightly lower than the all-steel bearing, it should be pointed out that the measurements fall within the experimental uncertainty bounds.

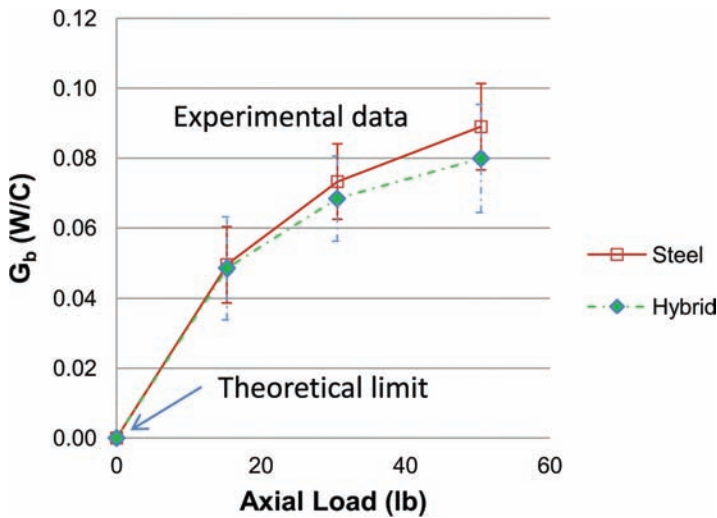


FIG. 7—Static dry steel and hybrid bearing, G_b versus axial load, for $T_{ave} = 20^\circ\text{C}$, $\Delta T = 20^\circ\text{C}$.

Static and Oil Lubricated Bearings

The next sections investigate the sensitivity of bearing thermal conductance to the presence and quantity of oil lubrication. The run-in process, i.e., the elapsed time and maximum speed the bearing reaches once it is in operation, determines the quantity of oil left inside the bearing mechanism. To gauge the impact of lubricant loss, thermal conductance was measured before the test bearing was run-in and after it was completely run-in.

Bearing Run-in

A “run-in” bearing, as used in this paper, indicates that the bearing has been operated at a constant speed for a sufficiently long period of time such that changes in measured torque are no longer observed when operated at a constant motor speed. Typically, the bearing initially exhibits a higher torque and heat generation for a given motor speed because of the uneven distribution of excess lubricant in the bearing, which causes resistance to ball motion. As the bearing is run-in, the torque and heat generation drop because of the lubricant loss both from the centrifugal forces, as well as oil being pushed out from the ball to race pathway because of ball motion, eventually reaching a minimum value. The exact length of time required to reach a fully run-in bearing depends on the initial lubricant quantity, maximum speeds achieved during run-in, and bearing size.

In the tests described here, thermal conductance measurements were taken (1) before any run-in was conducted on the bearings, and (2) after the bearings were completely run-in at 6000 rpm. The bearings were considered completely run-in once a minimum value of heat generation was observed and no further changes were detected. The total run-in time took a minimum of 16 h, and no torque changes were detected after 12 h of run-in.

Static and Oil Lubricated Bearings: Before Run-in

A number of 305-sized all-steel and hybrid bearings were lubricated with Nye Synthetic Oil 2001 and weighed. One all-steel bearing and one hybrid bearing were chosen for testing based on having the same initial lubricant quantity of 130 mg. Bearing thermal conductance was measured prior to performing any run-in and then comparisons were made to the dry bearing case.

Figures 8 and 9 illustrate the test results for a static bearing lubricated with 130 mg of oil for an all-steel and hybrid bearing, respectively. The figures show a lack of influence of average temperature on bearing thermal conductance for three different axial loads. All data in these plots were obtained for a constant bearing race temperature difference of 7 °C. The plots indicate that there is no substantial impact of average temperature on the thermal conductance of a static, oil lubricated bearing.

Figure 10 compares the influence of axial load on thermal conductance for hybrid and all-steel oil lubricated bearings. All the data in the graph represent constant thermal constraints, with an average temperature of 20°C and a

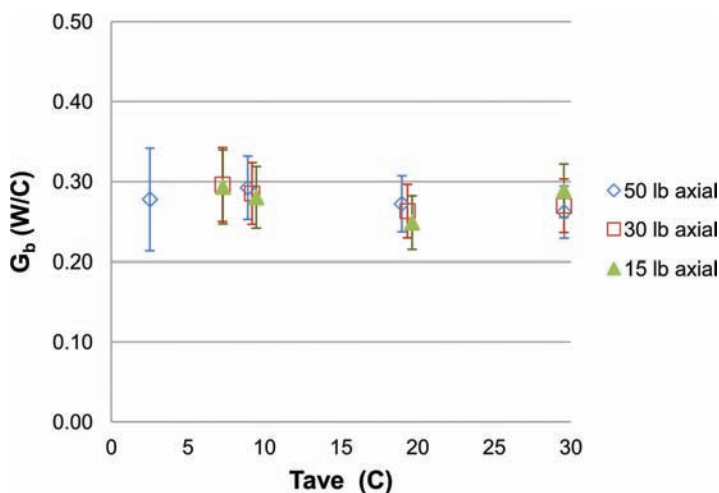


FIG. 8—Static non-run-in oil lubricated steel bearing, G_b versus T_{ave} , for $\Delta T = 7^\circ\text{C}$.

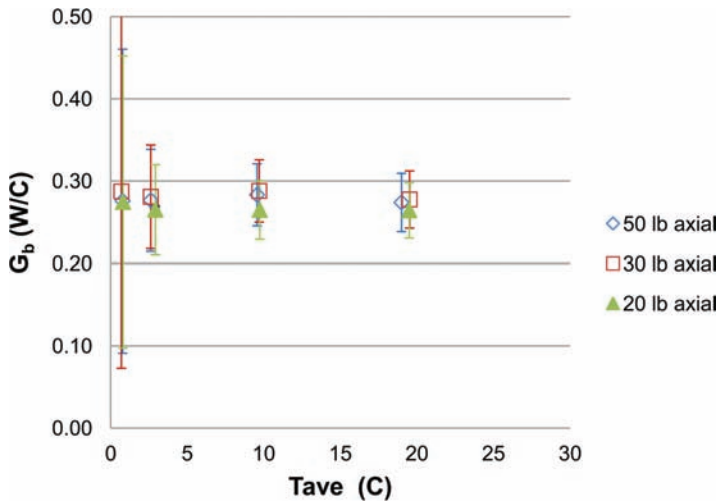


FIG. 9—Static non-run-in oil lubricated hybrid bearing, G_b versus T_{ave} , for $\Delta T = 7^\circ\text{C}$.

thermal difference of 7°C and total oil quantity of 130 mg. The results indicate that there is little effect of axial load or material composition on the oiled-bearing thermal conductance. Material composition had minimal-to-no impact on thermal conductance in oil lubricated bearings. The presence of lubricant also appears to negate any effect of axial load on bearing conductance, which is in sharp contrast with the data shown in Fig. 7 for dry bearings.

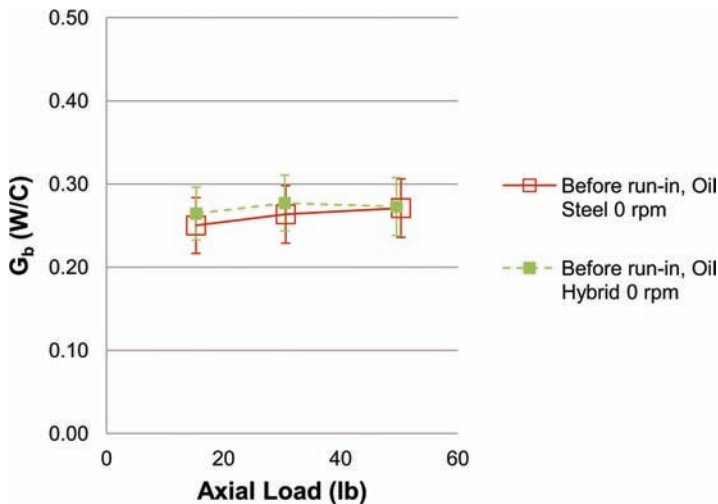


FIG. 10—Static non-run-in oil lubricated steel and hybrid bearing, G_b versus axial load, for $T_{ave} = 20^\circ\text{C}$ and $\Delta T = 7^\circ\text{C}$.

Figure 11 contrasts the effects of axial load on bearing thermal conductance for a dry and oil lubricated static bearing. Hybrid and all-steel bearing data are both shown. The plot shows that the dry hybrid and all-steel bearings are sensitive to axial load in a monotonic and non-linear manner, as predicted previously in Yovanovich's analytical results. This contrasts with the oil lubricated bearings, which are seemingly not substantially influenced by axial load to any measurable degree. Furthermore, the bearing thermal conductance values are approximately 3–5 times larger for the oil lubricated bearing as compared to the dry bearing. This suggests that the presence of lubricant dominates the heat transfer in the static bearing. The chosen ball materials do not appear to influence this trend, as a comparison of hybrid and all-steel bearing thermal conductance values fall well within error bounds.

Figure 12 depicts a possible explanation for these effects. In a static bearing, the mechanism of heat transfer across the system is conduction through the hotter race, across the hotter race-ball contact area, through the ball, across the cooler race-ball contact area, and finally through the cooler race. The thermal resistance across the bearing is dominated by the constrained contact areas. In a dry bearing, this contact patch consists solely of the metal-to-metal Hertzian

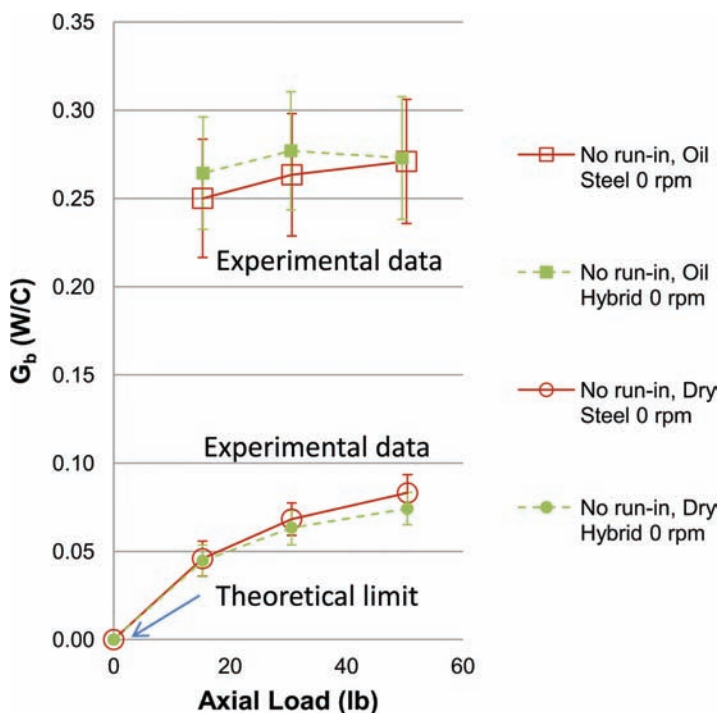


FIG. 11—Comparison between static dry and oil lubricated steel and hybrid bearings, G_b versus axial load, for $T_{ave} = 20^\circ\text{C}$.

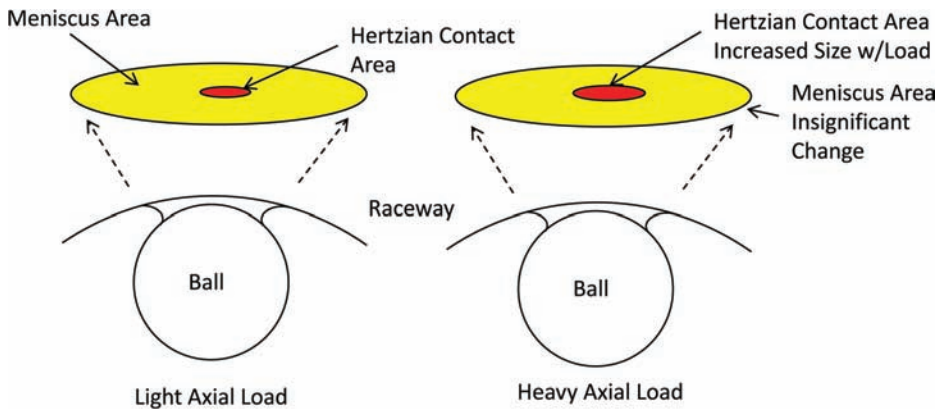


FIG. 12—Impact of oil lubrication on a static bearing – thermal conductance theory.

contact area, which is strongly influenced by the applied axial load. As explained earlier, bearing thermal conductance is proportional to square root of area. This means the increase in conductance diminishes as the contact area increases in size. Because load is proportional to area to the $2/3$ power, and thermal conductance is proportional to square root of area, the impact of axial load on bearing thermal conductance decreases as area increases. For a dry bearing, this area is very small, so small changes in applied load result in relatively large changes in bearing thermal conductance.

In a lubricated bearing, the ball-to-raceway contact region consists of both the metal-to-metal Hertzian contact area as well as the lubricant meniscus between the ball and raceway. This increased thermal pathway substantially magnifies heat transfer across the bearing; thus, resulting in a higher bearing thermal conductance. The large difference between the thermal conductances of dry and oil lubricated bearings suggest that the oil meniscus wetted area vastly dominates over the metal-to-metal Hertzian contact area. Thus, the lubricant provides the major thermal pathway for heat transfer between the ball and raceways.

As an example, the effective contact areas were determined based on the measured bearing thermal conductance values for a bearing with the following characteristics:

- (1) bearing size: 305-size,
- (2) bearing materials: all-52 100 steel bearing,
- (3) lubrication: Nye Synthetic Oil 2001,
- (4) lubricant condition: non-run-in, tested as received with 130 mg oil by weight, and
- (5) axial load: 50 lb axial load.

The calculated areas for the ball and inner race contact areas were based on the contact geometry required to generate the measured thermal conductance, using the Yovanovich bearing thermal model:

- (1) total effective area (metal-to-metal contact area and lubricant): 0.0017 in.²,
- (2) metal-to-metal Hertzian contact area: 0.0002 in.², and
- (3) effective oil lubrication contact area: 0.0015 in.².

As can be seen by the analysis results above, the oil contact area is an order of magnitude larger than the metal-to-metal Hertzian contact area. The sizable difference in contact area between non-run-in dry and oiled bearings suggests that the lubricant meniscus dominates and dictates the effective thermal resistance across the ball to race contact area.

The lack of responsiveness to axial load could also be explained in part by the sheer size of the lubricant meniscus as compared to the metal-to-metal contact patch. As the bearing thermal conductance is related to the square root of area, and a diminishing sensitivity in conductance occurs with larger areas, there is a lack of sensitivity to load. Theoretically, there should be some response to axial loads with the presence of lubricant, but the change would be outside the measurement sensitivity (within uncertainty bounds) for a static oil lubricated bearing.

Static and Oil Lubricated Bearings: After Run-in

Both the all-steel and hybrid bearings were run-in at 6000 rpm for a minimum of 16 h, then each were weighed and tested again at static conditions. At the beginning of the run-in period, bearing thermal properties (conductance and heat generation) change as lubricant is lost, but by the end of the run-in period, bearing thermal conductance and heat generation reached a minimum and no additional changes were observed. Thus, further losses were deemed to have ended.

The run-in process resulted in a loss of 30 mg of lubricant from the all-steel bearing (a total of 100 mg remained), and 40 mg of lubricant loss from the hybrid bearing (a total of 90 mg remained). The accuracy of the mass measurements was ± 10 mg. The amount of lubricant lost during run-in is most likely because of centrifugal forces ejecting excess lubricant from the bearing. Undistributed lubricant is also pushed out from the ball to race pathway during bearing motion.

After the run-in process was completed, thermal conductance was measured for static (0 rpm) all-steel and hybrid bearings over a range of temperatures and axial loads. Results of these measurements are shown in Figs. 13 and 14. These figures illustrate the effect of average temperature on bearing thermal conductance for several different axial loads applied to the all-steel and hybrid

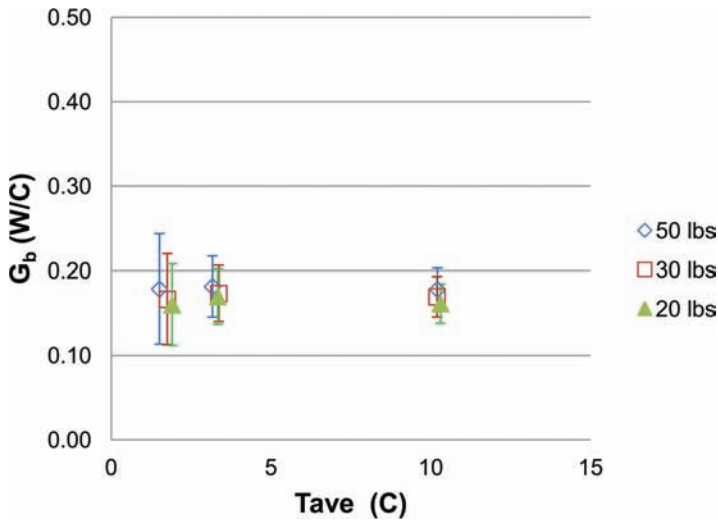


FIG. 13—Static run-in oil lubricated steel bearing, G_b versus T_{ave} , for $\Delta T = 10^\circ\text{C}$.

bearings, respectively. The results show that, again, the oil lubricated bearing is not sensitive to changes in temperature or axial loading.

Figure 15 compares the effect of axial load on bearing thermal conductance for the all-steel and hybrid bearings before and after run-in. The difference in thermal conductance magnitude is evident after the bearing has been fully run-in. The magnitude of the bearing thermal conductance is substantially lower

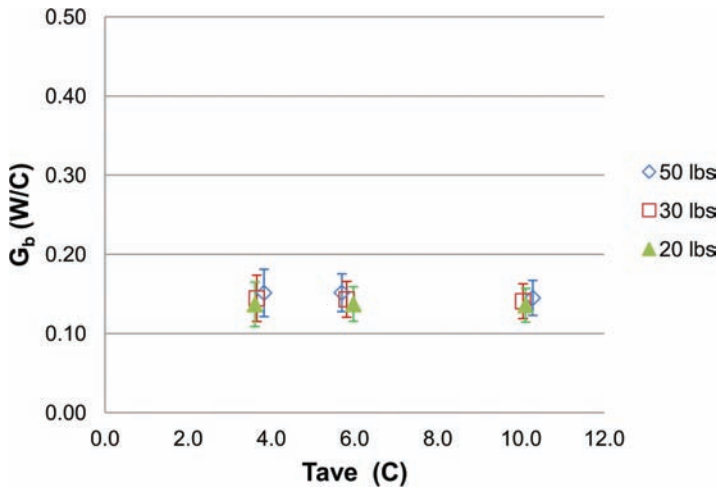


FIG. 14—Static run-in oil lubricated hybrid bearing, G_b versus T_{ave} , for $\Delta T = 10^\circ\text{C}$

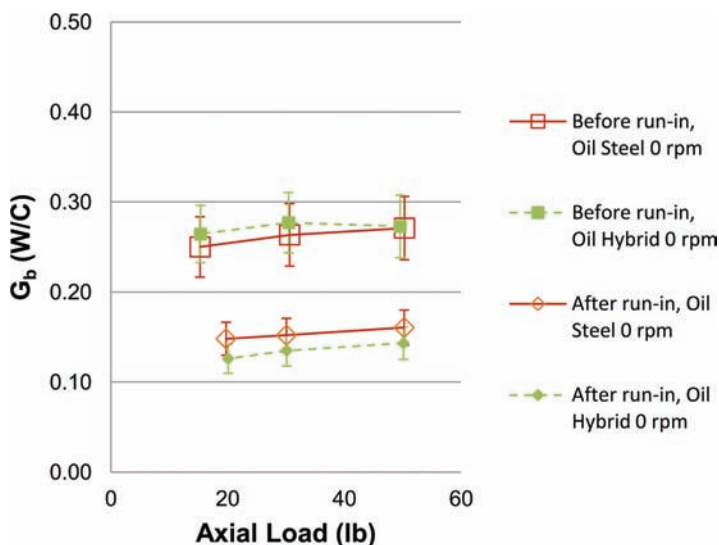


FIG. 15—Comparison between non-run-in and run-in oil lubricated, steel and hybrid bearings, G_b versus axial load, for $T_{ave} = 20^\circ\text{C}$.

(~40 %–50 % decrease) after the run-in period, which is as expected given the lower lubricant quantity. The thermal conductance of the run-in, oil lubricated bearing is still substantially higher (~2–3 times) than the conductance of the dry bearing.

