

UNCLASSIFIED

AD NUMBER
AD832929
NEW LIMITATION CHANGE
TO Approved for public release, distribution unlimited
FROM .Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; Apr 1968. Other requests shall be referred to RADC [EMEAM], Griffis AFB, NY 13440.
AUTHORITY
RADC ltr dtd 17 Sep 1971

THIS PAGE IS UNCLASSIFIED

AD 832 929

DEVELOPMENT OF A MATHEMATICAL MODEL FOR
PREDICTING LIFE OF ROLLING BEARINGS

Y. P. Chiu
J. A. Martin
J. I. McCool
et al

SKF Industries Incorporated

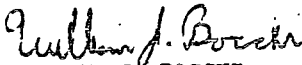
This document is subject to special
export controls and each transmittal
to foreign governments, foreign na-
tionals or representatives thereto may
be made only with prior approval of
RADC (EMEAM), GAFB, N.Y. 13440.

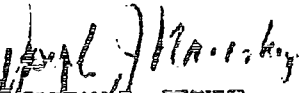
FOREWORD

This report was prepared by the Research Laboratory of  ESF Industries, Inc., King of Prussia, Pennsylvania, under USAF Contract No. F30602-67-C-0117, which was initiated under Project No. 5519, Task No. 551902. The work was administered under the direction of the Development Engineering Branch, Engineering Division, Rome Air Development Center, with Mr. W. J. Bocchi acting as project engineer.

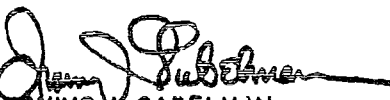
The following  ESF Industries, Inc. personnel contributed to this project, in addition to the authors: Dr. J. Y. Liu and Mr. F. R. Morrison. The project was under the supervision of Mr. O. G. Gustafsson and the direction of Mr. T. E. Tallian. This report covers research work conducted from December 22, 1966 through December 22, 1967. The secondary report designation is  ESF Report No. AL68P003.

This technical report has been reviewed and is approved.

Approved: 
WILLIAM J. BOCCHI
Mechanical Engineering Section
Development Engineering Branch

Approved: 
WILLIAM F. BETHKE
Chief, Engineering Division

FOR THE COMMANDER


IRVING J. GABELMAN
Chief, Advanced Studies Group

ABSTRACT

A description of rolling contact failure modes is given and the variables affecting the life of a rolling contact are identified. A mathematical model of subsurface and surface crack propagation is presented. The life to failure of volume elements in the vicinity of a defect is formulated. A term "severity" of a microdefect has been defined. The model is characterized by the inclusion of bulk material parameters, defect characteristics and parameters of geometry, stress, lubrication and surface topography. A statistical expression for the life of an entire rolling body is based on the defect life formula. The new model includes current standard bearing life prediction formulas as a special case. To assist in interpretation of the model, the stressed volume in a Hertzian elliptical stress field has been determined from the computed contours of equal reversing shear stress. A stress analysis has been conducted on the stresses near interacting asperities and around a surface defect (furrow).

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
I INTRODUCTION AND SUMMARY.....	1
II PRINCIPAL CONCEPTS.....	5
1. The Principle of Rolling Bearing Life Prediction.....	5
2. Principal Variables of Fatigue Failure.....	8
3. Fatigue Failure Model.....	9
4. The Matrix Factor.....	13
5. The Defect Severity Factor.....	15
6. Prediction of the Life at a Defect.....	17
7. The Statistics of Life for an Entire Rolling Body.....	19
8. Application of Life Equations.....	22
a. The Lundberg-Palmgren Case.....	22
b. Deviation from the Lundberg-Palmgren Type Weibull Distribution.....	23
c. Effect of Material Cleanliness (Inclusion Content).....	23
d. Two Competing Failure Modes.....	24
e. The Effects of Residual Compressive Stresses in Material.....	25
f. Effect of Hardness.....	25
g. Non-Hertzian Contacts.....	25
h. Lubrication Effects.....	26
i. Size Effect.....	26
9. The Stress-Strain Relationship.....	27
a. Contour of Equal Shear Stress Range in Elliptical Hertzian Contact.....	28
b. Near Surface Stresses Due to Asperities.....	28
c. Stresses Under Surface Imperfection.....	30
d. Plastic Strain Determination.....	30
10. Outlook.....	31
III SYNOPSIS OF LUNDBERG-PALMGREN THEORY.....	33
1. Failure Probability Distribution.....	33
2. Equivalent Load.....	39

TABLE OF CONTENTS (CONT)

<u>Section</u>	<u>Page</u>
3. Capacity of a Complete Bearing.....	40
4. Determination of Constants in the Life Formula..	
a. Point Contact.....	41
b. Line Contact.....	41
5. Factors Omitted from the Lundberg-Palmgren Theory.....	42
IV VARIABLES AND MECHANISMS OF ROLLING CONTACT FATIGUE.	43
1. Failure Mechanisms (General).....	43
2. Variables Affecting Fatigue in Rolling Contact..	49
a. Material Variables.....	49
b. Surface Microgeometry Variable.....	49
c. Design Variables Related to Design Dimensions	51
d. Operating Variables.....	51
3. Mechanisms of Failure in Rolling Contact.....	51
a. Subsurface Failure.....	52
b. Surface Distress and Surface Initiated Fatigue.....	53
4. Failure Process Diagram (Rolling Contact).....	55
V FORMULAS FOR FATIGUE CRACK GROWTH.....	57
VI STATISTICAL THEORY OF ROLLING ELEMENT FAILURE.....	65
1. General.....	65
2. Defect Severity Distribution.....	67
3. Distribution of "Defect Life".....	68
4. Asymptotic Distribution of Smallest Defect Life.	70
VII MACROSCOPIC STRESS CALCULATION FOR THE HIGHLY STRESSED VOLUME IN A HERTZ CONTACT.....	73
VIII DETERMINATION OF SHEAR STRESS NEAR ASPERITIES.....	87

TABLE OF CONTENTS (CONT)

<u>Section</u>	<u>Page</u>
IX DETERMINATION OF SHEAR STRESS BENEATH A FURROW.....	105
1. Computation of Contact Pressure Near a Surface Defect.....	108
2. Sub-surface Stress Distribution in the Vicinity of a Surface Defect.....	112
<u>Appendix</u>	
I Formulas for Stresses in a Hertzian Stress Field.....	117
II Formulas for Stresses Corresponding to an Infinitely Narrow Contact Ellipse in a Hertzian Stress Field....	121
III Plane Contact of Asperities.....	125
IV Compression of a Half Plane Containing an Idealized.. Surface Defect.....	129
V Average Shear Range in the Stressed Area Enclosed by a Contour of Equal Shear Range in a Hertzian Elliptical Contact.....	131
VI Plausible Defect Severity Distributions.....	133
<u>References</u>	139