The reason for a substantially higher thermal conductance across a run-in, oil lubricated bearing over its dry counterpart could be explained by the oil meniscus area, as illustrated by the following analysis. The effective contact areas were determined based on the measured bearing thermal conductance values for a bearing with the following characteristics:

- (1) bearing size: 305-size,
- (2) bearing materials: all-52 100 steel bearing,
- (3) lubrication: Nye Synthetic Oil 2001,
- (4) lubricant condition: run-in, 100 mg oil by weight, and
- (5) axial load: 50 lb axial load.

The calculated areas for the ball and inner race contact areas were based on the contact geometry required to generate the measured thermal conductance, using the Yovanovich bearing thermal model:

- (1) total effective area (metal-to-metal contact area and lubricant): 0.0006 in.^2 ,
- (2) metal-to-metal Hertzian contact area: 0.0002 in.^2 , and
- (3) effective oil lubrication contact area: 0.0004 in.^2 .

Whereas there was a dramatic reduction in lubricant meniscus area between the ball and raceway after run-in (from 0.0015 in.² for the non-run-in case to 0.0004 in.² after run-in), the oil contact geometry was still 100 % larger than the metal-to-metal Hertzian area. The oil contact area still dominates the heat-transfer pathway after run-in, and the ball-to-race thermal contact resistance is governed by the presence of the oil.

Changes in axial loading do not affect the conductance of the run-in bearings, which again could be because of diminishing sensitivity with larger contact areas as compared with the dry case. Thermal conductance for a static bearing, being proportional to the square root of area, responds less as the area increases in size. Theoretically, there should be some response to axial loads with the presence of lubricant, but the change would be outside the measurement sensitivity (within uncertainty bounds) for a static run-in oil lubricated bearing.

The impact of ball material composition on thermal conductance is not significant for a static bearing made of the materials considered. Instead, the driver amongst the parameters investigated is the lubricant quantity.

Dynamic and Oil Lubricated Bearings: Test Results

Two 305-size bearings, one all-52 100 steel and the other hybrid, were tested dynamically at a constant 6000 rpm inner race speed. The outer race was held stationary by the Flange Heat Meter. Before undergoing testing under dynamic conditions, the bearing was fully run-in as described in the previous section.

In Figs. 16 and 17, the two plots depict the influence of axial loading on thermal conductance for two different average temperatures but constant thermal difference. Figure 16 shows the all-steel bearing test results, whereas Fig. 17 shows the hybrid bearing test results. In the plots, it can be seen that the thermal conductance versus axial load curves fall on top of each other for both average bearing temperatures, indicating that the average bearing temperature did not have a substantial impact on thermal conductance over the test range specified for either bearing type.

The results for axial load; on the other hand, indicated that the loading influenced thermal conductance for a dynamic oil lubricated bearing. The plots show that thermal conductance trended higher as load increased. By contrast, the same oil lubricated bearing under static conditions, both before and after run-in, had shown that the effect of axial load on bearing thermal conductance was negligible over the conditions considered. This effect may be a result of differences in the heat-transfer mechanism across static versus dynamic bearings, as will be explained in more detail in the next section.

Furthermore, when comparing the static and dynamic thermal conductance magnitudes, it may be noted that the dynamic thermal conductance was higher

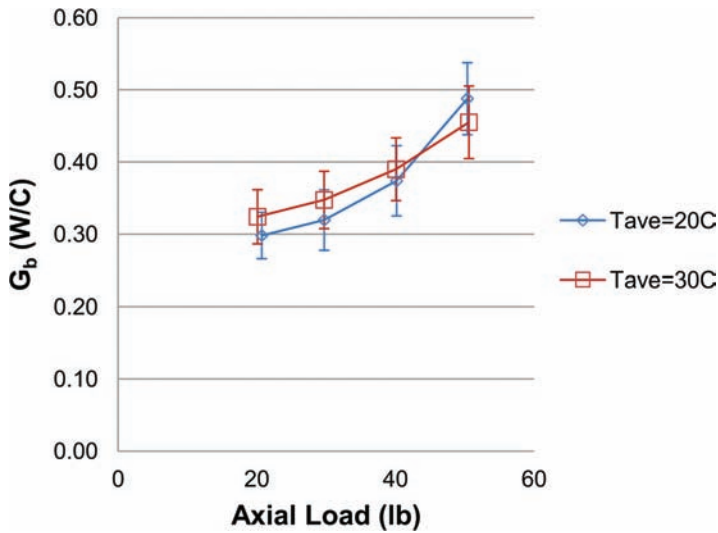


FIG. 16—Dynamic, run-in oil lubricated steel bearing at 6000 rpm, G_b versus axial load, for $\Delta T = 11^{\circ}C$.

than the static conductance, either before or after run-in, for any given axial load.

For a bearing rotating at a constant speed, the steel bearing initially seemed to respond to axial load slightly differently than the hybrid bearing. However, Fig. 18 overlays the thermal conductance results between the steel and hybrid

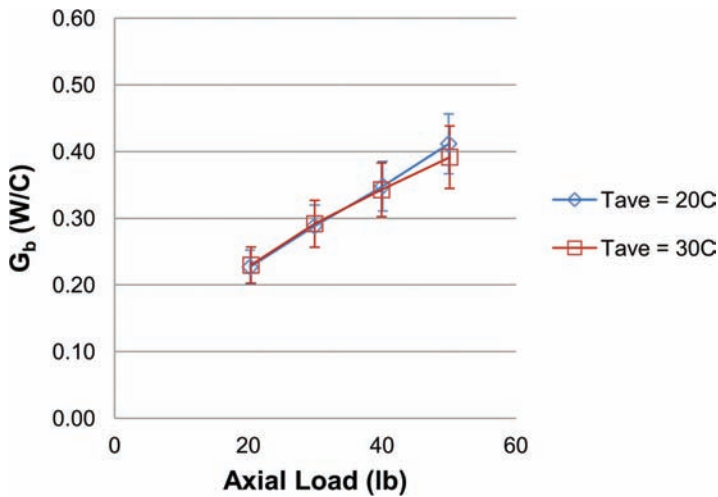


FIG. 17—Dynamic, run-in oil lubricated hybrid bearing at 6000 rpm, G_b versus axial load, for $\Delta T = 11^{\circ}C$.

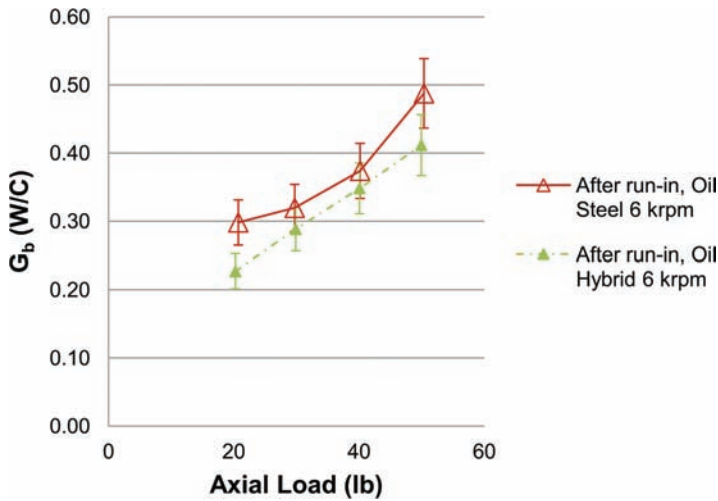


FIG. 18—Comparison of steel and hybrid bearing, oil lubricated at 6000 rpm, G_b versus axial load, for $T_{ave} = 20^\circ\text{C}$, $\Delta T = 11^\circ\text{C}$.

bearings; there is little difference in thermal conductance between the two bearing types.

The hybrid bearing seems to have a slightly lower thermal conductance than the steel bearing, and it may be noted that the lubricant quantity was slightly less than that of the steel bearing (90 mg after run-in for the hybrid; 100 mg after run-in for the steel). However, it should be noted that the differences fall within the uncertainty bounds of the experiment.

Dynamic and Oil Lubricated Bearings: Explanation of Results

As previously noted, there is a change in thermal conductance in response to axial loading between static and dynamic bearings. This section details a plausible explanation for the observed trends.

The mechanism of heat transfer for a static bearing is different than that for a moving bearing. For a static bearing, the heat-transfer mechanism is conduction through the races, balls, and the contact areas between the races and balls. The major source of thermal resistance in a static, conduction driven bearing is caused by the geometric constriction between the ball and raceway. The size of the contact area has been shown to drive the bulk thermal resistance across the static bearing. As previously noted, the degree of sensitivity to axial load is dependent on the size of the contact geometry caused by the thermal conductance being proportional to the square root of area, which is in turn proportional to the applied axial load to the 2/3 power. This means the smaller the area, as with dry bearings, the more sensitive thermal conduction is to axial loads, but this effect decreases with increased

size. The presence of lubricant vastly increases the contact area; thus, rendering its response to axial load outside the sensitivity range of the test hardware. For a dynamic bearing, mass transport dominates the heat-transfer process, as heat is absorbed by the lubricant and ball surface from the hotter race, is then transported by the ball motion, and finally is deposited at the cooler race. The conduction mechanism plays a much lesser role for a bearing moving at 6000 rpm.

With mass transport as the mechanism of heat transfer, a band of lubricant around the ball transports the bulk of the heat, as shown in Fig. 19. The width of the band is governed by the Hertzian contact length, a , which is the contact ellipse major axis between the ball and raceway. The major axis is oriented normally to the pathway; thus, governing the width of the band. As the load increases, the Hertzian contact length grows in size, as shown in Fig. 20, and thus, the lubricant volume carrying heat from the hotter to cooler race is increased, resulting in a higher bulk effective conductance across the bearing.

The length of the Hertzian contact ellipse varies proportionally to axial load to the $1/3$ power. Interestingly, over the load range considered in the experiments, i.e., from 20–50 lb, the contact ellipse length a increases in what seems to be an almost linear manner. This is a similar trend to the measured increase in bearing thermal conductance of a bearing rotating at a constant speed of 6000 rpm. However, it would be erroneous to extrapolate and consider the effect on thermal conductance as linear over a larger range of axial loads.

Furthermore, because of the higher Young's modulus of the ceramic ball as compared with the all-steel bearing, the Hertzian contact ellipse is slightly smaller, resulting in a smaller contact length a . Logically, this would mean that the hybrid bearing should have a slightly lower thermal conductance, which is consistent with measured observations. It should be noted, however, that, whereas the hybrid was consistently lower in thermal conductance over the

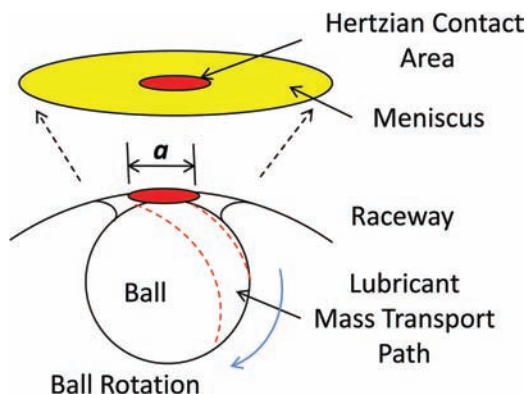


FIG. 19—Impact of oil lubrication on a dynamic bearing – thermal conductance theory.

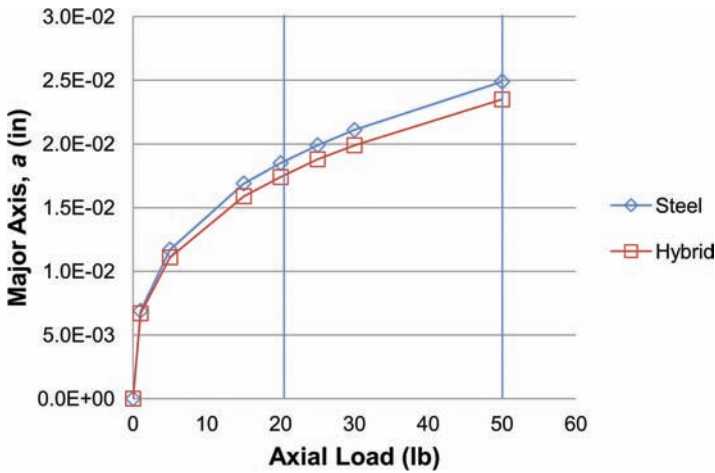


FIG. 20—Calculated hertzian contact ellipse major axis, a , versus axial load.

load range considered, the test data values are so close to each other that the curves fall within the uncertainty bounds of one another.

Summary

A number of parametric studies have been explored and results compared between all-52 100 steel and hybrid bearings. Over the examined temperature ranges, the bearing outer and inner race temperatures did not impact bearing thermal conductance.

The influence of axial loading, motion, lubricant presence and quantity had a substantial impact on bearing thermal conductance, as shown in Fig. 21. This graph summarizes the impact of axial loading on bearing thermal conductance for hybrid and 52 100 all-steel bearings for various lubricant conditions as well as the bearing state of motion.

Thermal conductance values were the lowest for the dry static bearing case, and responded to axial load to the $1/3$ power. The presence of oil greatly enhanced bearing thermal conductance and reduced the sensitivity of conductance to axial load in static bearings. Lubricant quantity also dominated the magnitude of static bearing thermal conductance. Larger oil quantities facilitated heat transfer across the bearing. For each static case considered, the mechanism of heat transfer was conduction and the dominant thermal resistance was attributed to geometric thermal constriction at the ball to race interfaces.

For the same oil lubricated bearing under constant 6000 rpm speed, the response to axial load was quite different as the mechanism of heat transfer changed from conduction to mass transport.

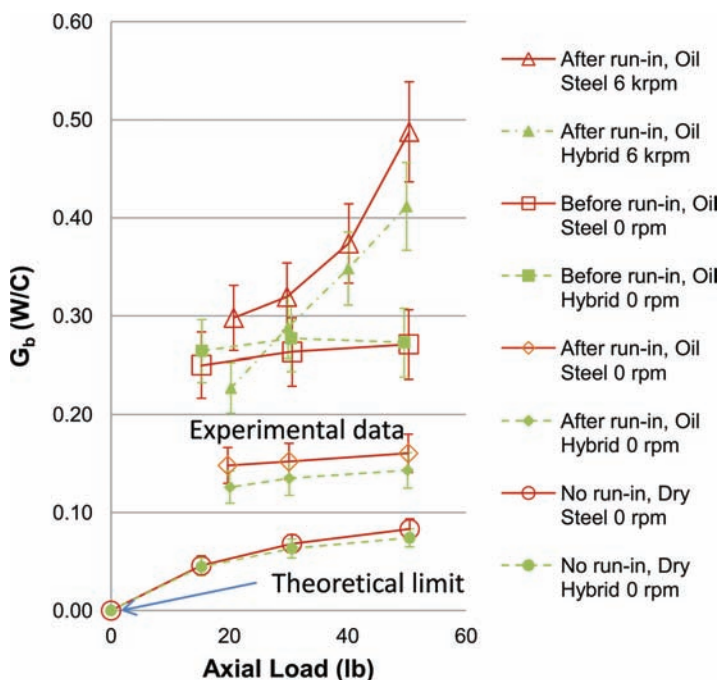


FIG. 21—Comparison of steel and hybrid bearing, G_b versus axial load, for $T_{ave} = 20^\circ\text{C}$; A Summary.

In each case, the results showed that the conductance magnitudes and trends are similar between hybrid and 52 100 steel bearings, indicating that ball material composition did not substantially impact the bulk effective heat transfer across the bearing.

Conclusion

Two thermal properties impact bearing temperatures in a vacuum environment: thermal conductance and heat generation. Heat generation tends to be the better understood quantity; thus, the focus of this study was the lesser known thermal conductance. The investigation showed that the thermal conductance magnitude and its response to operational and oil lubricant quantities were strikingly similar between steel and hybrid bearings. Thus, any observed differences in temperatures between hybrid and steel bearings may be assumed to be governed by heat generation and not by bearing thermal conductance.

Acknowledgments

The writers thank Mr. Donald Taylor and Mr. Michael Rosado, Research Associate and Laboratory Associate, respectively, at The Aerospace Corporation,

for their assistance in performing these experiments and maintaining the laboratory test rig. This work was funded by The Aerospace Corporation's Independent Research and Development (IRAD) Program.

References

- [1] Yovanovich, M. M., "Analytical and Experimental Investigation on the Thermal Resistance of Angular Contact Instrument Bearings," *Instrumentation Laboratory, E-2215*, M.I.T., Cambridge, MA, Dec 1967.
- [2] Yovanovich, M. M., "Thermal Constriction Resistance between Contacting Metallic Paraboloids: Application to Instrument Bearings," *Proceedings of Heat Transfer and Spacecraft Thermal Control*, M.I.T., Cambridge, MA, 1971, pp. 337–358.
- [3] Nakajima, K., "Theoretical and Experimental Investigation of the Thermal Contact Resistance of a Space-Use Deep Groove Ball Bearing," *Proceedings of the 18th International Symposium on Space Technology and Science*, AGNE, Tokyo, 1992, pp. 879–888.
- [4] Nakajima, K., "Thermal Contact Resistance between Balls and Rings of a Bearing under Axial, Radial, and Combined Loads," *J. Thermophys. Heat Transfer*, Vol. 9, No. 1, 1995, pp. 88–95.
- [5] Stevens, K. T. and Todd, M. J., "Thermal Conductance across Ball Bearings in Vacuum," *Report No. ESA-ESTL-25*, National Centre of Tribology, Risley, England, Feb 1977, pp. 1–51.
- [6] Delil, A. A. M., Heemskerk, J. F., and Vreeburg, J. P. B., "Design Report on the Esro Test Rig to Measure the Thermal Conductance and Friction Torque of Rotating Bearings in Vacuum," *Report No. NLR-TR-74069-U*, May 1974, pp. 1–90.
- [7] Heemskerk, J. F., Delil, A. A. M., and Vreeburg, J. P. B., "A Test Rig to Measure the Thermal Conductance and Friction Torque of Bearings in Vacuum," *European Space Tribology Symposium*, Frascati, Italy, 1975, pp. 1–18.
- [8] Rowntree, R. A. and Todd, M. J., "Thermal Conductance and Torque of Thin Section Four-Point Contact Ball Bearings in Vacuum," *Report No. ESA-ESTL-54*, National Centre of Tribology, Risley, England, Mar 1983, pp. 1–57.
- [9] Rowntree, R. A., "Stiffness, Torque, and Thermal Conductance of Thin-Section Four-Point Contact Ball Bearings for Use in Spacecraft Mechanisms," *Proceedings of the 1st European Space Mechanisms & Tribology Symposium*, Noordwijk, Netherlands, 1983, pp. 91–100.
- [10] Takeuchi, Y. R., Dickey, J. T., Demsky, S. M., and Eby, M. A., "Heat Transfer across Ball Bearings in Vacuum," *Proceedings of 2005 ASME Heat Transfer Conference*, San Francisco, CA, July 2005, pp. 1–9.

- [11] Demsky, S. M., Takeuchi, Y. R., and Dickey, J. T., "Thermal Conductance Measurement Techniques for High-Speed Bearings in Vacuum," *Proceedings of 2005 ASME Heat Transfer Conference*, San Francisco, CA, July 2005, pp. 1–8.
- [12] Takeuchi, Y. R., Eby, M. A., Blake, B. A., Demsky, S. M., and Dickey, J. T., "Influence of Oil Lubrication on Spacecraft Bearing Thermal Conductance," *Proceedings of the 38th Aerospace Mechanism Symposium*, Williamsburg, VA, May 2006.
- [13] Xinzhong, L., Denner, H.-D., Szucs, Z., and Klipping, G., "Static and Dynamic Thermal Conductance of Ball Bearing at LHe Temperatures," *Cryogenics*, Vol. 30, 1990, pp. 499–503.
- [14] Anderson, M. J., Roberts, E. W., and Rowntree, R. A., "The Thermal Conductance of Solid-Lubricated Bearings at Cryogenic Temperatures in Vacuum," *Proceedings of the 30th Aerospace Mechanisms Symposium*, European Space Tribology Lab., Cheshire, England, May 1996, pp. 31–45.
- [15] Winer, W. O., Bair, S., and Gecim, B., "Thermal Resistance of a Tapered Roller Bearing," *ASLE Trans.*, Vol. 29, 1986, pp. 539–546.
- [16] Khonsari, M. M., "Numerical Modeling of Thermal Effects in Journal Bearings," *Int. J. Modeling Simulation*, Vol. 8, No. 2, 1988, pp. 65–68.
- [17] Jędrzejewski, J., "Effect of the Thermal Contact Resistance on Thermal Behavior of the Spindle Radial Bearings," *Int. J. Machine Tools Manuf.*, Vol. 28, No. 4, 1988, pp. 409–415.
- [18] Jansson, R. M., "The Heat Transfer Properties of Structural Elements for Space Instruments," *Report No. E-1173*, Instrumentation Laboratory, M.I.T., Cambridge, MA, June 1962.
- [19] Hashimoto, E., Iwamoto, K., Honda, T., Ohashi, T., and Kushida, K., "Measurement of Thermal Conductance of Spacecraft-Use Bearings," *Proc. Space Sci. Technol. Conf.*, Vol. 28, 1984, pp. 524–525.
- [20] Eby, M. A., Takeuchi, Y. R., Davis, S. E., Fuller, J. K., Taylor, D. L., and Rosado, M. J., "Bearing Thermal Conductance Measurement Test Method and Verification," *2011 ASTM International Symposium on Rolling Element Bearings*, Anaheim, CA, April 2011.

*Christopher DellaCorte,¹ Ronald D. Noebe,¹
Malcolm K. Stanford,¹ and Santo A. Padula¹*

Resilient and Corrosion-proof Rolling Element Bearings Made from Superelastic Ni-Ti Alloys for Aerospace Mechanism Applications

REFERENCE: DellaCorte, Christopher, Noebe, Ronald D., Stanford, Malcolm K., and Padula, Santo A., "Resilient and Corrosion-proof Rolling Element Bearings Made from Superelastic Ni-Ti Alloys for Aerospace Mechanism Applications," *Rolling Element Bearings* on April 13–15, 2011 in Anaheim, CA; STP 1542, Yoshimi R. Takeuchi and William F. Mandler, Editors, pp. 143–166, doi:10.1520/STP103887, ASTM International, West Conshohocken, PA 2012.

ABSTRACT: Mechanical components (bearings, gears, mechanisms) typically utilize hard materials to minimize wear and attain long life. In such components, heavily loaded contact points (e.g., meshing gear teeth, bearing ball-raceway contacts) experience high contact stresses. The combination of high hardness, heavy loads, and high elastic modulus often leads to damaging contact stress. In addition, mechanical component materials such as tool steel or silicon nitride exhibit limited recoverable strain (typically less than 1 %). These material attributes can lead to Brinell damage (e.g., denting), particularly during transient overload events such as shock impacts that occur during the launching of space vehicles or the landing of aircraft. In this paper, a superelastic alloy, 60NiTi, is considered for rolling element bearing applications. A series of Rockwell and Brinell hardness, compressive strength, fatigue, and tribology tests have been conducted and are reported. The combination of high hardness, moderate elastic modulus, large recoverable strain, low density, and intrinsic corrosion immunity provides a path to the creation of bearings largely impervious to shock load damage. It is anticipated that bearings and components made from alloys with such attributes

Manuscript received March 28, 2011; accepted for publication March 13, 2012; published online August 2012.

¹National Aeronautics and Space Administration, Glenn Research Center, 21000 Brookpark Rd., Cleveland, OH 44135.

Copyright © 2012 by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

might alleviate many problems encountered in advanced aerospace applications.

KEYWORDS: materials, tribology, bearings, shock load resistance, mechanical components, superelastic materials

Introduction

The industrial revolution of the 18th and 19th centuries brought forth impressive new mechanical and materials technologies that vastly improved manufacturing and transportation but also created the need for an improved understanding of phenomena (stresses and deformations) encountered in highly loaded contacts (e.g., in bearings and gears). Hertz developed some of the earliest analytical stress models for contacting surfaces such as ball bearing-raceways and the wheel-rail interface [1]. The Hertz model is sufficiently general to include the radii of curvature of the surfaces, their elastic moduli, and the contact loads [2]. Hertz's calculations for contact stress, for example, explain what happens when two solid bodies contact each other and how damage occurs when the deformations and stresses exceed allowable material limits.

Such models have guided the design of mechanical components. For instance, Hertz calculations are used to estimate the contact stress in bearings between rolling elements and raceways, as well as between the contacting points of meshing gear teeth. Using "Hertz" stress levels as a guide, the contact geometry and material properties can be tailored so as to minimize stresses and distribute the load. Over decades of experimental development, so-called Hertzian contacts in machine elements have evolved to incorporate hard materials with commensurately high stiffnesses and relatively low values of ductility and recoverable elastic strain [3]. Though this approach yields long life and low friction, such components are susceptible to localized damage during high load events.

Contact that occurs between curved surfaces made from hard and stiff materials results in very small contact areas and high local stresses. In a bearing subjected to a shock or other type of overload situation, denting of the precision raceway can readily occur and is called Brinell-type damage [4]. Brinell denting of surfaces often leads to rough operation and premature fatigue failure and is avoided largely by limiting overloads and tailoring contact geometry so as to reduce the peak stress. The use of less rigid, softer materials could reduce local contact stresses but would also lead to the loss of load capacity and structural rigidity, higher operating friction, and generally lower fatigue life. Many of these negative effects can be traced to excessive deformation in the contact areas and the susceptibility of softer, less rigid surfaces to damage caused by foreign particles such as dirt and hard wear debris. Based upon the prevailing and successful use of hard, rigid materials in mechanical components, one can infer that hard and stiff is preferable to soft and less rigid.

An alternate method to improve bearing design might be to employ materials that are sufficiently hard to mitigate wear but, at the same time, sufficiently

compliant to minimize contact stress levels during high-load events. In effect, what might be more desirable than the use of hard, stiff materials for mechanical component applications is the use of hard but highly compliant materials. Such a material might seem like a paradox, given that hardness and rigidity typically go hand in hand. Furthermore, acceptable fatigue resistance is paramount for long life. Interestingly, but not widely recognized, the desire for such materials is not a new concept.

In the 1950s, Oberle was among the first to propose that a relationship exists between wear resistance and the ratio of material hardness (H) to elastic modulus (E), or the H/E ratio [5]. He called this ratio the Modell. In his landmark paper, he suggested that improved machine wear life could be attained not through the development of ultrahard materials alone but, rather, through the identification and development of materials more resistant to indentation-induced damage. Oberle assumed that the best materials for mechanical components were the ones that exhibited the high hardness needed in order to provide abrasion resistance but were still of sufficiently low elastic modulus to limit the localized stresses that occur when hard particles enter a contact. Recognizing that high hardness and high elastic modulus are generally coincident material properties, Oberle considered hard chrome plating on softer steel substrates as an alternate means to achieve a high effective Modell.

More recently, other researchers have followed this example in order to purposefully engineer coatings and surface treatments to enhance the H/E ratio and tribological performance [6]. Ceramic coatings, surface conversion treatments, and bonded metallic films on more elastic substrates have been extensively evaluated. It has been observed that the coatings that perform best in sliding are not necessarily the hardest but, rather, those with a high H/E ratio [5,6]. Though Oberle and more recent researchers seem to have concentrated upon sliding contacts, their approaches lead to equal, if not more compelling, arguments when applied to concentrated rolling contacts such as bearings, gears, and other mechanical components.

Engineering bearings from materials with a high Modell could enhance performance in several ways. A reduced elastic modulus will result in a higher degree of elastic deformation (strain) under a given load, leading to an increased contact area and reduced peak Hertz stress. Thus, in theory, higher load levels could be endured without exceeding local stress limits, provided that high levels of elastic strain are accommodated. Such a high Modell bearing should also be less susceptible to debris denting. Of course, an adequate material hardness would be required in order to mitigate more general types of wear such as abrasion. Successful bearing materials must also be compatible with lubricants such as oil or grease and be of sufficient strength to withstand typical machine loads. Other, equally critical attributes include the ability to be manufactured to great precision and of sufficient quality to provide reasonable fatigue properties.

The fact that no high Modell materials have yet been identified, even though Oberle expressed his desire for the development of such materials so long ago, suggests that such an avenue to improved wear performance is not possible. In fact, the recent proliferation of silicon nitride ball/steel race hybrid bearings appears to embody an approach opposite to the one promoted by Oberle. The adoption of ceramic rolling elements increases hardness and modulus and results in reduced contact areas with mating raceways, leading to higher Hertz stresses. Hybrid bearings are widely recognized as being prone to raceway denting because of these factors, and care must be taken to limit overall load levels. In the aerospace industry, shock loads can be unavoidable, and a recent example of such a failure can be readily found [7].

This paper evaluates the premise of using high Modell materials for bearings, gears, and other mechanical components by considering materials from the family of shape memory alloys. When properly prepared, some of these materials are hard but have a relatively low apparent elastic modulus and exhibit superelastic behavior. As an example, a nickel-titanium alloy, NITINOL 60, originally developed by the U.S. Navy, is evaluated for use in rolling element bearings. A series of experiments is conducted in order to establish the feasibility of using such materials for mechanical components, especially with respect to applications involving high dynamic loads.

In this paper, an ordered intermetallic binary nickel-titanium alloy is introduced that appears to satisfy Oberle's requirement of high Modell while offering additional significant benefits such as greatly extended elasticity, immunity from corrosion, tribological compatibility, and reduced density. Its feasibility for use in bearings and components is evaluated through a variety of mechanical and materials tests.

Materials: Background

Binary nickel-titanium (Ni-Ti) alloys are in widespread use in the medical and dental industries in applications in which their biocompatibility and unique superelastic and shape memory effect characteristics are readily exploited [8,9]. More recently, in the aerospace industry, shape memory alloy activated structures have been proposed and demonstrated for such applications as general flow control, adaptive inlets and nozzles, variable geometry chevrons, variable camber fan blades, and flaps and other hinged components [10,11]. These applications capitalize upon the temperature dependent large reversible strain change inherent in near equi-atomic Ni-Ti alloys (containing approximately 55 wt. % nickel) even when opposed by high constraining forces. The use of such alloys in mechanical components is problematic on two fronts. Near equi-atomic Ni-Ti alloys lack the dimensional stability required for precision mechanical assemblies such as bearings and gears. When they are used above or below their transformation temperature, large dimensional changes, due to

shape memory effects, are observed. Further, such alloys have low hardness, leading to high wear in tribological applications.

Equi-atomic Ni-Ti alloys, however, represent only a small fraction of the compositions in this family. Within the Ni-Ti system, and in closely related materials systems, there exist compositions that, when appropriately heat treated and processed, exhibit high levels of hardness and dimensional stability. For selected applications, the shape memory effect can be tailored so as to occur well below conceivable use temperatures. An example of one such alloy is the nickel-rich NITINOL 60 (60NiTi) alloy, which contains 60 wt. % nickel (55 at. %) and 40 wt. % titanium. Whereas the more commonly known NITINOL 55 (55NiTi) is soft and exhibits remarkable shape memory effects, NITINOL 60 can be made dimensionally stable and is easily hardened to Rockwell C 60 or higher through appropriate heat treatments. Recently, advanced manufacturing methods have been used to produce bearing quality rolling elements from NITINOL 60 that display excellent tribological behavior [12,13].

Table 1 lists relevant material properties for 60NiTi alongside the shape memory alloy 55NiTi and more conventional bearing materials. Although 60NiTi has about the same hardness as M50 tool steel and other bearing alloys, it displays a relatively low apparent modulus, similar to that of titanium. This yields a Modell (H/E) for 60NiTi that is about twice that of hardened steel and which satisfies Oberle's ideal of a bulk material with an enhanced Modell. In addition to having a high Modell, 60NiTi also exhibits a certain characteristic of shape memory alloys that is pertinent to applications in highly loaded contacts, namely, superelasticity. Superelasticity is defined here as a material's ability to undergo large (more than 3 %) strains that are fully recoverable. Typical bearing steels exhibit recoverable strain limits of less than 1 %.

In the following sections, a series of experiments are presented that are designed to determine the potential of a high Modell material such as 60NiTi to offer enhanced performance in applications such as bearings, gears, or other mechanical components. Hardness, tribology, fatigue, and basic materials tests are considered. These sub-component and indirect tests are used to support the concept of making shock resistant components utilizing high Modell materials. Furthermore, this paper explores the potential of other such materials to offer engineering solutions to longstanding problems and vulnerabilities facing precision mechanical systems.

Materials: Processing History

This study focused on a single composition of NiTi, namely, 60NiTi. Several diverse processing paths were used to fabricate samples for the various tests used to evaluate the potential of this material for bearings, gears, and other mechanical components. The goal was not to determine the best form of 60NiTi but, rather, to simply and conveniently generate a basic set of properties for this material.

TABLE 1—Nominal properties for conventional bearing alloys, 55NiTi, and 60NiTi.

Property	60NiTi	55NiTi	440C	Si ₃ N ₄	M-50
Density, g/cc	6.7	6.5	7.7	3.2	8.0
Hardness	56 to 62 HRC	35 to 40 HRC	58 to 62 HRC	1300 to 1500 Hv	60 to 65 HRC
Thermal conductivity, W/m °K	18	19	24	33	~36
Thermal expansion	~12.4 × 10 ⁻⁶ /°C	~10 × 10 ⁻⁶ /°C	10 × 10 ⁻⁶ /°C	2.6 × 10 ⁻⁶	~11 × 10 ⁻⁶ /°C
Magnetic	No	No	Magnetic	No	Magnetic
Corrosion resistance	Excellent	Excellent	Marginal	Excellent	Poor
Tensile/flexural strength, MPa	~1000	~900	1900	600 to 1200 (bend strength)	2500
Young's modulus, GPa	~95	~100	200	310	210
Poisson's ratio	~0.34	~0.34	0.3	0.27	0.30
Fracture toughness	TBD	TBD	22 MPa√m	5 to 7 MPa√m	20 to 23 MPa√m
Maximum use temperature, °C	~500	~300	~400	~1100	~400
Electrical resistivity, Ω cm	~80 × 10 ⁻⁶	~80 × 10 ⁻⁶	~36 × 10 ⁻⁶	Insulator	~60 × 10 ⁻⁶

Note: TBD, to be determined.

60NiTi balls were manufactured via a high temperature proprietary powder metallurgy process roughly similar to that described in the literature [14]. Pre-alloyed 60NiTi powder was hot iso-statically pressed into rough, spherical ball blanks that were then ground, hardened, polished, and lapped to produce high quality (grade 5) bearing balls 12.7 mm in diameter. A multi-step thermal process (heat treatment) was used to enable rough grinding of the bearing balls in a softened state followed by lapping to a very fine surface finish in a final hardened condition. The finished 60NiTi ball specimens, shown in Fig. 1, were bright and shiny in appearance and resembled conventional polished steel balls.

The 60NiTi plates were processed via two slightly different methods. In one case, plates were processed by means of vacuum induction skull melting followed by casting in graphite molds. Once removed from the molds, the ingots were subjected to hot isostatic pressing to consolidate any internal voids remaining from the casting process. Figure 2 shows a collection of castings from which test specimens were cut. This material is referred to herein as “cast” 60NiTi. In the second method, an electro-slag re-melt process followed by hot rolling was used to produce a plate material, referred to herein as “hot-rolled” 60NiTi. All samples, regardless of starting condition, were given a final hardening heat treatment (980°C for 1 h) followed by oil or water quenching. Some preliminary properties of the 60NiTi were also measured in the as-received condition (i.e., cast or hot-rolled), for comparison to the heat-treated material.



FIG. 1—*Photograph of 60NiTi polished ball specimens prior to testing.*



FIG. 2—60NiTi cast ingots from which test specimens were fabricated.

The density was measured as 6.71 g/cm^3 , which is about 15 % lower than the corresponding value for tool steel. Figure 3 shows the microstructure for a cross section of the 60NiTi ball specimens in their final hardened and polished condition. Hardness measurements indicated values in the range of 58 to 62 on

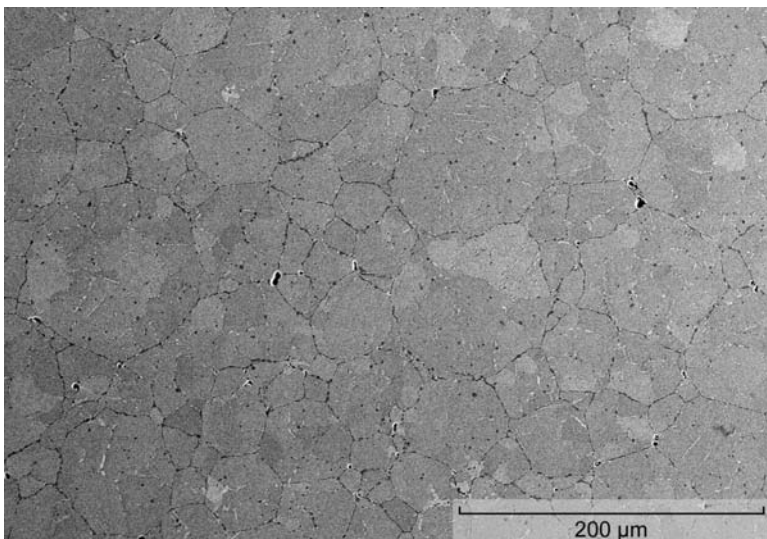


FIG. 3—Cross-section photomicrograph of a 60NiTi ball showing grain structure typical for powder metallurgy processed materials.

the Rockwell C scale (HRC). Figure 4 shows representative photomicrographs for several of the material conditions considered.

For comparison purposes, conventional bearing materials were also examined. Plates made from M50 tool steel, 440C stainless steel, and the cobalt alloy Stellite 6B (in the precipitation hardened condition) were prepared and tested. With regard to atmospheric attack, 440C stainless steel is considered a corrosion-resistant bearing material, whereas Stellite 6B is considered “corrosion-proof.” Additionally, indenter balls from bearing grade Si_3N_4 were procured from a commercial source.

Experimental Evaluations of 60NiTi Procedures and Results

Several conventional and unconventional experimental tests were conducted in order to determine whether a high Modell superelastic material such as 60NiTi

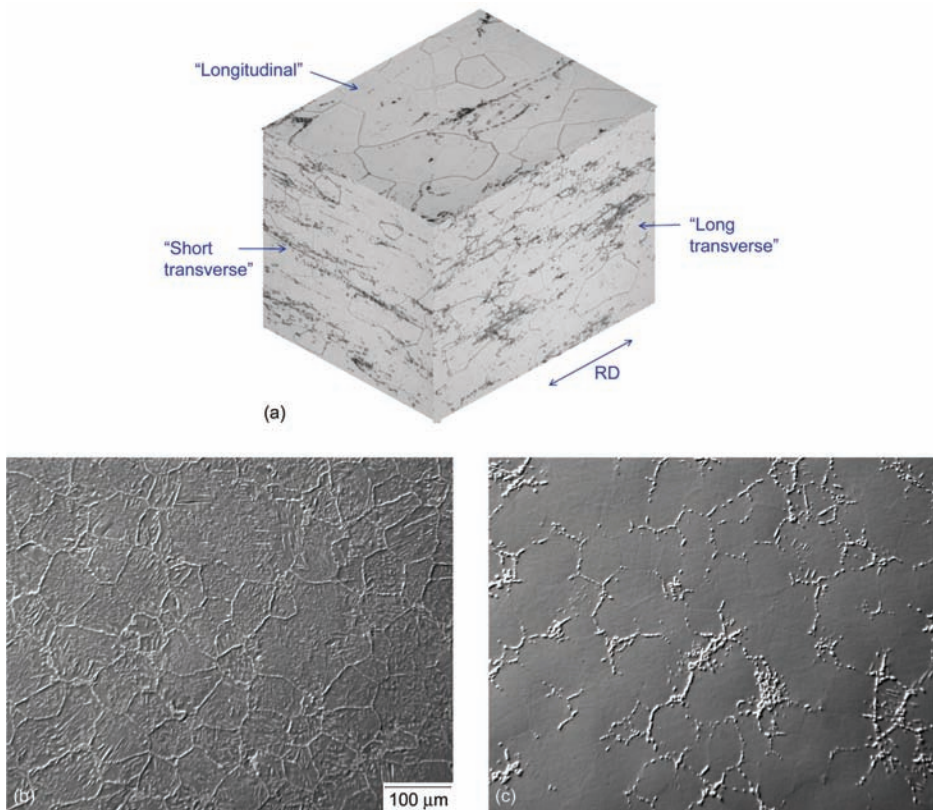


FIG. 4—Cross section of 60NiTi: (a) isometric view of hot-rolled material; (b) microstructure of cast and hot iso-statically pressed material; and (c) microstructure of the cast and hot iso-statically pressed material after heat-treatment.

can be a candidate material for mechanical components. These included conventional hardness tests, Brinell damage threshold tests, compression tests, rolling contact fatigue tests, and spiral orbit tribometer (SOT) tribology tests. The primary goal of these tests was to evaluate the possibility of making bearings and machine components from a superelastic and high Modell material that can better withstand dynamic overload forces without incurring permanent damage.