LIST OF ILLUSTRATIONS

<u>Figure</u>		<u>Page</u>
1	The Interdependence of Variables Affecting Contact Fatigue Life	54
2	Schematic Representation of Growth of Micro-cracks	61
3	Equilibrium of Stresser Acting in y-Direction	74
4	Variation of τ_R and θ_1 with x and z	76
5	Typical Variation of Orthogonal Shear Stresses with y for given x and z.	77
6	Contours of Equal Shear Range τ_R in a Hertzian Elliptical Contact ($a/b = 10$)	79
7	Contours of Equal Shear Range τ_R in a Hertzian Elliptical Contact ($a/b = \infty$)	80
8	Contours of Equal Shear Range τ_R in a Hertzian Elliptical Contact ($a/b = 7.5$)	81
9	Variation of S/ab with a/b and τ_R/p_{max} .	84
10	Variation of S/az_0 with a/b , $\tau_R/2\tau_0$ and $2\tau_0/\tau_R$	85
11	Asperity Slope Distribution on Ground, Honed, and Lapped Surfaces	88
12	The Contact of an Idealized Surface Asperity	89
13	Contours of Equal Octahedral Shear Stress in an Idealized Asperity Contact	91
14	Variation of $(\tau_{45^\circ})_{max}$ and its Depth y_0 with b/a for the Simple Asperity Model	92

LIST OF ILLUSTRATIONS (CONT'D)

<u>Figure</u>		<u>Page</u>
15	A Simple Model of Asperity Contact	94
16	Variation of Surface Deformation Outside the Contact Zone of the Simple Asperity Model	95
17	Variation of $\tau_{\max}$ and a/b with h/σ for a Ground Surface	99
18	Variation of $\tau_{\max}$ and a/b with h/σ for a Honed Surface	100
19	Variation of $\tau_{\max}$ and a/b with h/σ for a Lapped Surface	101
20	Variation of $\tau_{\max}$ with a/b	103
21	Schematic Representation of an Idealized Surface Defect	104
22	Schematic Representation of Contacting Bodies Containing Surface Defects (Limiting Case $R \rightarrow \infty$)	106
23	Variation of ω ($= c/b$ or defect width/distance between two contact edges) with C_0	107
24	Pressure Distribution Near a Surface Defect (for $c/b = 1.2$)	109
25	Variation of Maximum Contact Pressure and its Location Coordinate with Dimensionless Parameter c/b for an Idealized Defect	110
26	Variation of Maximum Contact Pressure as a function of Defect Geometric Parameter r/c and Nominal Pressure p_0	111
27	Application of Method of Superposition for Numerical Integration to Obtain Sub-surface Stress Distribution under a Surface Defect	113

LIST OF ILLUSTRATIONS (CONT'D)

<u>Figure</u>		<u>Page</u>
28	Contours of Equal von-Mises Yield Criterion in Material around an Idealized Surface Defect	114
29	Coordinate System and Defect Cells	66
30	Cumulative Distribution Function of Defect Severity with $d = 2.0$.	134
31	Probability Density Function of Defect Severity with $d = 2.0$	135
32	Variation of Average Shear Range in a Contour of Equal τ_R with Enclosed Area S	132
33	Plastic, Elastic and total Strain vs. Fatigue Life	47

LIST OF TABLES

<u>Table</u>		<u>Page</u>
I	Failure Modes of Rolling Con acts	44
II	External Variables Controlling Contact Fatigue Life	50
III	Variation of z_0/b and $r_0/p_{\max}$ with a/b in Hertzian Contact	83

NOMENCLATURE

A	Instantaneous crack area (in^2); ratio of major to minor semi-axes of contact ellipses
A_c	Crack area at end of Phase II life (in^2)
A_o, A_f	Initial and final areas of fracture cross section (in^2)
A_p	Crack area at completion of Phase I life (in^2)
a	Major semi-axis of contact ellipse; semi-width of contact region of asperity; semi-width of contact region of idealized surface defect (in)
a_i	Constants
B	A function of deterministic variables
b	Minor semi-axis of contact ellipse; semi-width of curved base of asperity; semi-width of free surface of idealized surface defect (in.)
C	Dynamic capacity (lb)
C_o	Dimensionless parameter
c	Semi-width of idealized surface defect (in); parameter of Pareto distribution; constant exponent
D	Ductility
D_a	Rolling body diameter
D_n	Raceway diameter
d	Defect severity
$\bar{d}$	Average defect severity
d_m	Bearing pitch diameter
d_o	Exponential distribution parameter

NOMENCLATURE (CONT'D)

E	Young's modulus (lb/in ²); elliptic integral of the second kind
E'	Reduced Young's modulus (lb/in ²)
e, (k)	Weibull slope
F	Elliptic integral of the first kind; bearing load
F(d)	Defect severity distribution
F(x)	General probability distribution
F ₁	Distribution of smallest sample value
f	Function; surface profile coordinate
G	Life distribution
G(λ)	Material function
g	Constant exponent
H	Rolling body life distribution
h	Minimum elastohydrodynamic film thickness (in); exponent
I	Function
J	Function
j	Exponent
K	Complete elliptic integral of the second kind; a function
k, (e)	Weibull shape parameter
L	Life; dimensionless elliptic coordinate; constant intercept
L ₁₀	90% reliable life

NOHENCLATURE (CONT'D)

l_{a, l_0}	Length of rolling track (in)
M	Material variables; fatigue ductility coefficient
M_x, M_y, M_z	Functions
m	Number of defect cells in a rolling body
N	Number of stress cycles; function
N_L	Number of cycles to failure
N_0	Shortest Phase I life
N_x, N_y	Functions
N^*	Weibull scale parameter
P	Total load (lb); function
p	Contact pressure (lb/in ²); exponent in load-life relationship
$P_c, (\sigma_0)$	Maximum Hertzian contact pressure (lb/in ²)
P_{max}	Maximum contact pressure due to a surface defect (lb/in ²)
Q	Rolling body load; function
Q_c	Rolling body load for a 90% reliable life of 10^6 ring revolutions
R	Radius of curved tip of asperity; radius of curved contact surface (in.)
r	Idealized surface defect shoulder radius (in.)
S	Size of highly stressed area in the cross section of a rolling element (in ²); asperity spacing (in)

NOMENCLATURE (CONT'D)

u	Contact cycles per revolution
V	Stressed volume (in^3)
v	Deformed profile of elastic half plane in contact with rigid asperity (in)
W	Load
w	Exponent
X, Y, Z	Dimensionless coordinates
x, y, z	Coordinate axes
X	Random variable
z	Fatigue ductility exponent
α	An angle
β, β'	Constants
γ	Matrix function
Γ	Function relating Phase I crack growth rate to defect severity
Δ	An increment of a function
δ_0	Depth of asperity tip
δ	Vertical distance between asperity tip and contact edge
ϵ_c	Plastic strain at micro-defect
ϵ_0	Total macro-strain
ϵ_p	Plastic strain
η	Number of defect cells per unit volume; reduction factor accounting for sliding

NOMENCLATURE (CONT'D)

Φ	Severity function
θ	Slope angle of asperity sides; an angle
Λ	Crack growth rate function
λ	Constant proportional to radius of asperity tip; reduction factor for roller end stress con- centration
$\lambda(N)$	Material condition after N cycles
ν	Poisson's ratio
ρ	Geometry parameter
σ_0	Surface roughness
σ_0 (p_0)	Maximum Hertzian contact pressure (lb/in^2)
$\sigma_x, \sigma_y, \sigma_z$	Normal stresses in x, y and z direction (lb/in^2)
$\sigma_{11}, \sigma_{22}, \sigma_{33}$	Principal stresses (lb/in^2)
σ_D	Von-Mises yield stress (lb/in^2)
σ_y	Yield strength (lb/in^2)
$\tau_{xy}, \tau_{yz}, \tau_{xz}$	Orthogonal shear stresses (lb/in^2)
τ_θ	Shear stress on plane inclined at angle θ to surface plane (lb/in^2)
τ_R	Reversing shear range (lb/in^2)
τ_0	Maximum amplitude of τ_{yz} (lb/in^2)
τ_{Oct}	Octahedral shear stress (lb/in^2)
z	Complex number ($= x + iz$)