Hardness Test

Standard indentation (Rockwell C indenter) hardness tests were performed on the 60NiTi samples in the as-received and hardened conditions. In this hardness test, a pyramidal diamond indenter is forced into the sample under a 150 kg load. The penetration depth after indentation is automatically measured and reported as a numeric hardness value (HRC). Soft materials such as mild steel, annealed nickel alloys, and titanium typically register in the HRC 15–30 range. Hardened bearing steels typically yield values of HRC 50 to 65. The hardness results for the hot-rolled material were essentially the same as those for the cast material. Therefore, no distinction among these processing routes is made hereinafter. The test material is simply designated as either “hardened” or “as received.”

As-received 60NiTi was fairly soft and had a hardness value of about HRC 24 to 30. Following hardening heat-treatment, its hardness increased to HRC 58 to 60. Despite such a large increase in hardness, casual laboratory measurements and observations suggest that the material was dimensionally stable during the heat treatment, with no significant observed warping, shrinkage, expansion, or distortion. For comparison, the hardness values for typical bearing materials were also measured and are catalogued in Table 2. Samples of 440 C stainless steel exhibited hardness that mirrored that of 60NiTi, with the hardness ranging from HRC 58 to 60. M50 tool steel was a bit harder (HRC 60 to 63). The precipitation hardened cobalt alloy Stellite 6B was slightly softer (HRC 44 to 46). For long life and maximum load capacity, a material hardness above HRC 55 is generally preferred for bearing and gear applications.

TABLE 2—Rockwell C hardness for 60NiTi and typical bearing materials.

Material	Rockwell C Hardness, HRC
Stellite 6B	44 to 46
440C	58 to 60
M50	60 to 63
60NiTi	58 to 62

Brinell Indentation Hardness Test

The Brinell hardness test consists of pressing a large-diameter (12.7 mm) ball into the surface of a plate of the test material beyond the yield strength, resulting in a round and permanent dent. The diameters of the ball and of the dent, along with the load, are used to calculate the Brinell hardness. The higher the number, the more resistant a material is to indentation damage. In these tests, Si_3N_4 balls were used to indent plates of 60NiTi, 440C and M50 bearing steels, and Stellite 6B. A multi-purpose load frame was used for this test. The test load was applied gradually over a 120 s period and varied from 10 to 3000 kg_f depending upon the material tested. The test setup is shown in Fig. 5.

The data (Table 3) indicate that 60NiTi had the highest Brinell hardness at 530, followed by M50 at 500 and 440C at 435. Stellite 6B exhibited the lowest value, 310. Clearly, 60NiTi has excellent resistance to indentation damage, but the Brinell number alone does not give a full picture of 60NiTi's ability to resist damage; this is more fully delineated in the following subsections.

A follow-on measurement was made with the Brinell test set up (Fig. 5) in which discrete loads were applied to the plate specimens using Si_3N_4 and 60NiTi indenter balls at levels below those needed to cause dents. In these tests, a low load was placed on the indenter balls and then removed. If no dent was observed (visually), a new test was done using a higher load. The process

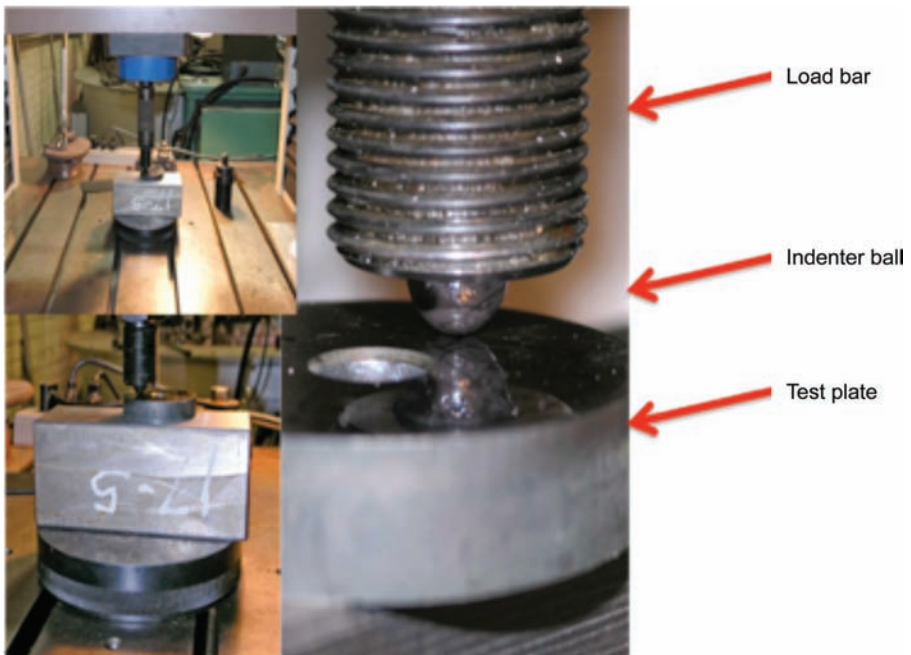


FIG. 5—Brinell hardness test setup used to press balls of Si_3N_4 or 60NiTi into various plate materials in order to measure the indentation resistance.

TABLE 3—Brinell hardness (BHN) and Modell (BHN/E) for 60NiTi and typical bearing materials.

Material	Elastic Modulus, GPa (MPsi)	Brinell Hardness Number	Modell Number
Stellite 6B	210 (30.4)	302	9.9
440 C	207 (30)	434	14.5
M50	207 (30)	493	16.4
60NiTi	100 (15)	531	37.9

was repeated until the damage threshold load was determined. Table 4 gives these threshold loads, along with the corresponding stress values calculated using the Hertz equations [1]. These values are of great interest in engineered systems because threshold load and stress—not a hardness number—are used to determine the machine design.

These results illustrate the value of using hard but highly elastic materials in concentrated contacts. Stellite 6B was indented with Si_3N_4 or 60NiTi balls at loads above only 10 and 15 kg, respectively. 440C withstood 51 or 150 kg, and M50 withstood 150 and 500 kg when loaded with Si_3N_4 and 60NiTi balls, respectively. Hardened 60NiTi proved to be the most dent resistant by far, withstanding 552 kg against Si_3N_4 and 1509 kg when contacted by the 60NiTi ball. Clearly, using highly elastic materials such as 60NiTi for one or both contacting surfaces leads to a dramatically enhanced loading capability. Such a loading capability exceeds that expected based upon hardness alone. In order to better understand the reasons for 60NiTi's exemplary resilience, one needs to consider its superelastic nature, and this is best illustrated through the standard compression test.

Compression Test

The most generally accepted and broadly understood test data come from simple compression tests conducted on small cylindrical samples. In this test, 5 mm diameter by 10 mm long cylinders of 60NiTi were placed between

TABLE 4—Hertz contact stresses and contact diameter at the threshold load for various plate and indenter material combinations.

Plate	Indenter	Threshold Load, kgf (lb)	Peak Stress, GPa (ksi)	Contact Diameter, mm (in.)	Average Stress, GPa (ksi)
Stellite 6B	Si_3N_4	10 (22)	2.06 (299)	0.30 (0.012)	1.37 (199)
440 C	Si_3N_4	51 (112)	3.48 (504)	0.52 (0.021)	2.32 (336)
M50	Si_3N_4	150 (331)	5.09 (738)	0.74 (0.029)	3.39 (491)
60NiTi	Si_3N_4	552 (1214)	5.56 (806)	1.36 (0.054)	3.71 (537)
Stellite 6B	60NiTi	15 (33)	1.56 (226)	0.42 (0.017)	1.04 (151)
440 C	60NiTi	150 (331)	3.33 (483)	0.92 (0.036)	2.22 (322)
M50	60NiTi	501 (1102)	5.02 (728)	1.37 (0.054)	3.35 (486)
60NiTi	60NiTi	1512 (3327)	5.90 (856)	2.19 (0.086)	3.94 (571)

hardened steel anvils in a servo-hydraulic test frame and deformed in displacement control at an initial strain rate of approximately $1 \times 10^{-4} \text{ s}^{-1}$ while the strain was measured as a function of applied stress. Repeated cycles were imposed on the same sample to a constant upper stress level determined from previous compression tests. The data are presented as stress versus strain, and the strength, apparent stiffness, and recoverable strain levels were assessed based on them. Figure 6 shows the compression test results for six repeated stress cycles on the same hot-rolled plus hardened specimen.

When initially loaded in compression, the 60NiTi sample deformed as the load increased in a slightly non-linear fashion up to the load limit capability of the test fixtures, in this case $\sim 2500 \text{ MPa}$. When the load was reduced, the majority of the strain was recovered, save for 1 %, most likely due to plastic deformation. Upon repeated loading, the sample deformed reproducibly between the same limits, with only a small hysteresis observed between the loading and unloading segments. Subsequent loading cycles showed little deviation from this reproducible behavior.

Two important aspects of the material behavior were revealed from this test. The first is that a small amount ($\sim 1 \%$) of non-recoverable strain occurred during the very first loading event. The implications of this one-time initial 1 % non-recoverable strain are discussed later. The other aspect is that 60NiTi withstood exceedingly large amounts of recoverable strain over repeated cycles, with no observed degradation. Typical bearing steels can withstand only a few tenths of a percent of strain with full recovery when the load is removed. The

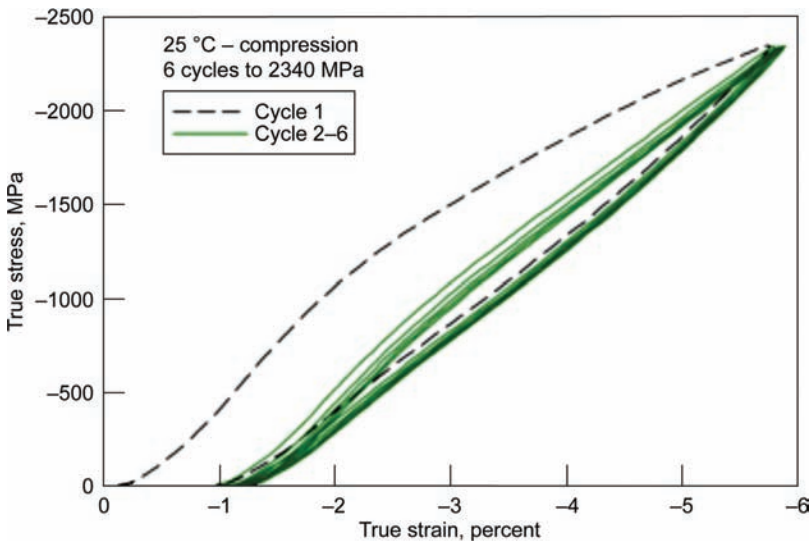


FIG. 6—Compression behavior of hot-rolled plus hardened 60NiTi. Note: The maximum stress level is limited by the test fixtures, not the specimen.

extraordinary levels of recoverable strain observed for 60NiTi are typical for superelastic materials and help explain why such high loads were required in order to achieve a dent in the Brinell test.

Rolling Contact Fatigue Test

Rolling contact bearing materials must possess acceptable rolling contact fatigue (RCF) behavior. To evaluate the RCF performance, 60NiTi rods were cut from the hot-rolled plus hardened plate using a wire electrode discharge machine and ground and lapped to achieve an acceptable surface finish (roughness $\sim 0.4 \mu\text{m rms}$). These were placed into a standard Federal Mogul style tester, lubricated by mineral oil, and loaded by three tool steel balls at three load levels corresponding to 266, 356, and 429 ksi contact stress. The test speed was 3600 rpm, and the time to first spall was evaluated as indicated by accelerometer detection. The results are summarized in Table 5. For comparison, high quality, fully commercialized bearing steels can withstand about 500 ksi without experiencing fatigue failures.

The 60NiTi results varied widely. At the highest stress, 429 ksi, life ranged from 2.2 to 8.4 h. Clearly, this stress level exceeds the capability of the currently available rods. At 356 ksi, the life ranged from a low of 4.2 h to well over 169 h in one test. At the lowest stress, 266 ksi, life ranged from 13.8 h to greater than 840 h. In this latter case, there was no failure and the test was intentionally ended, and the fatigue life was deemed essentially infinite. Examination of the spalls and close scrutiny of the rods using cross-sectional metallography coupled with scanning electron microscopy and energy dispersive spectroscopy x-ray analyses helped clarify the sporadic fatigue behavior observed.

Figures 7 and 8 show images of the wear tracks following rolling tests at 266 and 429 ksi, respectively. Under the low load, the wear was a mild polishing-type wear. At the higher load, the failure appeared as a classical fatigue spall likely originating at subsurface flaws, a characteristic common for hard, limited ductility materials undergoing repeated contact. Cross-section images of the microstructures shown in Fig. 4(a) revealed stringers of oxides and other hard inclusions, residual porosity, and multiple coarse phases within the NiTi matrix. The microstructure after the hardening treatment was essentially the same as that shown in Fig. 9.

TABLE 5—Preliminary RCF life for 60NiTi rods (loaded against M50 steel balls, 3600 rpm, mineral oil lubricant).

RCF Stress Level, Ksi	Wear Track Width, μm	RCF Life, h
429	850 ± 18	5.7 ± 3.2
356	701–/–25	5.4 (typical)
266	540	13.8 to 841 (no failure)



FIG. 7—Wear track surface following RCF test at 360 ksi; no failure observed.

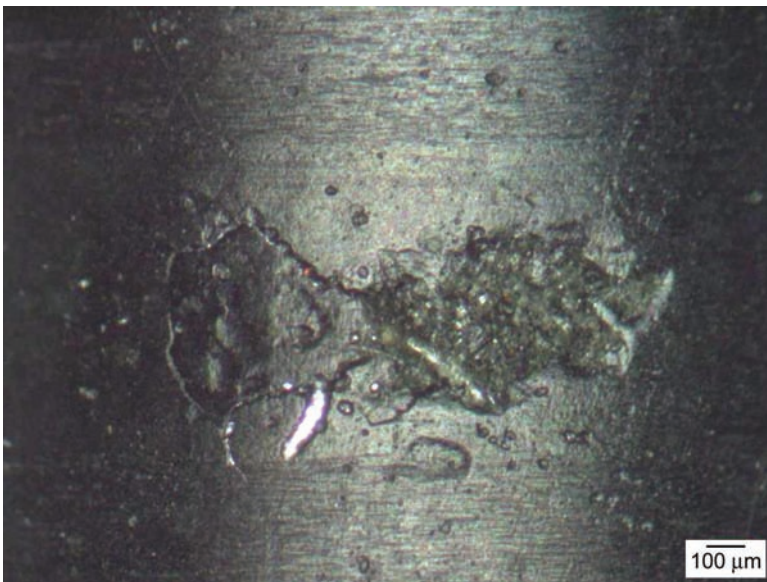


FIG. 8—Spall damage on a hardened 60NiTi rod surface following testing at 580 ksi.

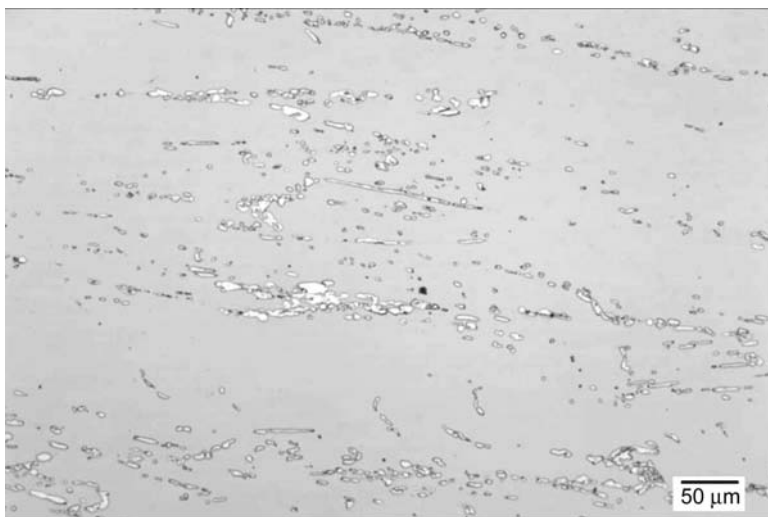


FIG. 9—*Optical micrograph of a cross-section of one of the hardened 60NiTi RCF rods showing voids and second phase precipitates.*

These flaws and the large variability in their size are undoubtedly contributing factors to the scatter and early failures seen in many of the tests. It is likely that those samples that exhibited superior fatigue performance, nearing that of commercial steels, had smaller subsurface defects than those with shorter lives. Further fatigue tests are planned that will utilize higher quality specimens made via powder metallurgy and other processing routes that will reduce the size and density of such defects. Despite the variability, even the current 60NiTi rods can be considered for modestly loaded applications with respect to fatigue. Furthermore, given the poor quality of the tested material in terms of the microstructure, the fatigue data are actually quite encouraging.

Spiral Orbit Tribometer Tribology Testing

The tribological behavior of 60NiTi in a lubricated rolling contact was evaluated using an SOT. In this test, a single powder metallurgy processed 60NiTi ball was loaded between two rotating 440C steel disks, lubricated with synthetic oil, and tested in a vacuum. The 60NiTi ball was marginally lubricated with a small quantity ($\sim 25 \mu\text{g}$) of oil. The test conditions fell within the boundary lubrication regime that helped assess the tribological and tribo-chemical compatibility of the material.

The SOT, depicted in Fig. 10 and described in detail elsewhere [12,15], is basically a thrust bearing with one ball and flat races (plates). One of the plates was stationary, and the other rotated to drive the ball into an orbit that was an opening spiral. The ball contacted a “guide plate” at the end of each orbit that

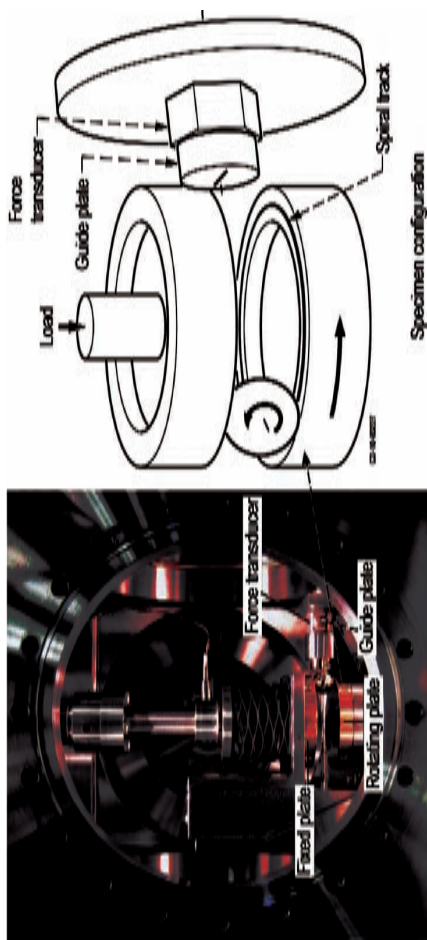


FIG. 10—Spiral orbit tribometer used to evaluate the relative lubricant life of various alloys under simulated bearing conditions.

forced the ball back into its initial orbital radius. A piezoelectric force transducer supporting the guide plate sensed the frictional force that developed on the ball as it slid on the rotating plate. During this contact, the coefficient of friction was obtained from this force and the load imposed on the system. Prior to testing, the ball specimen was dipped into a dilute oil-solvent solution, depositing a thin ($\sim 25 \mu\text{g}$) volume of lubricant onto the surface. Tests were typically run for several thousand cycles, until the oil was consumed and friction rose. The number of test cycles, prior to a steep rise in friction, was used as the indication of tribological performance.

Under these conditions, which are considered representative of precision bearings, the performance (life and friction) was comparable to that observed with 440C bearing balls. Figure 11 shows the friction results from several comparative tests. The results clearly show that 60NiTi met or exceeded the tribological capabilities of the standard 440C stainless steel. Typical lubricant lifetimes were approximately 30 000 to 60 000 revolutions, indicating a slow lubricant consumption rate. In one test (not shown) using both 60NiTi plates and balls, the rig was permitted to run well beyond the complete consumption of lubricant. The friction climbed to a coefficient of around 0.35 but was stable, and no galling tendencies were observed.

Also shown in Fig. 11 are SOT results for solid Ti-6Al-4V balls and 440C balls that had been previously sputter coated with a thin film of pure titanium. Both of these ball material configurations aggressively degraded the lubricant and experienced high friction and seizure within a few dozen revolutions.