NOMENCLATURE (CONT'D)

τ_{45°	Shear stress on plane inclined at 45° to surface plane (lb/in ²)
$\tau_{\max}$	Maximum value of τ_{45° (lb/in ²)
ϕ	A function of bearing geometry parameters
σ	Function of shear stress range and depth coordinate
ψ	Angular coordinate
ω	Ratio of surface defect shoulder width to the width of contact edge

Subscripts

I, II, III	Pertaining to Phases I, II and III of crack growth
1, 2	Pertaining to contacting bodies 1 and 2
i, o	Pertaining to bearing inner and outer ring
i	Pertaining to i-th cell
v, s	Pertaining to "weak" volume and surface elements (or cells)

EVALUATION

1. The present technique for predicting the life of a group of rolling element bearings does not consider the many factors known to affect bearing life, and for large bearings the technique is clearly inadequate. This contract is the first part of a two part effort to develop a practical engineering tool for the determination of the expected life of any group of bearings. This first contract was to consider all the variables that affect bearing life and the possible failure mechanisms involved, and then to develop equations which would contain parameters to account for those variables known to affect bearing life. These objectives have been accomplished and the results of this contract have provided a number of equations containing parameters characterizing material, geometry, load, defect severity, and environmental variables.

2. The above mentioned equations contain constants and function signs which must be evaluated and determined from test and field data before the technique can be used as an engineering tool. This is to be accomplished in the second effort. The results of these efforts will be included in a mechanical reliability handbook and should provide an improved prediction technique.

William J. Bocchi
WILLIAM J. BOCCHI

Mechanical Engineering Section
Development Engineering Branch

SECTION I

INTRODUCTION AND SUMMARY

This is the Final Report issued in fulfillment of Rome Air Development Center Contract No. F30602-67-C-0147 on "Development of Mathematical Models Predicting Life of Large Roller Bearings".

The objective of this work is to determine the variables which cause the life of (large) rolling bearings to vary from the life predicted by the existing design methods; to determine the effects of those variables on the life of rolling bearings and to formulate an improved mathematical model for the prediction of rolling bearing life.

The present study covers the first year of this effort and has led to a general model of bearing failure. A subsequent effort is currently underway (RADC Contract No. F30602-68-C-0147) to cover future development of the model and to make the formulas sufficiently specific for engineering use.

This report is divided in several sections summarized as follows:

Section II is a description of the principal concepts developed in this Contract. Using the currently accepted formula as a starting point, this section brings together all the new concepts generated in this study and explores the usefulness of the new model. An outlook on future research is given.

Section III presents a synopsis of the currently accepted Lundberg-Palmgren bearing fatigue life theory which forms the basis of the ASA standard for bearing rating and is the starting point for the present study. This section is included in recognition of the fact that the fundamental work of Lundberg and Palmgren may not be easily accessible to all readers.

Section IV extracts from recent literature, the principles of fatigue failure theory required for this study, as follows:

1. A Survey of rolling contact failure identities among which is spalling failure. This failure mode is the subject of the present study.
2. A listing is given of variables affecting contact fatigue life. These variables are grouped into four main categories. viz. material variables, surface

microgeometry variables, design variables and operating variables.

3. An evaluation is presented of the interdependence of the variables and their effect on subsurface and surface initiated spalling occurrences.
4. A model is offered of fatigue failures. Subsurface and surface initiated fatigue failures, are distinguished which compete to promote spalling failures in rolling contact. In both subsurface and surface failures, the fatigue process is described as a sequence of phases of fatigue crack generation, propagation, and final fracture, (i.e., spalling) at a "most critical" crack in the rolling element. Crack generation in rolling contact is postulated to result from localized plastic strain concentration around stress raisers.

Section V covers the formulation of an expression for the crack growth rate as a function of strength and stress parameters, ductility, and plastic micro-strain.

This concept is applied to a situation where defects of known "severity" exist in a uniform matrix, to yield a formula for the fatigue life of a defect.

Section VI gives a statistical theory of failure for an entire rolling body, based upon the relationship between life and defect severity developed in Section V. In this model, a rolling body is conceived as being built up of a large number of small cells each of which contains exactly one defect (including "defects" of no influence at all). The severity of the defect present in any cell is a random variable. The distribution of life over identically located cells in a population of rolling bodies similarly made and operated, is found through a transformation of distributions between severity and life.

The distribution of rolling body life is expressed as a compound of the individual cell life distributions. The asymptotic distribution of shortest cell life is derived for the case where all cell lives are commonly distributed.

Section VII covers a required stress analysis in a Hertzian contact. The micro-strain range in the highly stressed volume (or surface) is calculated.

Section VIII presents a specialized elastic stress analysis for determination of the maximum shear stress near an idealized surface asperity.

Section IX presents a stress analysis for the determination of maximum shear stress beneath a furrow-shaped surface defect.

Analytical details are supplied in several Appendices.

SECTION II

PRINCIPAL CONCEPTS

A mathematical model of rolling contact fatigue is a complex subject. Numerous aspects of metallurgical, mechanical, and statistical nature have to be considered in its development. At the present stage of the development, many of these details are still open. Others have been covered in quite some depth. The present report is a Summary of the studies to date. It will take up the several aspects of the problem in turn, at whatever depth is currently accessible. There is the possibility with this presentation that the reader may be diverted from the principal underlying concepts by the complexity of detail. In order to prevent this and to facilitate evaluation of work accomplished from the point of view of the engineer, who will ultimately use the theory for practical life predictions, a review of the principal concepts is offered in the present Section. No proofs or references will be cited: these are either given in the subsequent Sections or referenced there.

1. THE PRINCIPLE OF ROLLING BEARING LIFE PREDICTION

Rolling bearing life is defined here as fatigue life. Causes of failure other than fatigue are considered avoidable and are, therefore, eliminated from life prediction. Fatigue life is predicted on the basis of a cumulative damage concept, i.e. that with repeated application of cyclic stresses, irreversible material changes take place which ultimately result in failure. This concept, with its statistical implications, was first applied to rolling contact life prediction by Lundberg and Palmgren. The Lundberg-Palmgren concept, universally used today, revolves around a phenomenological equation of the following form, between numbers of cycles to failure, and macroscopic mechanical variables:

$$\log \frac{1}{S(N)} \sim N^e \varphi(\tau_0, z_0) V \quad (2.1)$$

where $S(N)$ = the probability of survival to N cycles

τ_0 = maximum shear stress amplitude

z_0 = depth co-ordinate of τ_0

V = stressed volume

φ = function sign

e = constant

This equation explicitly contains the number of cycles, a maximum shear stress and its depth co-ordinate, and the "stressed volume" which latter, however, is never expressed in absolute terms, only as a factor of proportionality.

The Lundberg-Palmgren life prediction theory consists of the application of Equation (2.1) to the required wide variety of geometrical and kinematic conditions which characterize a complex assembly such as a rolling bearing.