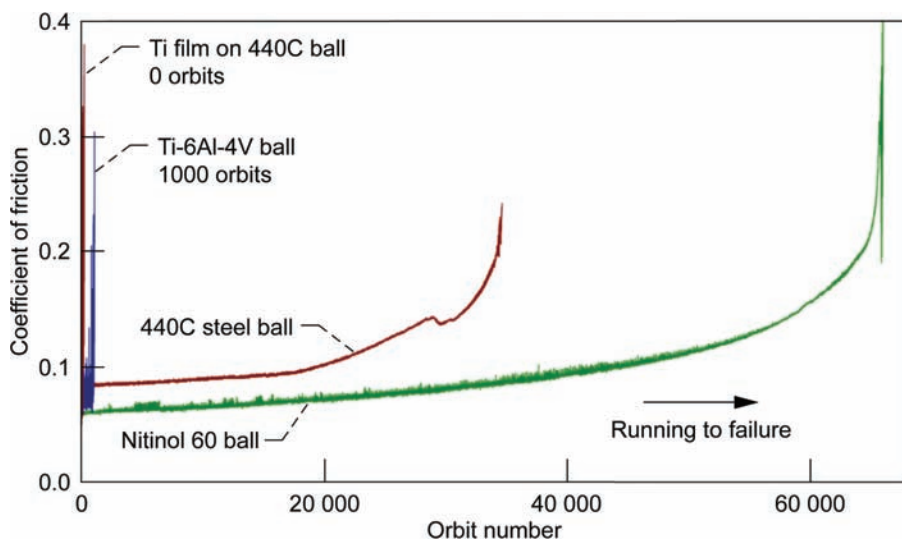


FIG. 11—SOT results for various balls running against 440C plates lubricated with Pennzane 2001A oil.

Discussion

The primary goal of this investigation was to assess the potential of highly elastic, hard materials such as 60NiTi to satisfy the demands of tribological contacts and to withstand extreme loading events in mechanical systems while remaining undamaged. A series of novel and conventional experimental measurements were used to provide input for such an assessment. The results of these experiments are remarkably consistent and support the validity of using hard superelastic materials for mechanical components.

The SOT tribology test results indicate that 60NiTi exhibits behavior comparable to that of high performance bearing steels. This result is welcome, although somewhat unexpected, given the significant differences in the surface chemistries of steels and 60NiTi. The long lubricant lifetimes exhibited might also indicate that further improvements might be possible through a more appropriate choice of lubricant or the use of additives tailored for nickel and titanium metallurgy. In addition, the tribology results indicate the potential for reduced friction.

Figure 12 shows the friction coefficients for 440C and 60NiTi balls rolling on 440C plates. The friction for 60NiTi is consistently lower than that for 440C. The exact reason for the lower friction with 60NiTi is not yet fully understood, but it might be related to the superelastic behavior of the 60NiTi ball in that the normal elastic deformation of the surfaces as they contact is

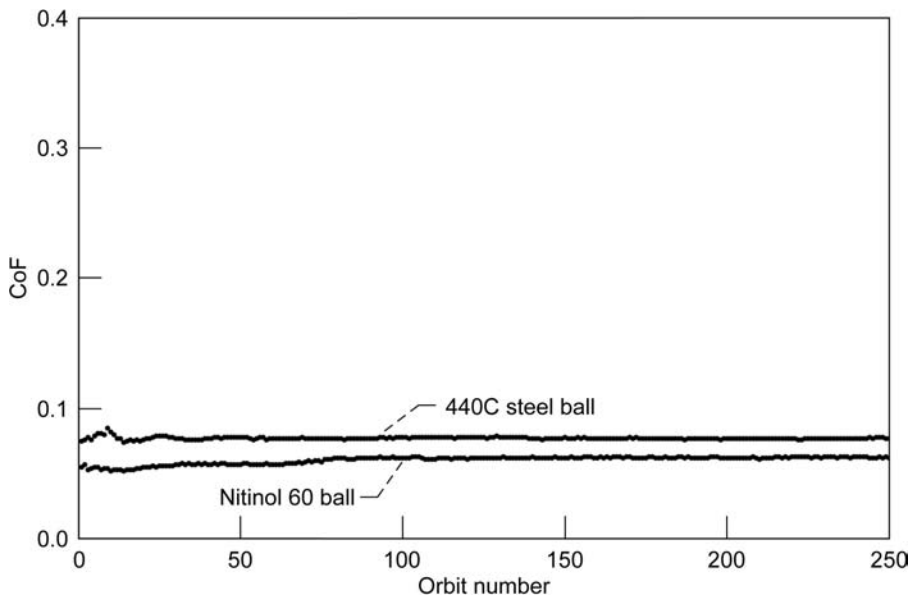


FIG. 12—Initial friction traces for a 440C steel ball and a 60NiTi ball rolling against 440C plates. Both balls were lubricated with $\sim 25 \mu\text{g}$ oil.

better recovered than in steel ball materials, which experience deformation losses (hysteresis). Based upon these considerations, the lower friction and higher overall efficiency of bearings, gears, drives, and mechanical components might be expected.

The results from the most basic standard Rockwell and Brinell hardness tests offer little evidence that 60NiTi differs to any appreciable extent from a common bearing material. The Brinell hardness of 60NiTi was around 530, and its Rockwell C hardness was 60 to 62. These values compare closely to those of M50 and 440C tool steels and indicate good prospects for the use of 60 NiTi in highly concentrated contacts.

The conventional compression test results illustrate the unique and unusual deformation behavior of 60NiTi. The compressive behavior of 60NiTi was found to differ substantially from that of other hardened bearing materials. Hard bearing materials have a low strain to failure, typically 1 % or less. They are stiff, unyielding, and brittle. 60NiTi, in contrast, is superelastic. Its stress strain behavior is complex, as shown in Fig. 6. In the first loading cycle, strain builds slowly over a broad strain range. At more than 5 % strain and 2.5 GPa load, the test fixture had reached its contact load limit. The load was then reduced, and we can see that at full unloading over 4 % of the strain was recovered. Subsequent loading cycles exhibited only fully recoverable “elastic-like” responses. In reality, this was superelastic behavior due to the formulation of stress-induced martensite during loading and its transformation back to austenite during unloading. In the current setup, the maximum allowable strain in compression was not reached or measured. Relative to conventional bearing steels, 60NiTi exhibited remarkably high levels of recoverable strain. It is hard, yet it combines a low apparent modulus and an extraordinarily extended region of recoverable deformation. Based upon the slope of the stress-strain curve and the result from the Brinell threshold load limit tests, which showed permanent deformation only above 5.7 GPa, one can extrapolate that the recoverable elastic strain limit lies near 10 % strain. This will need to be verified by future tests.

With respect to engineering applications, the superelastic behavior of 60NiTi offers significant benefits. Highly deformable materials that can also withstand high levels of recoverable strain are more forgiving of overload events. Under extreme loads, such surfaces will momentarily and elastically deform, creating a relatively large contact area with commensurately low contact stress and a reduced chance of permanent damage. Highly recoverable materials can also better withstand debris-induced damage. For instance, in a ball bearing, the superelastic nature of 60NiTi can allow a ball to overrun a hard debris particle by temporarily deforming rather than incurring a permanent dent through plastic flow. In order to realize this benefit, some pre-straining (e.g., coining or forging) will be required before a component reaches its finished geometry. This is necessary in order to eliminate the potential for

plastic deformation as shown in the first loading cycle in Fig. 6. Such treatment is common with high performance materials, and for 60NiTi it will ensure that all the deformation that occurs in use is of the fully recoverable type.

The Brinell style indentation experiments used to determine the load threshold for permanent damage provide compelling evidence that a superelastic material can contribute to the improvement of mechanical components and devices. The threshold load for damage for the 60NiTi system (ball and plate) is at least 1 order of magnitude higher than that of the best hybrid bearing system currently in use. The reason for this behavior can be best understood through the contact stress modeling embodied in the Hertz equations. For a given contact geometry, the stress between a ball and a plate is dictated solely by the elastic modulus of the ball, the plate, and the load. Table 4 lists the results. The lower stiffness of 60NiTi results in a larger deformed contact area, and thus the peak and average stresses are reduced compared to those of materials with higher elastic moduli. NITINOL's superelastic behavior enhances this effect in two ways. First, its superelasticity causes it to appear to have an even smaller "effective" modulus, further reducing peak stresses. Second, its ability to recover extremely large strains results in extraordinarily high damage threshold loads. The behavior of 60NiTi, as observed in this series of

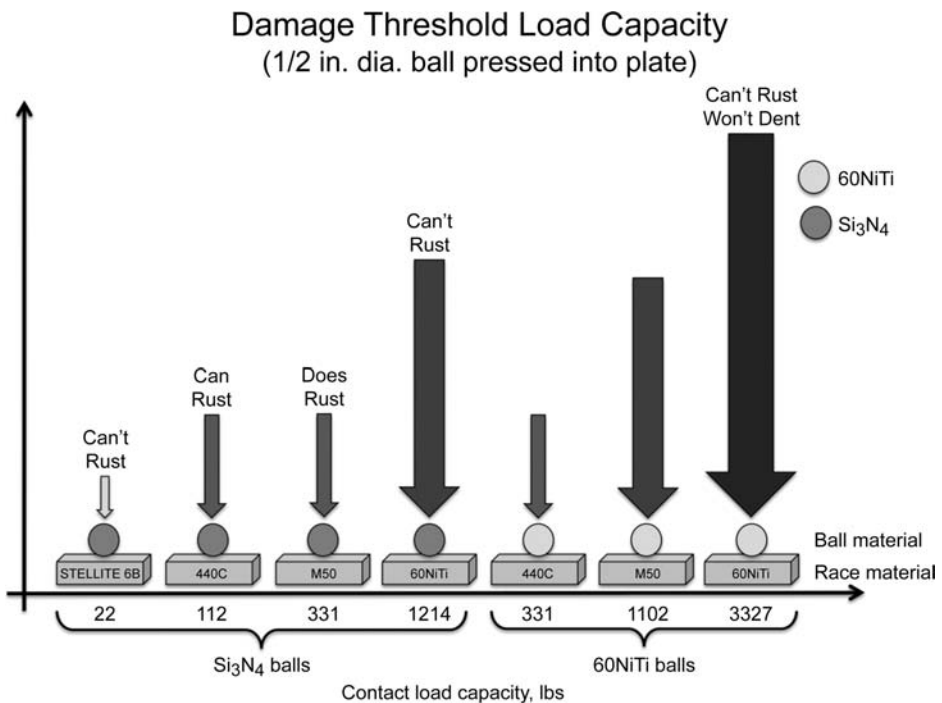


FIG. 13—Graphical depiction of the indentation load tolerance of various bearing material combinations.

experiments, is perfectly understandable once these effects are taken into account. The combination of low modulus and high recoverable strain provides load tolerance, and the high hardness gives good wear resistance.

In effect, the use of a high Modell material as postulated by Oberle in the middle of the 20th century appears quite valid. 60NiTi's additional attractive properties, such as low density, intrinsic corrosion resistance, and non-magnetic properties, add to the compelling case that it is a breakthrough in mechanical component materials. Figure 13 graphically depicts the combined benefits of this novel approach with respect to load damage tolerance and environmental corrosion.

Summary Remarks

This paper presents a novel approach for improving the resiliency, debris and shock load tolerance, and efficiency of mechanical components via the adoption of emerging hard yet superelastic construction materials. Through a series of conventional and unconventional materials tests, it has been shown that low modulus, hardened superelastic materials, such as 60NiTi, can withstand extreme loading conditions without incurring permanent surface damage. In contrast, bearing and gear materials currently in widespread use readily sustain damage at much lower load levels. The test results, when taken together, indicate that the superior performance of loaded contacts made from this hard superelastic material is directly related to its unique combination of high hardness, significantly reduced "effective" modulus, and extremely high levels of recoverable strain (*superelastic behavior*). Based upon the investigations, the following specific remarks are made:

- The use of hardened materials with a modest effective elastic modulus that exhibit superelastic behavior results in significantly enhanced load capability for contacting surfaces.
- The extreme levels of recoverable deformation that occur when superelastic low effective modulus materials such as 60NiTi are placed under heavy contact loads result in larger Hertzian contact areas that effectively distribute the load, reduce peak and average stresses, and prevent contact damage.
- 60NiTi appears to undergo a small amount of irrecoverable plastic deformation during initial loading. After that, it exhibits remarkably reversible, superelastic strain behavior, as evidenced by cyclic compression tests. These observations indicate that some mechanical working of the material, such as coining or forging, might be desirable prior to use.
- When the exemplary load tolerance of 60NiTi is considered with its previously proven benign tribological behavior, a compelling case for significantly enhanced machine component performance is expected.

- When one considers that the 60NiTi material is the least complex superelastic alloy currently available, it is likely that through alloying and other developments, materials with greatly enhanced performance will emerge.

Though much more research will be required before we can fully understand 60NiTi material and its metallurgical relatives, the results obtained thus far point to a very bright future for many applications in which extreme loading is unavoidable. Further, it is anticipated that many other materials with similar superelastic behavior will emerge to greatly expand the tools available to mechanical design engineers for problem solving and technological development.

Acknowledgments

The writers wish to acknowledge the financial support from the NASA Engineering and Safety Center (NESC) and NASA's Subsonic Rotary Wing and Supersonics Projects, as well as the Timken Corporation for granting permission to publish the preliminary RCF results.

References

- [1] Hertz, H., "On the Contact of Rigid Elastic Solids and on Hardness," *Verhandlungen des Vereins zur Beforderung des Gewerbefleisses*, Leipzig, 1882.
- [2] Dowson, D., "The Industrial Revolution," *History of Tribology*, Professional Engineering Publishing Limited, London, 1998, pp. 187–285.
- [3] Stone, D. H. and Steele, R. K., "The Effect of Mechanical Properties upon the Performance of Railroad Rails," *Rail Steels—Developments, Processing and Use*, ASTM STP 644, D. H. Stone and G. G. Knupp, Eds., ASTM International, West Conshohocken, PA, 1978, pp. 21–62.
- [4] Palmgren, A., "Bearing Failures," *Ball and Roller Bearing Engineering*, 3rd ed., SKF Industries, Philadelphia, PA, 1959, pp. 217–225.
- [5] Oberle, T. L., "Properties Influencing Wear of Metals," *J. Met.*, Vol. 3, 1951, pp. 438–439G.
- [6] Matthews, A. and Leyland, A., "Design Criteria for Wear-resistant Nanostructured and Glassy-metal Coatings," *Surf. Coat. Technol.*, Vol. 177–178, 2004, pp. 317–324.
- [7] Carter, D. L., "Status of the Regenerative ECLSS Water Recovery System," *NASA Report 2009-01-2352*, NASA Marshall Space Flight Center, Huntsville, Alabama, 2009.
- [8] Hodgson, D. E., Wu, M. H., and Biermann, R. J., "Shape Memory Alloys," *Metals Handbook*, Vol. 2, 10th ed., ASM International, Metals Park, OH, 1990, pp. 897–902.

- [9] Shabalovskaya, S. A., "Surface, Corrosion and Biocompatibility Aspects of NITINOL as an Implant Material," *Biomed. Mater. Eng.*, Vol. 12, 2002, pp. 69–109.
- [10] Stebner, A., Padula, S. A., Noebe, R. D., and Quinn, D. D., "Characterization of $\text{Ni}_{19.5}\text{Ti}_{50.5}\text{Pd}_{25}\text{Pt}_5$ High Temperature Shape Memory Alloy Springs and Their Potential Application in Aeronautics," *Proc. SPIE*, Vol. 6928, 2008, 69280X.
- [11] Mabe, J. H., Calkins, F. T., and Butler, G. W., "Boeings Variable Geometry Chevron, Morphing Aerostructure for Jet Noise Reduction," *AIAA—2006-2142*, American Institute of Aeronautics and Astronautics, Washington, D.C., 2006.
- [12] DellaCorte, C., Pepper, S. V., Noebe, R., Hull, D. R., and Glennon, G., "Intermetallic Nickel-Titanium Alloys for Oil-lubricated Bearing Applications," *NASA/TM—2009-215646*, National Aeronautics and Space Administration, Washington, D.C., 2009.
- [13] Pepper, S. V., DellaCorte, C., Noebe, R., Hull, D. R., and Glennon, G., "NITINOL 60 as a Material for Spacecraft Triboelements," *ESMATS 13 Conference*, Vienna, Austria, 13th European Space Mechanisms and Tribology Symposium 2009, September 23–25, European Space Agency, Paris, France, 2009.
- [14] McNeese, M. D., Lagoudas, D. C., and Pollock, T. C., "Processing of TiNi from Elemental Powders by Hot Isostatic Pressing," *Mater. Sci. Eng., A*, Vol. 280, 2000, pp. 334–348.
- [15] Jones, W. R., Jr., Pepper, S. V., Jansen, M. J., Nguyen, Q. N., Kingsbury, E. P., Loewenthal, S. H., and Predmore, R. E., "A New Apparatus to Evaluate Lubricants for Space Applications—The Spiral Orbit Tribometer (SOT)," *NASA/TM—2000-209935*, National Aeronautics and Space Administration, Washington, D.C., 2000.

J. T. Hanks,¹ D. W. Smith,¹ C. J. Stevens,²
and R. E. Winkel³

Evaluating the Impact of a Surprise Silicone Additive to a Synthetic Hydrocarbon Lubricant

REFERENCE: Hanks, J. T., Smith, D. W., Stevens, C. J., and Winkel, R. E., "Evaluating the Impact of a Surprise Silicone Additive to a Synthetic Hydrocarbon Lubricant," *Rolling Element Bearings* on April 13–15, 2011 in Anaheim, CA; STP 1542, Yoshimi R. Takeuchi and William F. Mandler, Editors, pp. 167–192, doi:10.1520/STP103909, ASTM International, West Conshohocken, PA 2012.

ABSTRACT: A fractional percentage of a silicone compound was inadvertently introduced in one lot of Rheolube synthetic hydrocarbon grease used in the spin bearings of our momentum control systems. As a result, a decision had to be made in a very short time as to whether the grease would meet the performance and life requirements. To accomplish this, a series of tests were undertaken to compare the performance of the subject grease or oil extracted from it to compare to the baseline lubricant. These tests represented each phase of bearing life: vibration, beginning of life, normal operational life, and end of life. Included were two different wettability tests, two different types of elastohydrodynamic lubrication tests and two different types of boundary tests. Acceptance test data of existing units with and without the subject grease were also used to determine how operational performance was affected. In addition to this test data, the Honeywell/Hanks lubricant life model was used to conclude that the subject grease would support our requirements.

KEYWORDS: lubrication, grease, silicone, contamination, bearings, space, tribology

Manuscript received April 28, 2011; accepted for publication January 17, 2012; published online April 2012.

¹Fellow Engineers, Honeywell International Defense and Space, Glendale, AZ 85308, USA.

²Principal Engineer, Honeywell International Defense and Space, Glendale, AZ 85308, USA (Corresponding author), e-mail: carlos.stevens@honeywell.com

³Senior Chief Engineer, Honeywell International Defense and Space, Glendale, AZ 85308, USA. 2011 International ASTM Rolling Element Bearings on 13-15 April 2011 in Anaheim, CA.

Copyright © 2012 by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

Nomenclature

ATP =	acceptance test procedure
BOL =	beginning of life
COF =	coefficient of friction
CMG =	control moment gyroscope
EOL =	end of life
EHD or EHL =	elastohydrodynamic lubrication
IV =	induced vibration
NOL =	normal operating life
OPL =	operating preload
Ra =	Roughness average
RPM =	revolutions per minute
rss =	root-sum-square
RWA =	reaction wheel assembly
SRV =	Schwingung, Reibung, Verschleiß (German), oscillating, friction, wear (English)
WAM =	Wedeven Associates machine
Ue =	entrainment velocity

Introduction

Honeywell uses Rheolube 2000 synthetic hydrocarbon grease in the bearing systems of Reaction Wheel Assemblies (RWA's) and Control Moment Gyroscopes (CMG). Each of these devices must pass a rigorous acceptance test procedure (ATP) prior to delivery. The ATP data and factory family data is routinely collected to characterize performance that may indicate potential future problems, so any out-of-family test data observed is thoroughly evaluated. Recently an RWA exhibited an uncharacteristic emitted vibration spectrum after random vibration exposure based on factory characterization tests. As part of the anomaly evaluation, a sample of grease was sent to the Aerospace Corporation for chemical examination, a routine practice at Honeywell. Results of the lubricant analysis showed that 0.5 % to 1 % of a dimethylphenyl silicone compound known as OCF-452 was present in the lubricant; this combination will henceforth be referred to in this report as the "subject" lubricant.