Physically, Equation (2.1) teaches that the cumulative probability of survival decreases with increasing number of cycles N , and with increasing size of the rolling contact system (stressed volume). The specific choice of the function $\varphi(\tau_0, z_0)$ was made by Lundberg and Palmgren once and for all, and is given in Equation (2.2)

$$\varphi(\tau_0, z_0) = \tau_0^c z_0^{-h} \quad (2.2)$$

c, h : constants > 0

This equation states that the survival probability decreases with increasing shear stress range, but increases for greater values of the depth co-ordinate of the maximum shear stress range.

By applying elastic analysis to the contact situation, Lundberg and Palmgren derived detailed statements regarding the effect of the pertinent macro-geometry parameters influencing normal surface pressures in the contact, the subsurface shear stresses resulting from these pressures and the kinematic parameters determining numbers of cycles in terms of bearing ring revolutions. Equation (2.1) is readily modified to take account of time-variable or space-variable loading by using the "Palmgren-Miner hypothesis" of damage accumulation, stating that fatigue damage accumulates at a rate depending only on load conditions at the current time, so that Equation (2.1) can be written in the form:

$$\log \frac{1}{S(N)} \sim \int \left[\int \varphi(\tau_0, z_0) dN \right]^6 dV \quad (2.3)$$

$(V)(N)$

The phenomenological nature of Equations (2.1) through (2.3) results in an impasse if one attempts to incorporate into life prediction, newly acquired knowledge regarding the effect of parameters other than contact geometry and kinematics, since these equations offer no clue as to the proper role of such parameters in defining life. For this reason, past attempts at

refining the Lundberg-Palmgren theory have relied on the fact that Equations (2.1) and (2.3) are proportionalities, i.e. they contain a freely available constant multiplier relating the absolute magnitude of life to the probability of survival. This multiplier is intended by Lundberg and Palmgren as the material constant, but has, from time to time, been used to incorporate a variety of correction factors.

The limitations of this approach are obvious, and it has therefore been decided in the present study of improved life prediction methods to abandon attempts at modifying the basic Lundberg-Palmgren equation. Rather, it was decided to generate novel equations from which the Lundberg-Palmgren equations can be obtained as a special case.

The new equations were derived using a more detailed physical model of fatigue failure rather than as purely phenomenological equations. It is recognized that the use of such a model has many pitfalls, the most obvious being that its details may not be verifiable. However, the drawbacks are more than compensated by the heuristic value of a detailed model and can be rendered harmless by insisting that non-verifiable details of the model should not enter into the final engineering formula for life prediction.

Before leaving this brief review of Lundberg-Palmgren theory, it is noted that Equation (2.1) is equivalent to:

$$H(N) = 1 - S(N) = 1 - \exp \left\{ -\left(\frac{N}{N^*}\right)^e \right\} \quad (2.4)$$

where N^* = constant "scale parameter" of the life distribution

$H(N)$ = cumulative probability of failure within N cycles

i.e. failures are distributed according to a Weibull distribution with zero lower bound, characteristic life N^* and dispersion exponent e . This distribution appears in the Lundberg-Palmgren formulation as a result of deliberate choice, as a useful distribution for the description of fatigue phenomena, and its appearance does not stem from extreme value considerations. This point will be of interest later.

2. PRINCIPAL VARIABLES OF FATIGUE FAILURE

Section IV gives a detailed review of the variables governing a fatigue failure situation. There, it is deduced that the variables fall in four main categories: material variables, surface microgeometry variables, design variables, and operating variables.

Material variables are those influencing the "strength" of the rolling system. Current fatigue investigations (chiefly of the non-rolling type) consider yield strength and ductility of the material as dominant bulk (or matrix) strength variables. Modifying these are residual stresses and work hardening effects, acquired, in part, during the course of fatigue life. In rolling contact, the materials used are of high hardness, and such materials do not react with their matrix strength. Rather, rolling contact life appears to be determined by the strength of the material in the vicinity of inevitable material imperfections such as inclusions in the matrix or microcracks. Thus, the nature of these imperfections is a dominant variable.

Surface microgeometry variables determine the detailed nature of the contact through which loads are transmitted to the material. The generalized roughness of the surface determines the topography of the contacting surface elements, the plastic behavior of the material immediately adjacent to the surface, and interacts with lubrication, as will be seen presently.

There are localized imperfections on surfaces, mostly in the form of sharp depressions ("furrows") which form stress raisers near their edges and are influential in failure. Of course, there can be many other types of surface variables, some of them artificial as induced by coating, special treatment of the surface, etc.

The design variables of the rolling contact are dealt with extensively by the Lundberg-Palmgren theory, and they are, therefore, quite familiar. Track length, conformity between rolling elements, dimensions and number of these elements, contact angles, parameters defining the precise cross track geometry and many others are influential, primarily because they determine the macroscopic (Hertzian) stress field and the number of cycles as a function of bearing ring rotation. They also determine the magnitude of the highly stressed volume.

Operating variables encompass the external environment in which the rolling contact system must endure, including load, speed, lubrication, temperature, and atmospheric conditions. The influence of load in determining the stress field is obvious, and so is that of speed in determining the number of cycles per unit time. However, these two variables also interact with the lubricant to determine the hydrodynamics of the often-present pressure-bearing elastohydrodynamic (EHD) lubricant film in the contact, which redistributes stresses and has important effects on the microbehavior of the contact area (asperity interaction). Temperature enters by influencing both the material strength and the lubrication, and atmospheric conditions can be of consequence if they influence lubricant behavior or cause corrosive effects.

This list of influential parameters deters the theorist by its multiplicity. The only practicable approach to the development of a life prediction formula in the presence of such a multitude of variables is to find a flexible, simple concept describing failure mechanism, and then solve the problem of introducing each variable by defining its impact on that mechanism. Success of such an attempt depends on the proper choice of the mechanism and will necessarily be limited. There will always be variables that the model cannot accommodate, and as time passes, the influence of these will become more and more recognized, leading ultimately to the abandonment of the model. However, the model will serve well in the interim if it permits account to be taken of the most important parameters recognized to date. In what follows, such a model will be outlined. It appears at the present stage of the study to have the required flexibility and to accommodate many of the most important parameters, including all those which the Lundberg-Palmgren theory utilizes. It will remain for further study to develop the specific formulations for the incorporation of new parameters into this model and to show whether it is sufficiently free from inherent contradictions to be practically used. This further effort is currently underway.

3. THE FATIGUE FAILURE MODEL

Our failure model visualizes fatigue damage as the growth of a crack. There are plastic flow occurrences, carbon migration, and, of course, first of all, dislocation motions in the matrix as a result of cyclic stressing, which precede or are concurrent with crack formation. However, for modeling purposes, these subtler occurrences are not helpful in fixing ideas of fatigue damage because no way is known to measure the degree to which they are shortening life expectancy. The effect of a crack on life is intuitively clear: when the crack has become large enough a piece

of material will separate from the surface, forming the fatigue spall defined in rolling contact technology as fatigue failure. Crack growth is visualized as irreversible, so that a suitable measure of crack size satisfies the concept of "damage" as irreversible progress towards failure. The fatigue phenomenon starts, accordingly, with the initiation of a crack, and proceeds through stages of its growth until the crack has become large enough to form a spall. Fatigue life, as determined by the crack in question, begins with the onset of cycling and terminates when the spall forms. Of course, it can be argued that a rolling contact system may be functional in the presence of a spall of tolerable size. There is room to accommodate this argument in the model, as will be pointed out later, but the current description of the failure terminates with first spalling.