Anecdotal, conventional wisdom and some research data had previously given the combination of silicones and hydrocarbons in a lubricant an uncertain reputation, primarily due to the concern of dewetting of the bearing surfaces and possibly bearing failure. The silicone presence now became a reason for further investigation even though no unusual performance from units lubricated with this grease formulation had been observed on the delivered hardware.

Honeywell launched an investigation to determine the impact on operating performance over life of this surprise additive and to identify affected hardware, specifically how many lots of grease were affected and how many units

were lubricated with grease that had this surprise additive. Further investigation found that only one grease lot was affected however, 150 RWA units were lubricated with the subject grease lot and were either in ATP, had been shipped, or were installed and ready for launch.

Because of this, a determination needed to be quickly made as to whether or not this formulation would support product performance and life requirements.

Approach

Short term performance data was available to address early life performance, but many of the RWA's have life requirements in excess of 15 years; therefore a real time life test was not an option. Because of this, tests and analysis were devised to evaluate the effects of the lubricant performance at "phases" for Beginning of Life (BOL), during Normal Operating Life (NOL), and at End of Life (EOL) of the bearing system.

Life Phase Evaluation

Test and analysis methods were devised to evaluate each phase of life and compare performance of the standard and subject lubricant. Table 1 presents the tests and analyses identified for this comparison.

TABLE 1—*Phases of life.*

Phase of Life	Method To Address	Test Description	Lube State
Vibration	SRV test	High stress, oscillatory, ball-on-plate test	Solvent/oil dip, reduced grease quantity
Beginning of Life (BOL)	ATP Induced vibration and drag torque data	Extensive review of ATP data from units lubed from standard and subject grease lubricated RWAs	Normal Lubrication quantities & procedures
WAM Boundary test	Low speed ball-on-plate boundary lube test	Solvent/oil dip, reduced grease quantity	
Normal Operating Life (NOL)	WAM EHL test EHL Analysis	Optical ball-on-plate EHL test	Solvent/oil dip, reduced grease quantity
Late Life-End of Life (EOL)	Aerospace EHL test	Stick-out EHL test with reduced grease quantity and no pre-oil to simulate starved lube condition	10% normal grease, no solvent/oil dip
Dynamic Wettability test	Raceway wettability check before and after rotation	solvent/oil dip	

For each of the tests performed, the subject grease or oil extracted from the grease was used and compared with the standard grease or oil. To simulate the initial exposure of the bearing system to random vibration, Schwingung, Reibung, Verschleiß—oscillating, friction, wear (SRV)—testing was performed, which simulates the oscillatory motion of the ball against the raceway using a ball and plate. Boundary lubrication properties for BOL were compared by performing testing on the Wedeven Associates machine (WAM) test machine, which was also used to determine the elastohydrodynamic lubrication (EHL) film thickness for NOL at operating speed. The WAM test machine uses a ball on disk similar to the SRV testing but with a rotating disk. A separate EHL test fixture that had a 304 size bearing was used to measure the film thickness and determine if EOL lubrication could be compromised in the subject grease. Finally, a visual test was performed to determine the wetting ability of bearing surfaces with oil from the subject grease and the ability to rewet the surfaces after bearing operation.

In addition to these tests, acceptance test procedure (ATP) drag torque and induced vibration (IV) data from the RWA's undergoing test was evaluated. This data set was used as an additional figure of merit for comparison to the grease level testing results previously described for BOL. An EHL analysis was also performed for comparison of the test results obtained for NOL operation.

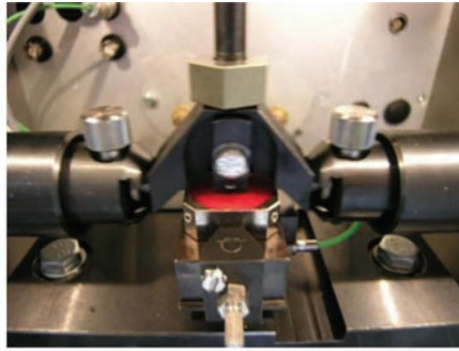
Test Methodology and Results

Vibration Phase

In the normal life cycle of a RWA or CMG, an exposure to random vibration occurs due to launch. During this period, the wheels are not rotating so the balls oscillate in one spot on the raceway with reasonably high loads and many oscillation cycles. The test series addressed the question: Did the presence of the silicone alter the vibration performance of the lubricant?

Nye Lubricants has an SRV tester that can apply different loads to a ball on a plate, oscillate it back and forth across a surface, and detect changes in the coefficient of friction (COF), Figs. 1(a) and 1(b). This test was chosen to simulate the motion between the ball and raceway during random vibration. This test subjects the contact to pure sliding and is considered to be conservative when compared with the motion in the RWA bearing that includes both rolling as well as sliding at the ball/race interface.

A total of 32 tests were run consisting of standard oil and grease and the subject oil and grease. Four test runs of each of these configurations were performed with the 52100 and hybrid bearing materials, i.e., 16 runs for each configuration. Testing was performed at 20 °C (contact disc temperature controlled) at 100 Hz using 0.25 mm stroke of the ball on the disc.



(a)



(b)

FIG. 1—(a) SRV linear reciprocating tribometer, Nye Lubricants. (b) SRV test ball in holder and stationary plate.

For the all tests, consisting of grease and oil, a film was applied to the disk using the appropriate oil-solvent mixture then drained vertically for a controlled period of time before testing or application of the grease.

For the grease tests 0.5 % of the total bearing grease pack (about 10 mg) was applied to the plate in the form of a disc onto which the ball was subsequently placed, Figs. 2(a)–2(c). This is a nominal-to-high range estimate of the amount of lubricant expected to reside in the ball/raceway interface for a bearing that has been operated prior to vibration.

The applied load was incrementally increased by 50 N every 2 min during the test while monitoring the coefficient of friction. A COF of approximately 0.25 was used as the criteria for termination of each test. Once this value was attained the elapsed test time was recorded. Figures 3(a) and 3(b) show plots of the collected test data for the 52100 and CRU/Ceramic systems. It should be noted that the load step profile was the same for each material combination.

The maximum load attained for these tests was 400 N (310 ksi for the 9/16 in. ball on a flat plate) for the 52100 system and 450 N (370 ksi) for the hybrid

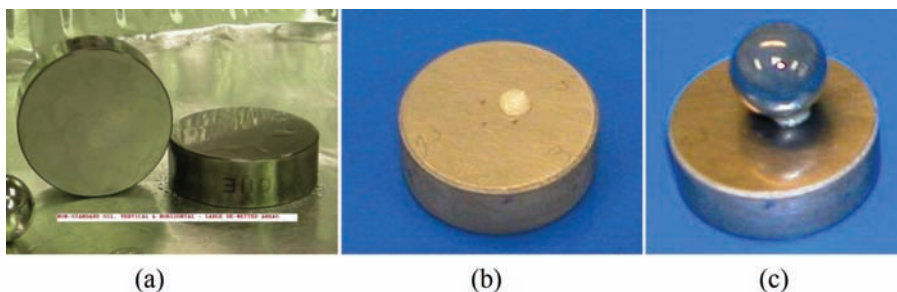


FIG. 2—(a) Disc coated with subject oil: Drained, dewetted surface (left) prior to test with a horizontally drained sample for comparison; (b) standard grease: 10 mg grease applied to disc; and (c) disc shown with ball for sense of scale.

system. Test times ranged from 15 to 3700 s for individual samples, and from 200 to 900 s for the four-sample average. The energy dissipated in the contact during the test is proportional to the area under the load-time test profile.

Table 2 shows the total relative energy at termination, ranked highest to lowest using the standard grease and 52100 combination as the reference. The standard grease ranks highest in the hybrid configuration, followed closely by the standard grease and 52100 material. Overall, the performance ultimately ranked in the order of the original expectations: standard grease, subject grease, standard oil, and subject oil.

The subject grease performance presented one surprise. Considering that the bearing configuration during random vibration does have grease in the ball contact locations, the grease test data is most applicable. The subject grease with the 52100 system had the most repeatable results of all the tests; this consistency was noted as well in an earlier round of preliminary tests, SRV1, at Nye. For the subject grease, the standard deviation was 5 % of the mean energy absorption, versus 35 % of mean energy absorption for the standard grease. This subject grease performance is within the bounds of the performance of the standard grease previously shown in Fig. 3(a). This characteristic is not thought to be related to test system repeatability or lubricant application method, but an interesting feature of the subject grease performance. The relative lifetimes for the subject oil tests were about half that of the standard oil. Both are characterized by high standard deviations which are nearly equal to the mean value of energy absorption for the sample of 4.

To date, we have not established the basis of equivalency between the acceptance test and the SRV test. In the applications using similar 9/16 in. balls, mean and 3 sigma stresses during the 60 s acceptance random vibration are 193 ksi and 278 ksi, respectively, compared to 270 ksi and 310 ksi, respectively, for the average and max stresses experienced on the ball-plate contact during the

SRV ASTM D5706mod - 52100 specimens 100Hz, 0.25mm, 50 N steps, 20°C

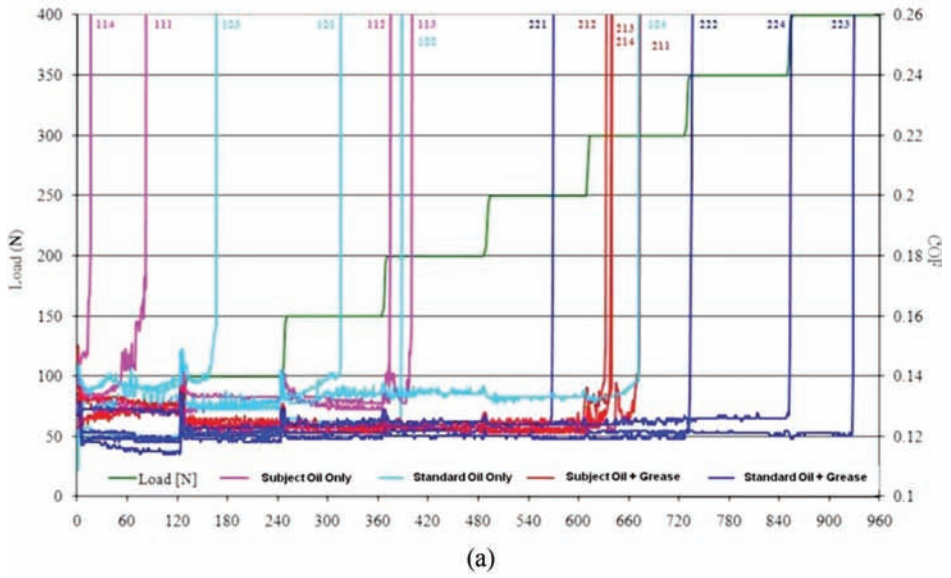
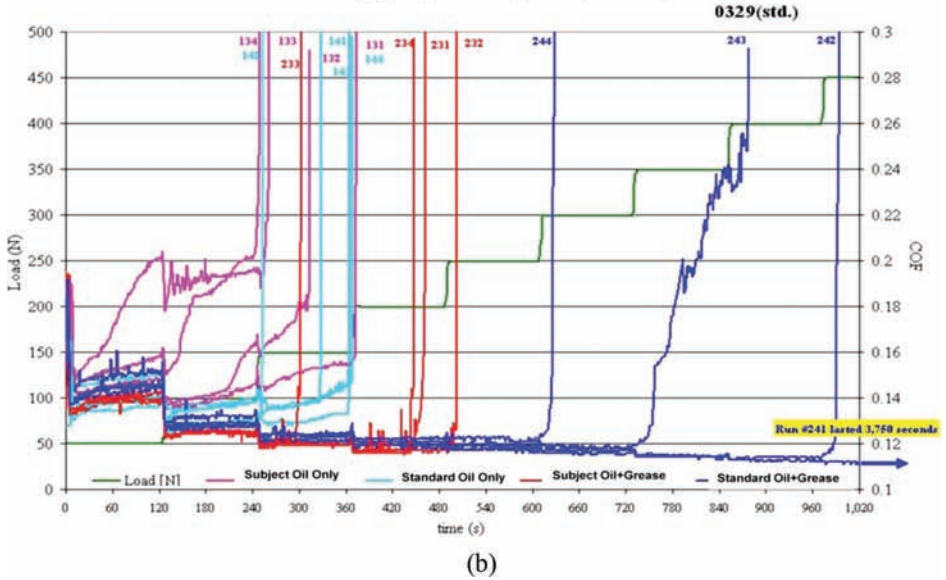
SRV ASTM D5706mod - Si_3N_4 ball, CRU disc, 100Hz, 0.25mm, 50 N steps, 20°C

FIG. 3—(a) SRV2 step load test results (52100). (b) SRV2 step load test results (Cru/ Si_3N_4).

TABLE 2—SRV step load test results.

Sample Configuration	Materials	Ranking Total Energy Dissipated in Contact Prior to Failure: Mean of 4 Samples	Ratio: Normalized to 52100 Standard Grease
		(Watt-s)	
Oil + Grease (standard)	CRU/Si ₃ S ₄	108,299	1.20
	52100	90,007	1.00
Oil + Grease (subject)	52100	63,227	0.70
	CRU/Si ₃ S ₄	30,315	0.34
Standard Oil	52100	31,551	0.35
	CRU/Si ₃ S ₄	22,326	0.25
Subject Oil	CRU/Si ₃ S ₄	21,667	0.24
	52100	14,296	0.16

400–800 s SRV tests where grease is used. For the purpose of judging health of the bearings, we noted the bearing induced vibration (IV) signature tests show no out-of-family spectrums related to the two lubricant populations discussed later in this paper.

Beginning of Life (BOL) Phase Testing

For the BOL evaluation one of the first activities was to evaluate test data from all the units lubricated with the subject lot of grease and compare those results with units lubricated with uncontaminated grease lots. Operating drag torques and bearing disturbance (induced vibration) signatures from hardware built with the subject grease were used to determine if the lubrication system was functioning as expected or if the bearings were potentially damaged.

During the BOL phase, the grease “channels” and runs in during bearing operation. As the bearing runs in, bearing temperatures, drag torques, and emitted disturbances begin to stabilize. Based on bearing models and test data an assessment of bearing health can be made from the drag torque and induced vibration from the bearings.

The ATP drag torque data evaluated was from 38 RWA’s with standard grease and 38 RWA’s with the subject grease (measurements taken twice). Evaluated data included operating speeds of 1000 and 6000 rpm as well as ambient, hot, and cold operating temperatures. The standard grease bearing lot to lot drag torque difference can be greater than 20 %. In comparison, the largest difference seen was 14 % more drag torque with the subject grease than with standard grease (1000 rpm, cold test). The average drag torque increase with subject grease was about 4 % throughout all testing. It was expected that some

drag torque change could occur due to bulk property differences between the two lubricants. The witnessed magnitude of change is either within the expected deviation or it could be a real difference.

If it is a real difference, it is a very small increase compared with the increase that would be expected if the EHL film were deficient. Figures 4(a) to 4(d) show the results of the drag torque comparison at 6000 rpm. Figures 5(a) to 5(d) show the results of the drag torque comparison at 1000 rpm. Each figure shows an assessment of the ratio of both median subject/standard grease drag torque ratio as well as an assessment of the relative data set distribution deviation.

On the basis of the results of the ATP drag torque test data, there is nothing to suggest that the subject lubricant performs significantly different from the standard lubricant for BOL.

Pre and post random vibration IV data was also evaluated for 15 RWA's with the standard grease and 12 RWA's with the subject grease. This data set was collected via accelerometers installed on the unit during testing and provides responses over a spectrum of frequency that was generated by the bearing system and structure.

Specific increments of frequency ($N*f$) are associated with defects located on a specific component. Evaluation of the magnitude of the responses at specific frequencies indicates the bearing surface condition in regard to defects for each bearing component (ball, inner race, outer race). As such, it is the best non-intrusive method to assess the change in surface conditions due to random vibration exposure. The energy of the defects located on a specific component is assumed to be indicated by the root-sum-square (rss) of the magnitudes at the defect frequencies associated with that component within the analyzed frequency range.

Data were taken in a 0 to 2000 Hz frequency range and defect energy was assessed for the inner race, outer race, and ball, Figs. 6(a)–6(c). The blue data points represent RWA's with standard grease. The red data points represent RWA's with subject grease. It is obvious that the two data sets (subject grease and standard grease) are largely coincident with each other.

An additional assessment was made using the six largest peaks (regardless of defect location) in a frequency range of 0 to 10000 Hz, Fig. 6(d). Again, it is obvious that the two data sets (subject grease and standard grease) are largely coincident with each other.

In addition to evaluation of flight unit drag torque and IV data, a test was devised with Wedeven Associates in which their Wedeven Associates Machine (WAM 9) would be used to perform a test that would simulate BOL boundary lubrication of the bearing system, Fig. 7.

This testing was performed at ambient conditions using a 13/16 in. diameter 52100 ball with a surface finish $<1 \mu\text{-in.}$ Roughness average (Ra) and 52100 disc with a surface finish of $2 \mu\text{-in.}$ Ra.

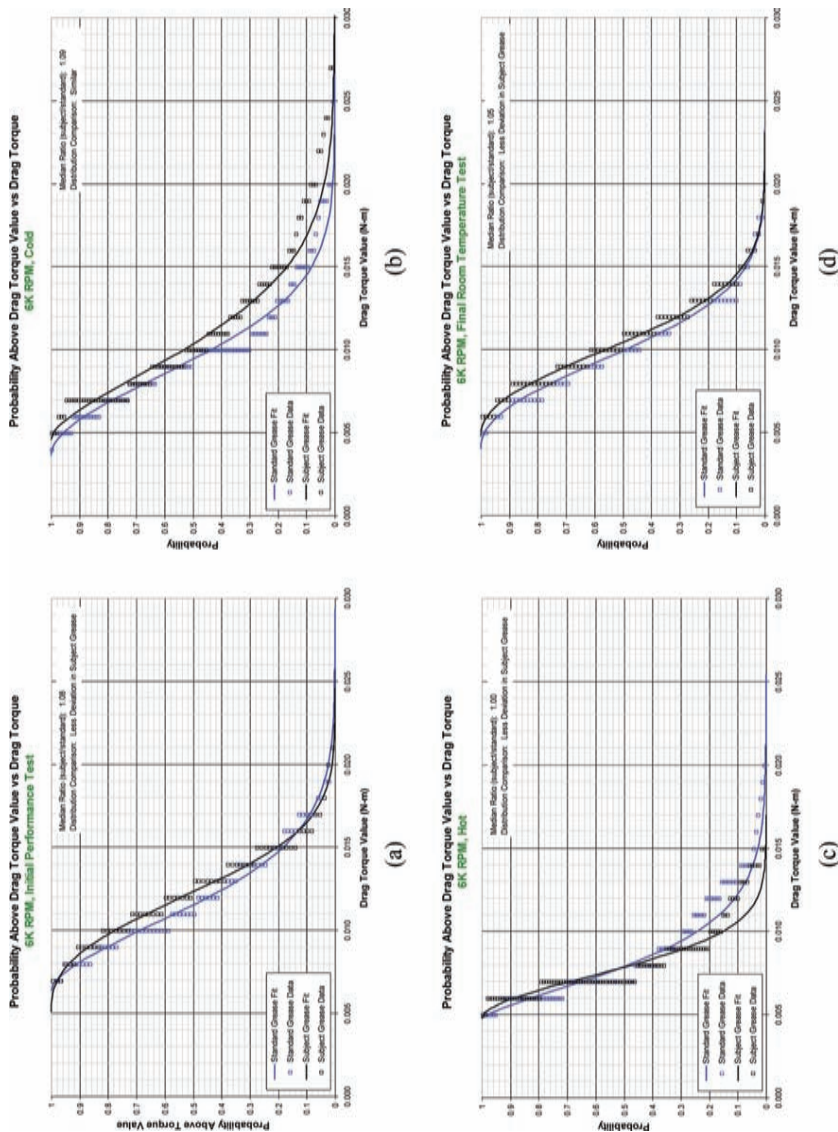


FIG. 4—(a) 6000 rpm initial performance test drag torque, (b) 6000 rpm cold test drag torque, (c) 6000 rpm hot test drag torque, (d) 6000 rpm final performance test drag torque.