It is convenient to describe failure generation in three phases:

Phase I begins immediately upon the onset of cyclic stressing, and is consumed by the formation and growth of a microcrack. A definition of a microcrack will be given below. From the point of view of the model, it is characterized by the fact that it is small enough not to interact with other microcracks that may be present in the rolling element.

In Phase II, the crack grows macroscopically until, at the end of this phase, it has reached a critical size, defined by the fact that it is now large enough to cause the initiation of precipitous crack growth (in Phase III).

Phase III is occupied by precipitous crack growth at a rate greatly in excess of Phase II growth. This precipitous cracking forms the spall itself. This Phase may be virtually instantaneous or consume substantial length of time, depending on whether one specifies a minimum spall size, which is accepted as a failure, and depending on a variety of material and operating conditions.

The objective of a mathematical life formula is to describe crack growth through the above three phases. In order to describe crack growth, one selects a measure of crack size A and formulates an equation of the form:

$$A = f_1 (N, X_i) \quad (2.5)$$

where A = crack size
 f_1 = function sign
 X_i = undefined parameters
 N = number of stress cycles

At the present state of the study, it seems best to select a form of Equation (2.5) common in current fatigue theory, viz. one defining the first derivative of crack size (the crack growth rate) in terms of the relevant variables:

$$\frac{dA}{dN} = f_2(N, A, X_1) \quad (2.6)$$

where f_2 = function sign

Life prediction is then accomplished by determining from Equation (2.6) that value of the number of cycles N_L which corresponds to a critical crack size A_c , causing immediate spalling, i.e.

$$N_L = f_2(A_c, X_1) \quad (2.7)$$

where N_L = life at failure

A_c = critical crack size at spalling

Many current theories of fatigue failure use the crack growth Equation (2.6) in the following simple form:

$$\frac{dA}{dN} = \Lambda(e_c, D) \approx \Lambda(e_c(N), D(N)) \quad (2.8)$$

where Λ = function sign

e_c = plastic strain at the propagating crack front

D = ductility

According to Equation (2.8), the only variables entering the crack propagation equation are a plastic strain (measured at the propagating crack front) and a measure of material ductility. Specific definitions of these two variables in terms of measurable physical quantities are open at this point, both because appropriate definitions for the rolling contact situation have not previously been determined, and also because the plastic strain e_c is a microparameter which is not directly measurable. Note that Equation (2.8), for all its simplicity, contains many assumptions. Only the first derivative of crack size appears explicitly. (The crack size itself enters by way of its influence on e_c .) Only variables measurable at the location of the propagating crack front appear, and these only with their values assumed at the time of the N -th stress cycle. (Of course, the equation is compatible with a dependence $e_c(N)$ and $D(N)$, and the specific form of this dependence determines whether this formula satisfies the Palmgren-Miner hypothesis.) There is hardly room for concern about restrictiveness at this point, however, since even Equation (2.8) is much too general to be practically applicable.

In order to add definition to Equation (2.8), the concept of "defects" is introduced. A defect is a location in the material or at the surface, at which there is a tendency of crack generation. It will be assumed that this tendency manifests itself in Equation (2.8) by some property of the defect causing ϵ_c to be higher at the defect than elsewhere in the vicinity.

A simple form of this relationship can be written by assuming that one can select (as in uniaxial tension) a critical scalar ϵ_0 of the total (elastic plus plastic) macroscopic strain field as it would exist at the defect location, in the absence of the defect, such that ϵ_c depends only on ϵ_0 , on defect severity and on the yield strength σ_y of the matrix, i.e.

$$\epsilon_c = f(\theta, \epsilon_0, \sigma_y) \quad (2.9)$$

where ϵ_c = plastic strain at defect
 ϵ_0 = critical "undisturbed" total (elastic plus plastic) strain
 θ = defect severity measure
 σ_y = (micro) yield strength of the matrix
 f = function sign

The defect severity factor θ is defined as a "strain raising" factor characteristic of the defect. Conveniently, θ is defined for all real defects with strain raising properties, but also for an "ineffective defect" with no severity at all, i.e. one which does not raise the magnitude of the strain. For purposes of statistical treatment, it is convenient to define such "ineffective" defects as the limiting case of defects with real stress raising properties.

Introducing Equation (2.9) into Equation (2.8), one has

$$\frac{dA}{dN} = \Lambda(\theta, \epsilon_0, \sigma_y, D) \quad (2.10)$$

It is convenient to separate the variables influential in crack growth into the two groups: variables related to defects, and matrix variables. In Equation (2.10), θ is the variable related to defects, whereas ϵ_0 , σ_y , and D are related to the matrix. For simplicity, the matrix effects are consolidated into a single function γ , i.e.

$$\frac{dA}{dN} = \Lambda(\theta, \gamma); \quad \gamma = \gamma(\epsilon_0, \sigma_y, D) \quad (2.11)$$

The new assumptions underlying Equations (2.9) through (2.11) are that there is a scalar ϵ_0 of the macrostrain field which, alone, among strain field characteristics, determines ϵ_c , and that all matrix parameters exert their influence on crack growth via a single quantity γ . Neither of these assumptions is essential in order to arrive at a workable model, but are made here in the absence of a more refined understanding of the actual physical situation, to arrive at a relatively simple formulation.

Further development of Equation (2.11) requires use of a further restrictive concept such as the hypothesis of multiplicative effects on fatigue life. This concept, also used by Lundberg and Palmgren, asserts that the rate of fatigue damage (crack growth) can be expressed as a product of a number of independent factors, i.e.

$$\frac{dA}{dN} = \prod_i (\varphi_i^{a_i}) \quad (2.12)$$

where φ = unspecified independent factors

$\prod$ = multiplication operator

a_i = constants

Applying this concept to Equation (2.11), one may, by suitable definition of the functions Θ and γ , absorb in them the function Λ , and write

$$\frac{dA}{dN} = \Theta \cdot \gamma \quad (2.13)$$

We will proceed now to the examination of the matrix factor γ and the defect factor Θ .

4. THE MATRIX FACTOR γ

From the definition of γ given in Equation (2.11), it is a function of a critical total macrostrain scalar at the location of the growing crack front, and of a yield strength and a ductility measure.

The yield strength measure σ_y will be a microyield stress, since small scale plastic occurrence are at issue in rolling contact fatigue. It must be taken with its value at the time of the N-th stress cycle, to account for work hardening or work softening of the matrix.