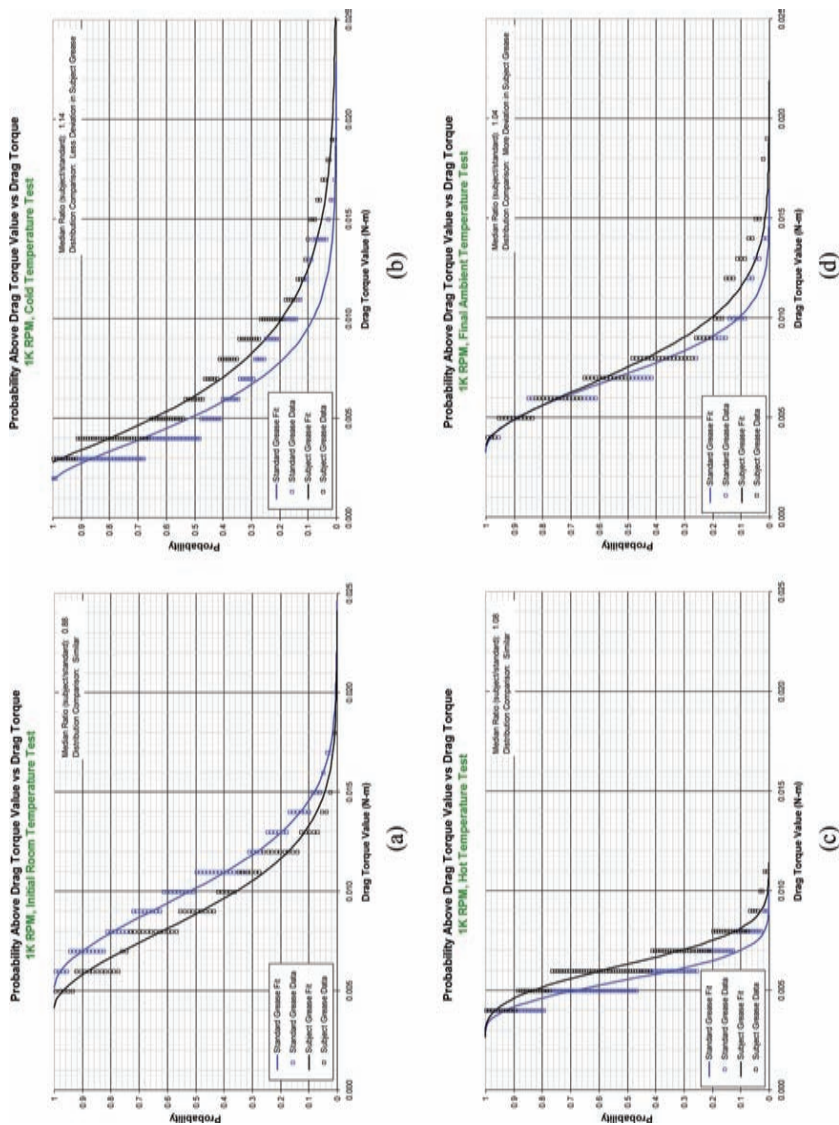


FIG. 5—(a) 1000 rpm initial performance test drag torque. (b) 1000 rpm cold test drag torque. (c) 1000 rpm hot test drag torque. (d) 1000 rpm final performance test drag torque.

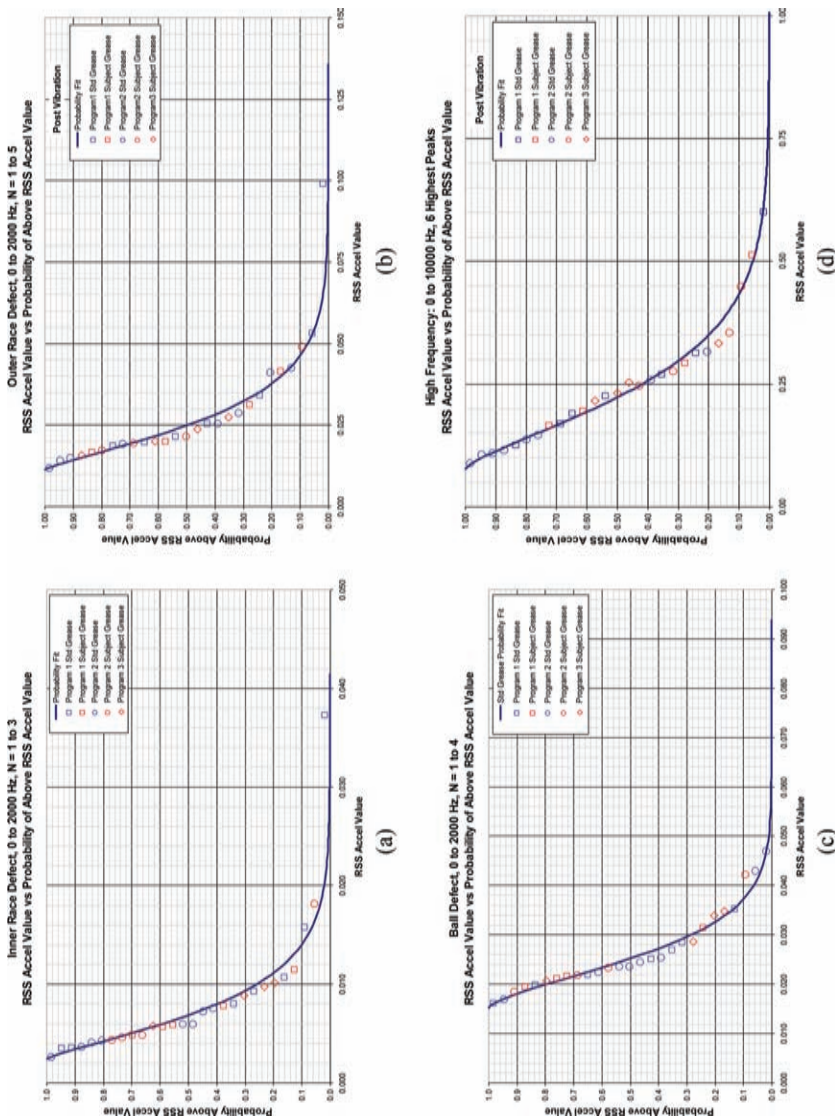


FIG. 6—(a) Inner race defect rss magnitude. (b) Outer race defect rss magnitude. (c) Ball race defect rss magnitude. (d) High frequency rss magnitude.

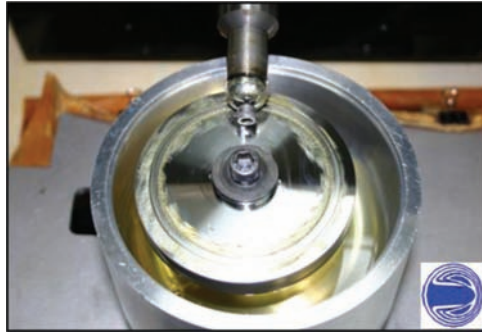


FIG. 7—WAM 9 test machine ball on disc configuration.

For these tests the surface of the ball and the disk were coated with a thin film of virgin Pennzane[®] oil to simulate actual bearing processing. Subsequently, a thin film of the subject grease is placed on the surface of the disk and the testing was initiated with 0 % slip for the first 60 s of testing followed by inducing 2 % slip with a operating contact stress of 80 ksi.

Three test runs of both the subject and standard grease were performed individually. The traction coefficient was monitored during the test and was used as the indicator of test failure and termination, Fig. 8.

The time-to-failure ranged from approximately 1400 to 2400 s for the standard lubricant and 950 to 1650 s for the subject lubricant. Comparing the average time-to-failure for each lubricant to prior data acquired for the standard

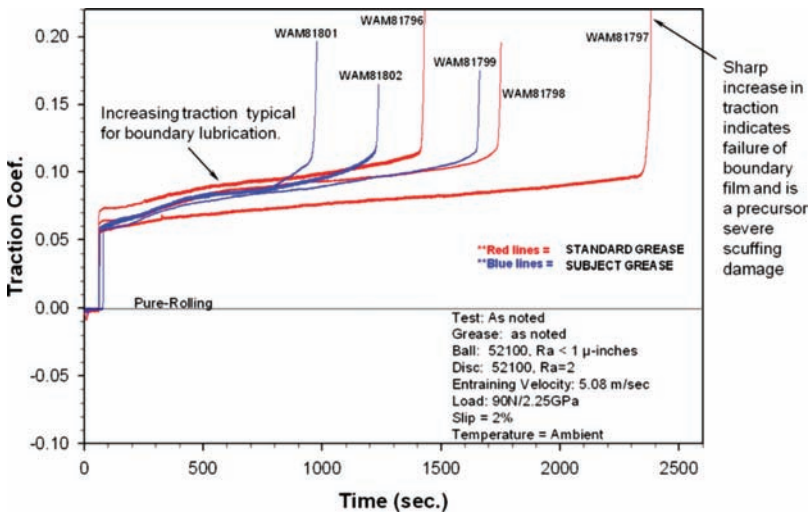


FIG. 8—Traction coefficient versus time to failure for subject and standard grease.

grease in 2005, the subject grease time falls in the middle of the data set, Fig. 9.

These results suggest that the subject lubricant performance is in a family with the heritage standard grease. A verification of this test data was made by comparing the wear scars on the ball and on the plate after termination of the test, Figs. 10(a) and 10(b).

The wear scars on the ball and the disc had similar features and width dimensions regardless if the test was performed with the standard or subject grease. This information further supported the test data that indicated the performance of both lubricants was similar.

Normal Operating Life (NOL) Phase Testing

Normal operating life encompasses the vast majority of the life span. The bearing drag torque is at the lowest level in this time of life. The grease has been channeled during NOL and the bearing is now running on a film of oil that is continuously supplied by the nearby grease pack. The biggest concern during this phase is that a silicone compound may cause dewetting which starves the inlet to the ball-to-race interface and allows direct ball-to-race contact, leading to wear and reduced life.

To evaluate this, Wedeven Associates was engaged to perform elastohydrodynamic lubrication (EHL) testing using their WAM 9 test machine with optical interferometry. This instrument measures the EHL film using interferometry through a semi-transparent rotating Pyrex disk that is coated with a sputtered chromium film upon which the ball freely rolls. Knowing the index

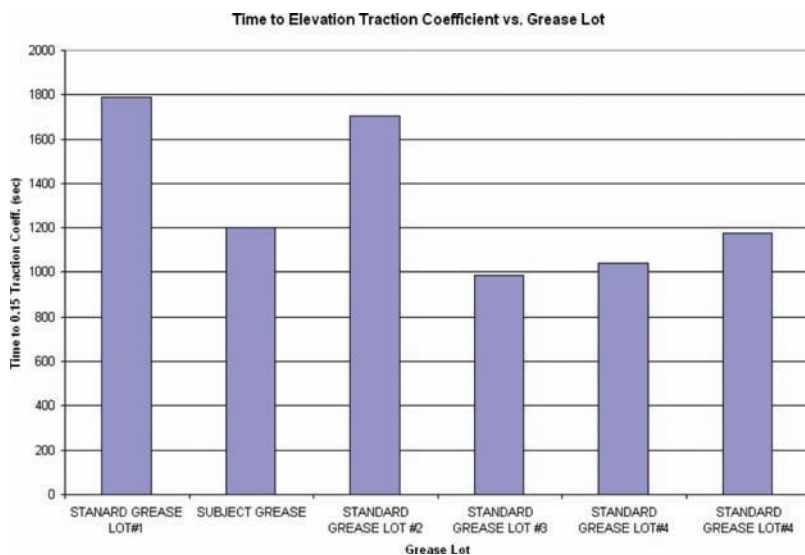


FIG. 9—Average time to failure for subject grease and standard grease.

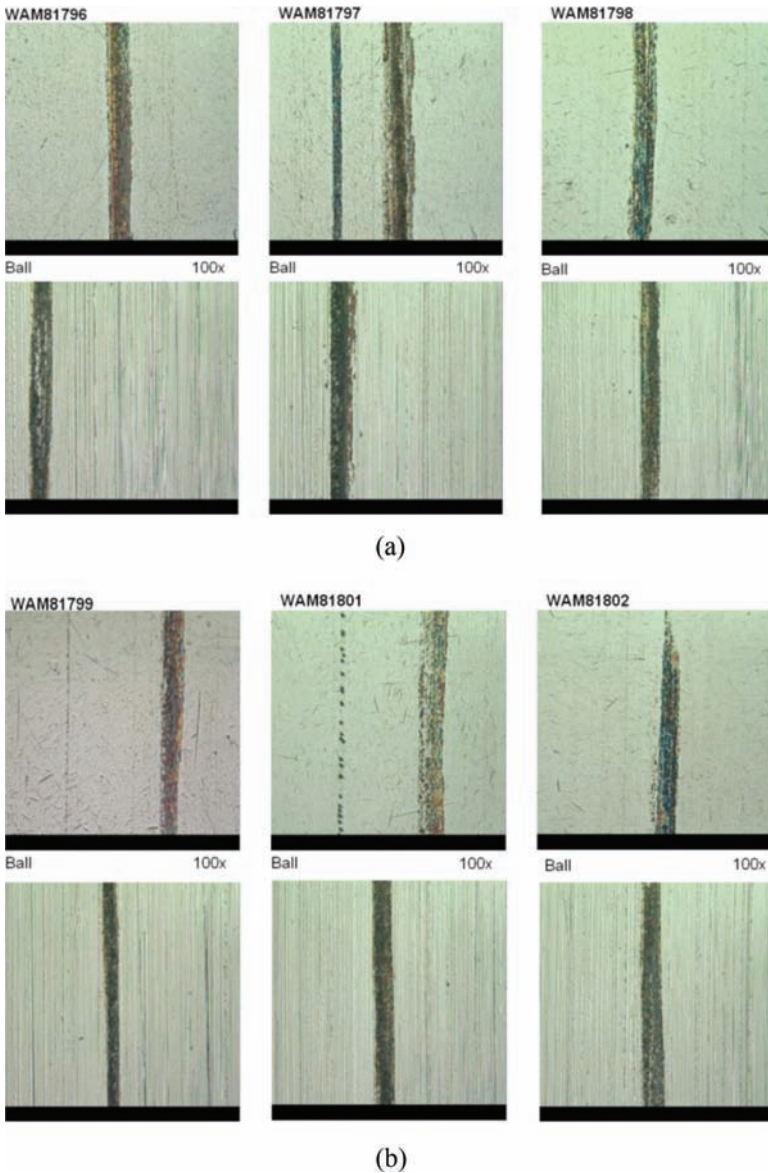
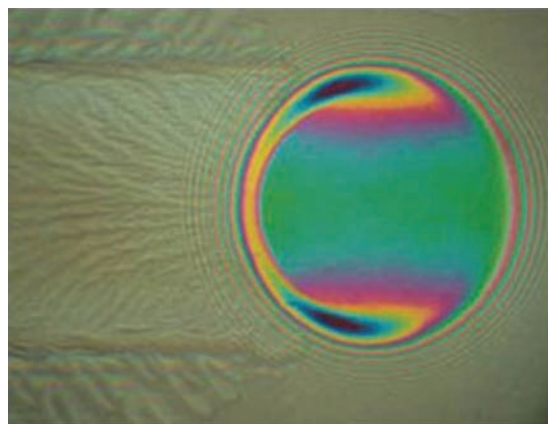


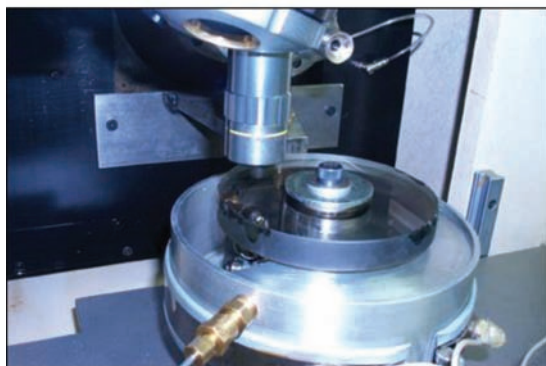
FIG. 10—(a) Ball and disc wear scar standard grease. (b) Ball and disc wear scar subject grease.

of refraction of the lubricant and the number of interference fringes that are observed, an assessment can be made of the operating film thickness between the ball and plate; see Figs. 11(a) and 11(b).

This testing was performed at ambient conditions using a 13/16 in. diameter 52100 ball with a surface finish $<1 \mu\text{-in. Ra}$. For these tests, the surface of



(a)



(b)

FIG. 11—(a) Interferometry photograph of ball running on Pyrex disc. (b) WAM 9 test machine EHL film thickness configuration.

the ball and the disk were coated with a thin film of virgin Pennzane[®] oil to simulate actual bearing processing. Subsequently a thin film of the subject grease was placed on the surface of the disk and testing performed with a 0 % slip condition between the ball and disc, i.e., freely rotating ball driven by the disc. A total of eight different configurations were run of both the subject and standard lubricant operating at entrainment velocities (U_e) of 0.09 m/s, 0.28 m/s, 0.85 m/s, and 2.54 m/s with a Hertzian contact stress of 80 ksi, Figs. 12(a)–12(d). The maximum entrainment velocity was limited by the test equipment capability. For the outer race this range encompasses the entrainment velocity for both the 101 and 304 size bearings at nominal speed. For the inner race the maximum tested entrainment velocity was approximately 70 % for a 101 bearing and approximately 40 % for a 304 bearing at nominal speed.

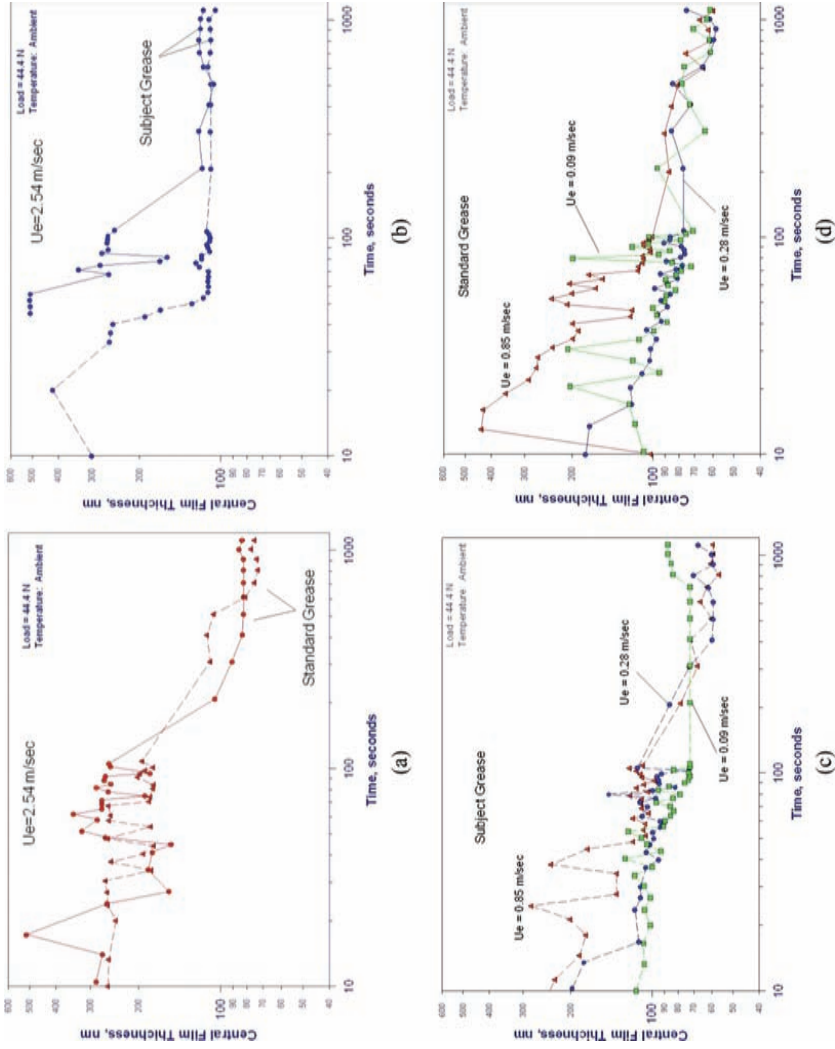


FIG. 12—(a) Standard grease film thickness $U_e = 2.54$ m/s. (b) Subject grease film thickness $U_e = 2.54$ m/s. (c) Subject grease film thickness $U_e = 0.09, 0.28, 0.85$ m/s. (d) Standard grease film thickness $U_e = 0.09, 0.28, 0.85$ m/s.

The data presented show a decrease in film thickness the first 200 to 300 s in which lubricant is channeling and the oil film in the ball to disc contact is stabilizing. For an entrainment velocity of 2.54 m/s, the subject lubricant forms approximately 40 % more film thickness.

Over the entrainment velocity range, the subject grease formed as much or more stabilized film thickness than the standard grease; This data set is summarized in Fig. 13. The subject grease formed a thicker film at the upper and lower end of the entrainment velocity, however, it had a similar film thickness between 0.28 and 0.85 m/s as the standard lubricant.

The results of these tests show that the subject lubricant will form a similar or better film during NOL of the bearing system.