The ductility is defined in static tensile tests as the reduction in cross sectional area at fracture. It is not obvious that this simple definition will apply under the conditions of contact fatigue, but it is intuitively convincing that one should include in the formula a material ductility property measuring the amount of plastic strain the matrix can absorb before it cracks. A variety of metallurgical parameters, but also some operating conditions, will determine ductility. The most important operating condition is hydrostatic compressive stress. It is generally believed that the high hydrostatic compression component existing in most of the Hertzian contact stress field retards crack formation. This fact will manifest itself in a point-wise varying value of the ductility parameter D when examining material elements located at different points within the Hertzian stress field. Inasmuch as it may depend on work hardening, ductility can be a function of N . Thus, the ductility parameter is already known to depend on material constants, on a parameter of the stress field, and can depend on N . It may also be related to other operational parameters, e.g. to cycling rate. These relationships are symbolized by the following equation:

$$D = D(M, \sigma_h, N, \varphi) \quad (2.14)$$

where M = material variables
 σ_h = hydrostatic compressive stress
 φ = operating factors

Turning to the critical macrostrain parameter ϵ_0 , it is obviously dependent on the variables of load W , contact geometry ρ , position under the contact $\vec{x}$, and elastic modulus E , defining between them the elastic Hertz stress field. With a quasi-elastic assumption, these parameters give a relationship of the form

$$\epsilon_0 = \epsilon_0(W, E, \vec{x}, \rho) \quad (2.15)$$

where W = load
 E = elastic modulus
 $\vec{x}$ = position vector
 ρ = contact geometry parameters

The "quasi-elastic" assumption operates on the scale of the whole Hertzian stress field and disregards the vicinity of defects. It postulates that the macroscopic total strain ϵ_0 can be calculated from the elastic (Hertzian) stress field of the rolling contact. This is the case if the loads are sufficiently low that, at most, very small amounts of macroscopic plastic flow take place so that the plastic component of total strain is negligible

and that plastic flow does not result in a significant redistribution of elastic stresses. Except for the generation of residual stresses due to cyclic stressing, this is a reasonable assumption in all practical rolling contact fatigue situations. The question of residual stresses will require separate examination. Generally, they are handled by assuming that, after a small number of cycles, the residual stresses have "shaken down" to a constant value. Then, they act as a superimposed static stress field and combine with the cyclic stresses. The resulting time-variable stress field is, of course, different from that existing without residual stresses, and assumptions must be made regarding the effect of this difference on crack propagation.

A common assumption in fatigue theory is that superposition of a static stress field does not alter the plastic strain ϵ_c influencing crack growth rate. However, the hydrostatic compression component of the residual stress field may modify the ductility D .

Everything said above about quasi-elastic behavior is restricted to the matrix at locations remote from defects. Due to the stress raising effect of defects, it is, of course, possible that localized plastic occurrences take place in small volumes in their vicinity. Such microplasticity is, in fact, the condition of cracking in the proposed model.

For any given defect Θ , and given matrix strength (σ_y), it is possible to delineate that portion of the Hertzian stress field within which the macrostrain ϵ_0 is high enough to cause plastic microstrain ϵ_c . For a given population of defects, there will be a "realistic" maximum severity Θ . One can delineate a highly stressed area in the Hertz stress field within which all microplastic occurrences occasioned by defects of "realistic" severity will be confined. This definition of a "highly stressed zone" will be adopted in what follows, and the cross sectional area of this highly stressed zone, in a plane perpendicular to the rolling direction, will be designated by S .

5. THE DEFECT SEVERITY FACTOR Θ

The effect of a "defect" in generating plastic strain in its vicinity is manifestly very complex. A simple relationship for Θ will be proposed as follows:

$$\Theta(N) = \Theta(d, A(N), S) \quad (2.16)$$

where d = initial defect severity
 $A = A(N)$, crack size after N cycles
 S = size of the highly stressed zone

The variables of Θ are: d , the severity of the defect at the onset of cyclic stressing; $A(N)$, the crack size at the time of the N -th stress cycle; and S , the cross sectional area of the highly stressed zone in the Hertz contact.

As mentioned before, crack propagation can be envisioned as occurring in three phases. The first phase, microcrack propagation, can be defined using Equation (2.16) by postulating that the crack is so small that, by comparison, the size of the highly stressed zone can be considered infinite, so that for Phase I:

$$\Theta_I = \Theta_I [d, A(N)]; \quad A \leq A_p \quad (2.17)$$

where A_p = self-propagating crack size at the end of Phase I

Phase II, on the other hand, can be defined as extending from that point in time where the crack has grown sufficiently large to outweigh, in its effect on propagation rate, the original defect. Such a crack does not require the defect to propagate, it is "self-propagating". For a crack of this size, or larger, the size of the highly stressed zone can no longer be considered infinite, so that one has:

$$\Theta_{II} = \Theta_{II}(A(N), S); \quad A_p \leq A \leq A_c \quad (2.18)$$

where A_c = critical crack size

Here, A_c is the crack size at the termination point of Phase II. This size crack is assumed to lead to a spall "instantaneously" by a precipitous fracture mechanism. This does not suggest that the rolling contact system becomes inoperative immediately, although this may be the case. However, there is a "catastrophic" growth step between the crack of size A_c and the completed spall, i.e.

$$\frac{dA}{dN} \rightarrow \infty \text{ for } A \rightarrow A_c \quad (2.19)$$

One can assume that the critical crack size is related to the size of the highly stressed zone, or, in its simplest form:

$$A_c = k_1 S \quad (2.20)$$

where k_1 = constant

6. PREDICTION OF THE LIFE AT A DEFECT

Substituting Equations (2.17) and (2.18) respectively, into Equation (2.13), one obtains the following formulas for crack propagation rate during Phases I and II:

$$\left(\frac{dA}{dN}\right)_I = \gamma_I(\epsilon_0, \sigma_y, D) \cdot \Theta_I[d, A(N)]; \quad A \leq A_p \quad (2.21a)$$

$$\left(\frac{dA}{dN}\right)_{II} = \gamma_{II}(\epsilon_0, \sigma_y, D) \cdot \Theta_{II}(A(N), S); \quad A_p \leq A \leq A_c \quad (2.21b)$$

where subscripts I and II apply to Phases I and II respectively.

Using the previously explained multiplicative hypothesis on the function Θ , these equations may be rewritten as follows:

$$\left(\frac{dA}{dN}\right)_I = \gamma_I(\epsilon_0, \sigma_y, D) \cdot f_1(A) \cdot \Gamma(d); \quad A \leq A_p \quad (2.22a)$$

$$\left(\frac{dA}{dN}\right)_{II} = \gamma_{II}(\epsilon_0, \sigma_y, D) \cdot f_2(A/S) \cdot f_3(A); \quad A_p \leq A \leq A_c \quad (2.22b)$$

In Equation (2.22b), two functions f_2 and f_3 are shown, one representing the effect of relative crack size by comparison to the size of the highly stressed zone, and the other any remaining direct effect of absolute crack size (as hypothesized e.g. by Lundberg and Palmgren when introducing the effect of the depth co-ordinate z_0 of the maximum shear stress range.) Note that in Equation (2.22b), the original defect severity d does not appear. Therefore, this equation contains only macrostrain and matrix variables, and is independent of the original defect population. In Equation (2.22a), on the other hand, $\Gamma(d)$ is different for each individual defect, and the equation is, therefore, dependent on the defect population.

Integration of the differential Equations (2.22a) and (2.22b) leads to the following forms:

$$f_I(A) = N \gamma_I \Gamma(d) \quad (2.23a)$$

$$f_{II}(A, A/S) = N \gamma_{II} \quad (2.23b)$$

Substituting A_p into Equation (2.23a) yields a value N_I , the life at the end of Phase I. Substitution of A_c into the Equation (2.23b) yields a value N_{II} , the duration of Phase II life.

The life N_L from the beginning of cycling through the end of Phase II life is then obtained as the sum of these two phase-lives $N_L = N_I + N_{II}$. Phase III life may or may not be 0, depending on the definition of failure as discussed above.

In principle, then, it is possible to obtain a prediction of life to failure at a particular defect with severity d , from Equations (2.23a) and 2.23b).

It is noted that the apparent arbitrariness in the selection of the self-propagating crack size A_p will not influence the total life $N_L = N_I + N_{II}$ if the hypotheses outlined previously are correct, because the Equations (2.23a) and (2.23b) were both obtained from Equation (2.16) by neglecting, for Equation (2.23a), the influence A_p , and for Equation (2.23b) the influence of d . Inasmuch as these approximations are valid, the two equations merely describe two portions of the same function $A(N)$, and their domains of validity overlap so that the selection of A_p is, within limits, discretionary.

Solving Equations (2.23a) and (2.23b) for N , and substituting as described above, one obtains the following expressions for life to failure N_L (at the end of Phase II)

$$N_I = \frac{f_I (A_p)}{\gamma_I \Gamma(d)} \quad (2.24a)$$

$$N_{II} = \frac{f_{II} (A_c, A_c/S)}{\gamma_{II}} \quad (2.24b)$$

$$N_L = N_I + N_{II} \quad (2.24c)$$

One must settle on a suitable value of A_p and determine that value of A_c/S at which the crack becomes critical. Assuming these decisions can be reached generally, γ_I and $\Gamma(d)$ remain, in Equation (2.24a), as functions of external parameters and S and γ_a in Equations (2.24b). All parameters of these remaining functions are observable.

7. THE STATISTICS OF LIFE FOR AN ENTIRE ROLLING BODY

In determining the life of a finite-size rolling body, one starts with Equations (2.24) for life in the vicinity of a given defect, and uses statistical theory to obtain life for a volume of material containing a multitude of defects.

Failure of the rolling body will occur through "competition" between a multitude of potential defects acting as failure nuclei. According to the definition of Phase I fatigue, microcracks grow at a multitude of defects, at differing rates, and independently of each other. One of these microcracks, or several, will reach the beginning of Phase II, within the life of the part. These will then proceed to accumulate Phase II life until such a time as one of them has generated a crack of critical size A_p , at which a spall forms, whereupon the rolling body is considered failed. All other defects which have entered Phase II show, at the time of failure, cracks of less than critical size. In the proposed model, variations in original matrix strength within a rolling body are considered small enough to be neglected (this position can be revised later if necessary). This leaves two main sources of variability: the systematic point-wise variability of the macro-strain field in the contact zone (and the consequent variability of work hardening and residual stresses) and a random variability of defect severity and location with reference to the contact zone. It is the effect of these variables on the life calculated from Equations (2.24a) and (2.24b) that determine the outcome of the competition among defects for the generation of the crack leading to failure.

A statistical treatment which can be used to describe this competition will be illustrated for Phase I.

Phase II life will be considered a deterministic quantity, calculable for each point in the rolling body, from the knowledge of macroscopic strain and matrix parameters alone.

To express the statistics of Phase I life, consider the highly stressed zone to be composed of elementary "cells" of uniform size, selected small enough to contain only one defect, but large enough for a crack of size A_p to be wholly confined within the cell. Then, Phase I fatigue damage, originated within a cell, will remain confined within it. Fatigue damage existing in one cell will not influence the behavior of adjacent cells. On this assumption, Phase I life of each cell is independent of the life of all other cells.

Assume now that there is a known probability distribution of defects of varying severities d for each cell, i.e. there is a known cumulative distribution function $F(d)$ such that in any cell

$$\text{Prob} (d_1 \leq d) = F(d) \quad (2.25)$$

where F = cumulative distribution function of d

Then, Equation (2.24a), establishing a functional relationship between d and N_I , permits determination of a transformed probability distribution $G(N)$ such that

$$\text{Prob} (N_1 \leq N_I) = G(N_I | \bar{x}) \quad (2.26)$$

where G = cumulative distribution function of N_I
 $\bar{x}$ = position vector

In the general case where the functions f_I , V_I , and Γ in Equation (2.24a) vary from point to point in the rolling body because of the non-uniformity of the macrostrain field, or for any other reason, the distribution $G(N_I)$ will depend on the position vector $\bar{x}$ shown in Equation (2.26).

Equation (2.26) states the (cumulative) probability that the cell with position co-ordinate $\bar{x}$ will reach the end of Phase I life in N_I cycles or less.

From Equation (2.26), statements can be made regarding the probability of failure of the entire rolling body. The rolling body will fail if exactly one of its cells fail. It is, therefore, required to express the probability distribution of the life of that cell in the rolling body which fails first of all cells.

If the fatigue phenomena in each cell are independent as assumed, the probability that the rolling body survives is the product of the survival probabilities of all cells in it, i.e. the life distribution of an entire rolling body is:

$$H(N_I) = 1 - \Pi [1 - G_i(N_I)] \quad (2.27)$$

where H = cumulative distribution function of N_I for the entire rolling body
 Π : multiplication operator

Equation (2.27) follows from Equation (2.26) by observing that the probability of survival is obtained by subtracting the probability of failure from unity.

Equation (2.27) is general. A special case is that for which all G_i are identical. This is a simple approximation of the case of a thrust loaded bearing ring (the Hertz stress field is independent of contact position) in which there is a single defect distribution throughout the ring and the defects are infrequent so that a "cell" can be represented by a short "slice" of ring between two closely spaced cross-sections. In this simple case, Equation (2.27) reduces to the following:

$$H(N_I) = 1 - [1 - G(N_I)]^m \quad (2.28)$$

where m = number of cells in the rolling body

As is shown in a later section of this report for certain general classes of distributions $G(N_I)$, and for increasing m , Equation (2.28) approaches the form of a Weibull distribution:

$$H(N_I) = 1 - \exp \left(- \frac{(N_I - N_0)^k}{N^*} \right) \quad (2.29)$$

where $N^* \sim m^{-1/k}$ = a "characteristic life" or scale parameter

k = constant dispersion exponent

N_0 = minimum life, ($N_0 \approx 0$)

The constants are determined by the specifics of the distribution function $G(N_I)$. N_0 , the minimum Phase I life, can be considered zero, since a crack of size A_p may pre-exist in the matrix. Equation (2.29) emerges from general theorems on the asymptotic properties of extreme value distributions and is not a separate hypothesis.

In the model presented, the (Phase I) life distribution of a rolling body is not merely observed as a phenomenological fact, it is related to the physically meaningful distribution of defects. The relationship between the distributions $F(d)$ in Equation (2.25) and $G(N_I)$ in Equation (2.26) permits inferences from one distribution to the other, thereby identifying suitable measures of defect severity. As shown in subsequent sections, it is possible to conjecture appropriate defect severity distributions and test these conjectures by the effect they have on the life distribution of the part.

Equation (2.29) contains a volume effect on Phase I life. For cells of fixed size, their number m is proportional to the stressed volume. Therefore

$$N^* \sim V^{-1/k} \quad (2.30)$$

i.e. the scale parameter of the Weibull life distribution is an inverse function of the stressed volume as in the Lundberg-Palmgren theory.

The observable life N_L of an entire rolling body is, according to Equation (2.24c), the sum of Phase I and Phase II lives at the defect where N_L will be shortest. This can be calculated simply, only under one of the following two conditions for Phase II life N_{II} :

(a) $N_{II} \ll N_I$, or

(b) $N_{II} \approx \text{constant}$

For either condition, the shortest N_L will coincide with the shortest N_I , so that N_I can be taken from Equation (2.27), (2.28) or (2.29). N_{II} is then calculated from Equation (2.24b) for the $\bar{x}$ at which the shortest N_I occurred and the two are added.