End of Life (EOL) Phase

If the film of standard oil, initially applied to the bearing, is gradually replaced by oil containing the silicone, eventually the oil lubricating the ball-race interface has roughly the same concentration of silicone as the oil in the grease. This can be thought of as near end of life and would be the worst case condition for dewetting potential, resulting in EHL breakdown and ball-to-race contact.

With support of The Aerospace Corporation, an EHL film thickness test was performed with the standard and subject grease with a special test fixture, Fig. 14. The fixture used a 304 size bearing that was lubricated with a known

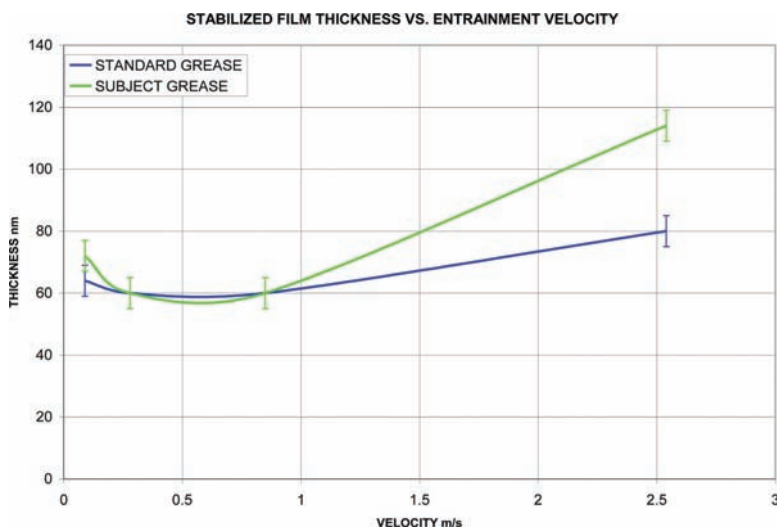


FIG. 13—*Stabilized EHL film thickness versus entrainment velocity summary.*

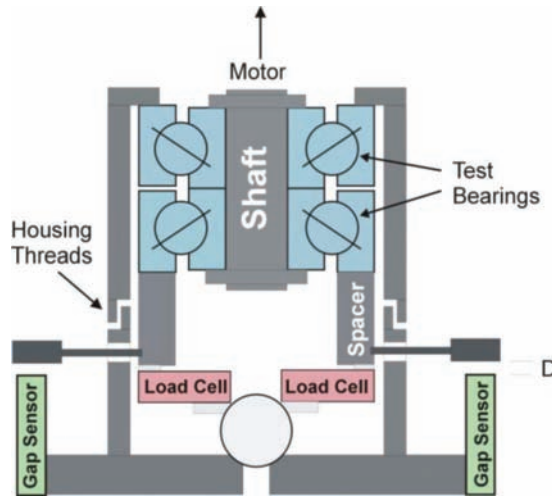


FIG. 14—The Aerospace Corporation EHL tester.

amount of lubricant less than the initial grease pack installed in the bearing at beginning of life.

The reason for this was to simulate a limited oil supply from the grease. The bearing preload was set and monitored using a load cell to ensure the operating preload was in the range of interest. Additionally, gap sensors were used to determine the amount of deflection that occurs from zero to operating speed. Knowing the contact angle of the bearing and the deflection due to operation, the lubricant film thickness could be determined. Testing was performed at a speed of 6000 rpm and an ambient lab temperature of 75 °F. Tests on both the standard and subject lubricant were run up to 1900 h, Fig. 15.

These tests are still in progress at the writing of this paper. The film thickness measurements in tests stabilized between 400 and 600 h. For the subject lubricant, the average film thickness was approximately 20 μ -in., which is approximately twice the value of the data obtained to date for the standard lubricant. The results of this data set also support the information from the WAM boundary testing that the film thickness of the subject lubricant is as good or better than the standard lubricant.

Wettability

There was concern that the silicone would cause permanent dewetting of the bearing contact surfaces. Static wettability tests showed that an oil sample extracted from the subject grease dewet bearing surfaces that were initially wet, but the question of whether operation of the bearing after dewetting of the surface reversed the condition remained.

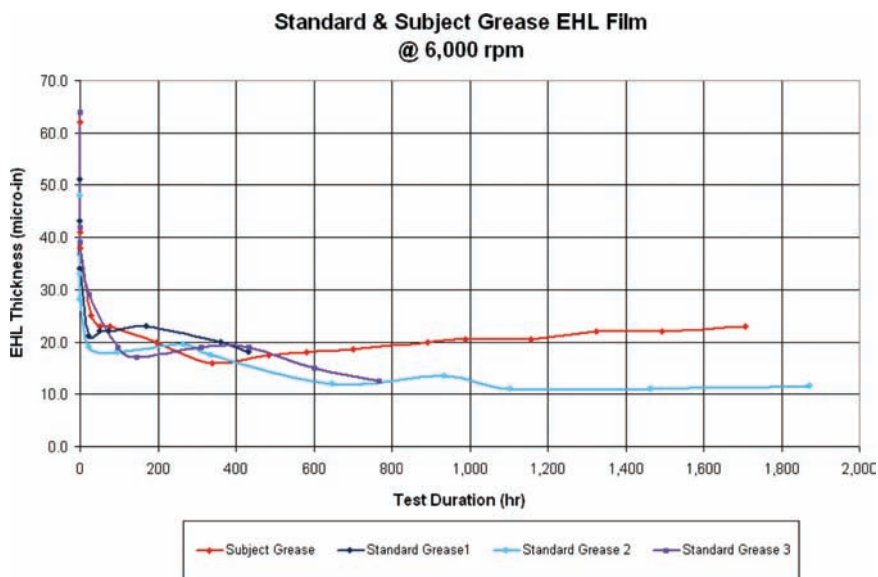


FIG. 15—EHL film thickness of subject and standard grease on Aerospace test fixture.

A test was set up using a 305 size bearing which has a raceway running contact area that can be easily seen to ascertain wettability. First the bearing was cleaned and dipped in a solvent to standard oil mixture which did not show dewetting of the raceway as expected, Fig. 16.

Following this, a small quantity of oil extracted from the subject grease was applied to the running surface of the rings, which resulted in dewetting. The bearing was then rotated by hand 50 times and the contact surfaces were observed. It was found that the surfaces that were initially dewet from the silicone oil had become wet once again immediately and after a specified elapsed time in hours.

These results were contrary to the initial thought that once the surfaces were dewet by a contaminant, they would not wet again. This data set provided additional confidence that the presence of silicone in the lubricant at the levels noted would not result in permanent dewetting of the critical contact surfaces of the bearing.

Lube Life Model

Before we could select test methods for evaluating each life phase we had to have a mental model of how bearing lubrication works to help us define how a new additive may affect performance during each phase. Figure 17 shows nomenclature regarding EHL film generation. The following model was developed based on both the standard Hamrock/Dowson fully flooded EHL film

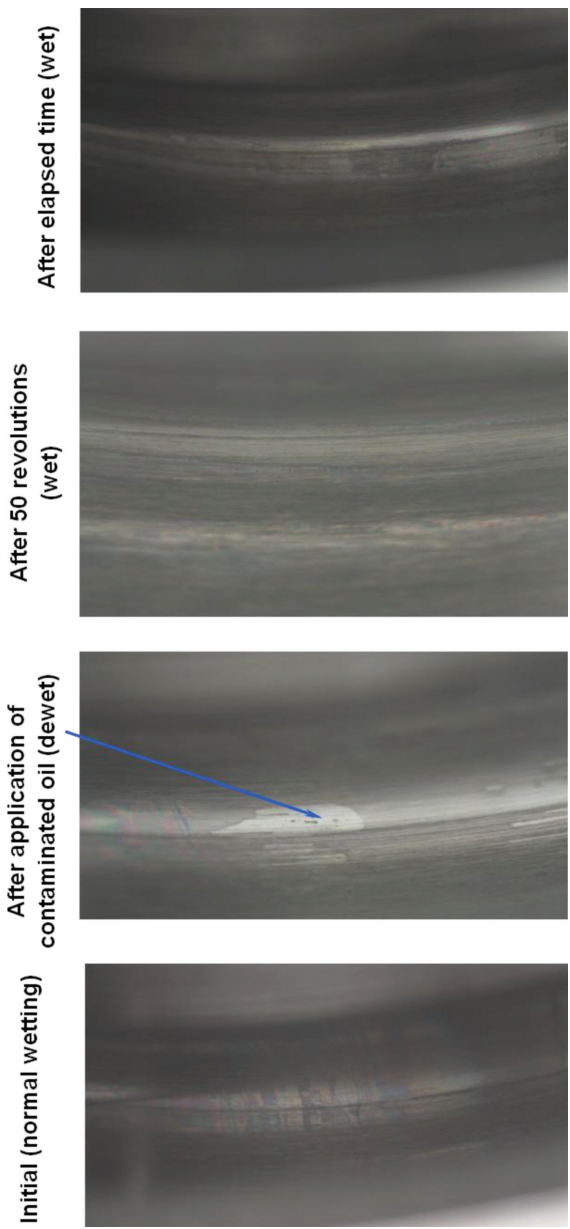


FIG. 16—Wettability of bearing ring before and after dynamic motion.

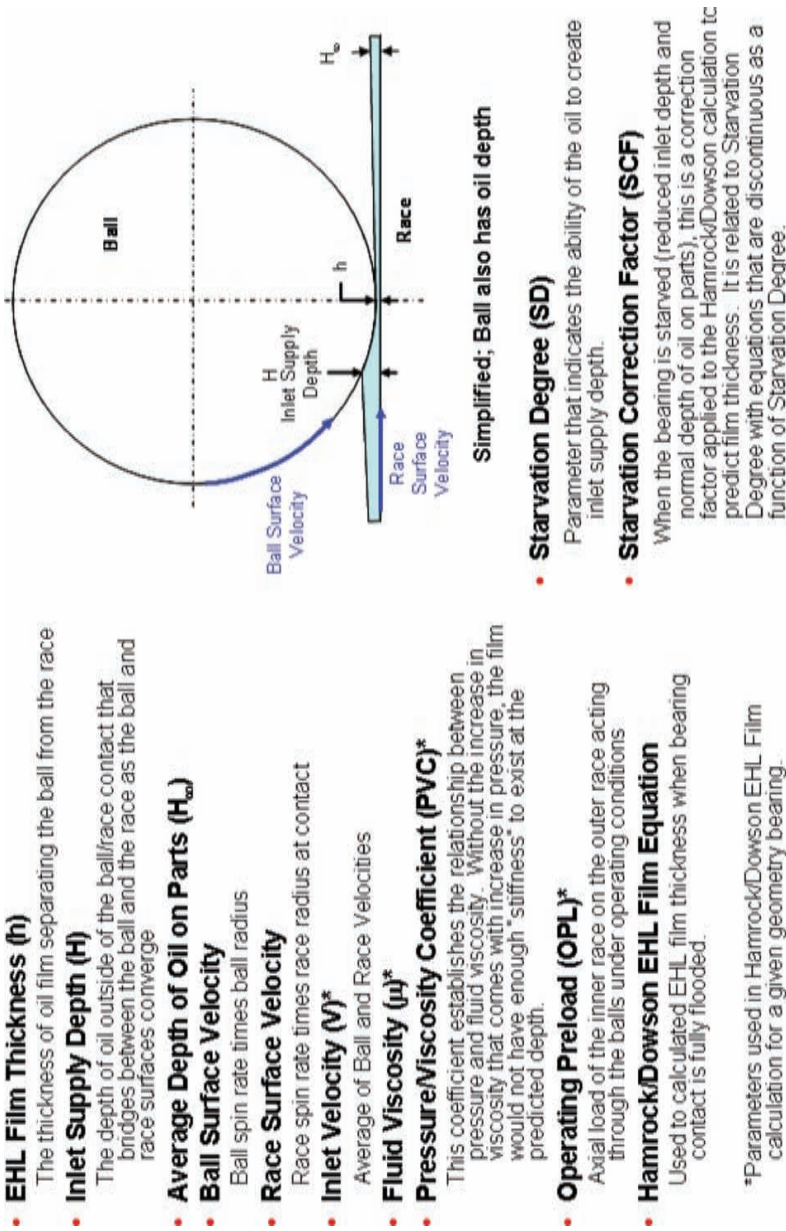


FIG. 17—EHL film generation nomenclature.

model [1] and testing performed by Cann, Damines, and Lubrecht [2] that demonstrated a relationship between specific bearing operating parameters and a starvation correction factor that is applied to the Hamrock/Dowson calculation to calculate EHL film thickness under starved inlet supply conditions.

The EHL film thickness that is generated when the depth of oil on parts is not limited by oil supply is predicted by multiplying the fully flooded Hamrock/Dowson EHL film equation result by the appropriate starvation correction factor.

The starvation correction factor is a function of starvation degree. The general equation for starvation degree is:

$$\text{Starvation degree (SD)} = C * \mu * V * \text{OPL}^{1/3} / f(H_{\infty})$$

As shown in the Fig. 17, H_{∞} is the average oil depth on parts (ball, race) away from the contact. When the oil supply rate to the bearing components is adequate, the oil depth on parts is relatively constant at a given temperature and speed because the loss mechanisms are a function of depth. Because, in a properly greased spin bearing, the oil supply rate is adequate to provide an oil depth that creates a constant film thickness for multiples of the life requirement, this is the oil depth condition that the bearing runs in for most of its useful life (NOL, normal operating life). The C value in the above equation is chosen with test data taken under these NOL conditions and the H_{∞} term is removed from the equation. As such, it is an empirically derived constant for a given bearing system.

There are normally two time periods that a bearing could run with less than steady state oil depth on parts.

- If inadequate free oil is applied to the bearing at Beginning of Life (BOL), it will take some time (perhaps thousands of hours) for the oil depth on the parts to increase to the steady state value because the flow rate of oil from grease is relatively small. Honeywell normally provides free oil to the bearing parts with a solvent/oil solution dip prior to final grease lubrication.
- As End of Life (EOL) is approached, the supply rate becomes too small to maintain the oil depth on parts needed to create the Normal Operating Life (NOL) film thickness. Once the depth of oil is no longer adequate to provide an inlet depth necessary to create a sufficient film, metal to metal contact occurs and drag torque increases. Bearing failure of high speed bearings, based on operational experience, follows a few years later.

As previously mentioned, the starvation degree equation has been calibrated for Honeywell bearings based on test data (capacitance measurement, axial deflection, thermal conductance, temperature difference between inner and outer race) in the NOL condition. In addition, the relationship between starvation degree and starvation correction factor has been determined using

TABLE 3—*Starvation correction factor vs starvation degree.*

Regime	Regime Description	SD Range	SCF Equation	SCF Range
Flooded	Film Thickness Increasing with SD Increase	0–1.4	1.0	1.0
Flooded to Starved	Film Thickness Increasing with SD Increase	1.4–2.28	$1.239 - 0.1705 * SD$	1.0–0.85
Starved	Film Thickness Decreasing with SD Increase	2.28–6.0	$(2.225 - 0.323 * SD) / (SD^{0.68})$	0.85–0.085
Starved	Film Thickness Constant	>6.0	$0.288 / (SD^{0.68})$	<0.085

data generated by others [1,2]. This relationship is non-linear and discontinuous. Table 3 shows the relationship.

The operational regime with most risk of not creating adequate oil film is the starved regime described as Film Thickness Constant. This is the regime in which the inlet depth is approximately equal to the film thickness. In other words, the “shape” of the oil on the parts is a constant, not having adequate time to move before next component pass.

There are 3 potential deleterious effects of the silicone oil.

- The steady state oil depth is less with the subject grease, due to higher oil loss rate and/or lower supply rate at a given speed and temperature.
- The surfaces de-wet resulting in very little oil depth on parts.
- The silicone/Pennzane[®] oil mixture properties are inadequate to produce an adequate EHL film thickness.

The worst case test for the first potential effect is to operate a bearing with no free oil (low oil depth on parts) in the worst starved regime discussed above. This can simulate an EOL condition (low oil depth on parts). This test was run in the Aerospace EHL test fixture. A description of the results of that testing is shown on Fig. 15. The subject grease created a film thickness similar to that of the standard grease at beginning of the highly starved test when oil depth on parts was at a minimum.

The test that refutes the second potential effect is the Dynamic Wettability Test.

The Aerospace EHL testing and the WAM EHL testing both refute the third potential effect. In fact, the Subject grease created somewhat more film thickness than the standard grease.

Test and Analysis Results Summary

Several tests were performed in this investigation that showed the level of silicone present in the lubricant was not of concern. Data from these tests were validated by evaluation of flight unit data as well as performing analysis. The following summarizes the results of these activities:

The SRV testing revealed two characteristics of the subject grease that are of interest to the comparative application of the standard and subject grease to the RWA/CMG application. First, the minimum lifetime for both the standard and subject grease are similar because, though the subject grease has a 20 % lower mean life it also has a tighter life scatter. This characteristic was observed on the original SRV test series. Second, both the standard and subject oil-only relative lifetimes are a fraction, roughly 30 %, of the grease lifetimes, with the subject oil-only exhibiting greater scatter, a possible area of interest. To judge the impact of the scatter we looked at the RWA BOL spin bearing noise spectrums to judge the presence of bearing impact and found none to be present.

Boundary testing with the WAM test machine showed a time of 1220 s for the subject grease to attain a traction coefficient of 0.15; this value compared to the average time of 1225 s for prior tests performed on standard grease samples. The latest standard lubricant average time to failure was 1780 s. Overall, comparing the test results the standard lubricant time to failure enveloped the non standard lubricant time. This data set is also validated by the visual inspection of the wear scars on the balls after completion of each test, all of which look very similar.

EHL testing using the WAM machine indicates the initial film thickness of both grease types is elevated at the beginning of the test. It is suspected that the additional film thickness during this time may be due to the early life presence of grease thickener in the EHD contact. Eventually, the thickener is expelled from the contact and film thickness becomes a function of oil properties. Both grease types show a steady EHD film after approximately 200 s from test start. For an entrainment velocity of 2.5 m/s the standard grease had ~28 % lower film thickness (~80 nm) than the subject grease. For entrainment velocities of 0.85, 0.28, and 0.09 m/s both lubricants had film thicknesses of ~60 nm after 1000 s. The data collected also suggest that the higher the entrainment velocity the thicker the initial film thickness for the first 100 s. Overall, both lubricants have similar film thickness until the entrainment velocity reaches 1.5 m/s, at which the film thickness of the subject grease is greater than the standard lubricant.

EHL testing performed using the Aerospace test machine shows that the film thickness of both lubricants may be close to stabilization at approximately 800–1000 h. It is possible that it would further increase with time as the depth of oil on parts increases if the steady state depth had not been reached. The standard lubricant, at test termination, has a film thickness of 10–12 $\mu\text{in.}$, which compares to 20 $\mu\text{in.}$ for the subject grease at test termination.

Static wettability tests show than an initially wet surface can be dewetted from silicone oil in the levels experienced. Subsequent operation of the bearing with the dewetted surfaces present demonstrates that mechanical action causes surfaces to become wet once again. It is certainly plausible that the level of the silicone present can affect the outcome of this test. If higher levels of silicone

were present, the threshold at which this would occur is unknown and further testing would be required.

Review of RWA drag torque and post-random vibration induced vibration data reveals that there is no significant difference between the bearings with subject grease and bearings with standard grease.

Some customers had the ability to replace the spin bearings on their hardware and have chosen to do so. To date one customer has chosen to fly three of their RWA's with the subject lubricant based on the results of this test data. In addition to this there are three additional units that have been in operation prior to findings of this investigation and have been functioning normally for over 2.0 years.

Conclusions

Several tests were undertaken to simulate the life of a bearing and the conditions it is exposed to over this timeframe to determine if the subject lubricant could have deleterious effects on performance or meeting life. All of the testing that was performed in this investigation shows that the subject grease with the particular silicone contaminant in the quantities noted performed as good or in some cases, better than the standard grease. This has also been validated by on-orbit data from three units that have been in operation for 2.0 years and are functioning normally.

Acknowledgments

The writers gratefully acknowledge support of this work and the collaborative efforts of Nye Lubricants Inc., Fairhaven, MA; Wedeven Associates Inc. Edgemont, PA; and the Aerospace Corporation, El Segundo, CA.

References

- [1] Hamrock, "Elastohydrodynamic Lubrication-Status of Understanding", *Status and New Directions in EHL*.
- [2] Cann, Damines, and Lubrecht, "The Transition Between Fully Flooded and Starved Regimes in EHL," May 2004.



www.astm.org

ISBN: 978-0-8031-7528-0

Stock #: STP1542