If neither of the above conditions for N_{II} is satisfied, then the (deterministic) variation of N_{II} with $\bar{x}$ may cause the shortest N_L . For this case, the statistical treatment must include N_{II} , and this may lead to difficulties in satisfying the assumption of independence of all failures required for Equation (2.27).

6. APPLICATION OF THE LIFE EQUATIONS

Equations (2.24a), (2.24b), and (2.24c), and the statistical interpretation contained in Equations (2.25) to (2.29), represent the framework of the proposed fatigue life prediction model. In order to apply this model, it is necessary to make specific assumptions for all functions. This work will be pursued in the next project year. It is instructive, however, to illustrate the approaches that can be taken towards application of the model by considering a few special examples at this time.

a. The Lundberg-Palmgren Case

The following is an example of one of several possible methods by which the Lundberg-Palmgren formulas can be obtained as a special case of the proposed model.

If, in Equations (2.24a), (2.24b), and (2.24c), one assumes that $\Gamma(d)$ is a material constant, γ_I depends only on the maximum shear stress range τ_0 , A_p is proportional to a linear dimension of the highly stressed cross section, say,

the depth co-ordinate z_0 of the maximum shear stress range and $N_{-I} = 0$, then one obtains

$$N_L \sim \frac{f_I(z_0)}{\gamma(\tau_0)} \text{ or } N_L \gamma(\tau_0) f_I^{-1}(z_0) = \text{const.} \quad (2.31)$$

If one uses power functions for γ and f_I , then:

$$N_L e^{\frac{\tau_0 c/e}{z_0 h/e}} = \text{const.} \quad (2.32)$$

The designation of the constant exponents is that used in Equations (2.1) and (2.2). Equation (2.32) states that the life distribution of each defect and consequently that of the entire rolling body, is scaled by the maximum alternating shear stress τ_0 and its depth co-ordinate z_0 exactly as specified by the Lundberg-Palmgren theory. Equation (2.30) states that the stressed volume is another scale factor for life. Equations (2.29) and (2.32) are equivalent to Equations (2.1), (2.2), and (2.4). Thus, the Lundberg-Palmgren formula appears as a special case of the new model.

b. Deviations from the Lundberg-Palmgren Type Weibull Distribution

It has been found experimentally that there is a non-zero minimum life prior to which there is no finite probability of bearing failure. This fact is not explicable by the Lundberg-Palmgren theory. However, it is immediately obvious if Equation (2.24b) is assumed to give a non-zero value for N_{II} . In this case, N_I may still be distributed according to a Weibull distribution with $N_0 = 0$, Equation (2.29), but the total life to failure N_L will have a positive minimum value.

c. Effect of Material Cleanness (Inclusion Content)

Assume that there are two materials, one "cleaner", the other of lesser cleanness, i.e. possessing different inclusion severity distributions $F(d)$ per Equation (2.25), but otherwise identical. The probability of encountering inclusions of great severity is higher for the material of lesser cleanness. One can then expect a larger number of effective stress raisers in any given rolling body made of the less clean material. According to Equation (2.24a), the relationship between Phase I life N_I and defect severity is inverse. Consequently, the distribution of lives $G(N)$ in Equation (2.26) will show shorter lives with higher probability for the steel of lower cleanness.

The distribution of cell life will be scaled to lower values. The scale parameter N^* in Equation (2.29) will be a smaller number, i.e. typical rolling body life will be reduced.

d. Two Composing Failure Modes

Assume there is not one family of defects (e.g. inclusions) but two families (e.g. inclusions and surface defects). Assume that these two families of defects operate independently of each other and each has a severity distribution. It is possible to define cells such that they have either one or the other type of defect, but not both. This state of affairs is realistic: failure in rolling contact has been shown to occur either subsurface (e.g. from inclusions) or to start at the rolling surface (due to surface defects). Cells with inclusions are volume elements not extending to the surface, cells with surface defects are surface areas underlaid by a thin "skin" of material. The distribution of the two types of defects is independent.

One obtains two cell life distributions of the type of Equation (2.26): one for inclusions and the other for surface defects:

$$\text{Prob } (N_I \leq N_{I,V}) = G_V (N_I | x, y, z) \quad (2.33a)$$

$$\text{Prob } (N \leq N_{I,S}) = G_S (N_I | x, y) \quad (2.33b)$$

Equation (2.33a) applies to subsurface volumes (three co-ordinates), and Equation (2.33b) applies to surface areas (two co-ordinates).

The probability of rolling body failure can be obtained by considering that survival of the rolling body necessitates survival of all cells from both populations. Accordingly, in analogy to Equation (2.27)

$$H(N_I) = 1 - \left(\prod_V [1 - G_V(N_I)] \prod_S [1 - G_S(N_I)] \right) \quad (2.34)$$

or substituting from Equation (2.29)

$$H(N_I) = 1 - \left([1 - G_V(N_I)]^{m_V} [1 - G_S(N_I)]^{m_S} \right) \quad (2.35)$$

where m_V = number of cells with inclusions

m_S = number of cells with surface defects

One can determine whether the resulting asymptotic distribution for a large number of cells is a Weibull distribution by comparing the characteristics of the cell life distributions G_V and G_S . $H(N_1)$ will be Weibull only if G_V and G_S are sufficiently similar.

e. The Effect of Residual Compressive Stresses in the Material

If there are residual stresses in the highly stressed zone, their hydrostatic pressure component can be expected to influence ductility. Accordingly, γ_I and/or γ_{II} will change. This may have an effect on crack initiation (Phase I), or only on crack propagation (Phase II). Depending on experimental evidence, appropriate modifications can be introduced in the first or the second of Equations (2.24).

It is also possible to account for the effect of the non-hydrostatic component of residual stresses if a reasonable assumption can be made regarding the effect of a static stress component on the microplastic strain ϵ_0 . As said earlier, many fatigue theories assume that this effect is nil.

f. Effect of Hardness

Three parameters: ϵ_0 , σ_y and D enter the matrix strength function γ , Equation (2.11). Clearly, the assumption made by Lundberg and Palmgren that the only stress-variables influencing life are the maximum alternating elastic shear stress and its depth co-ordinate, is tenable only if the material is kept a constant so that σ_y and D do not vary. For different materials, γ will depend on the excess of elastic stress above the microyield stress σ_y , which is known to be related, although not equivalent, to indentation hardness. Thus, a general expression for γ will contain a material strength parameter of the type of hardness.

g. Non-Hertzian Contacts

Practical contacts are non-Hertzian in two respects:

Roller to race contacts are macroscopically non-Hertzian because roller profiles are not correctly approximated by second order surfaces (roller crowning, edge effects.)

All contacts are microscopically non-Hertzian because of the presence of surface asperities, which will be discussed later.

No Hertzian assumption is inherent in any of the derivations given above. Provided that the total strain can be defined point-wise in the contact, it is possible to calculate cell life and hence rolling body life, using the proposed model.

h. Lubrication Effects

The most thoroughly explored lubrication effects to date are those of elastohydrodynamic lubrication, i.e. the transmission of contact pressure via a pressurized viscous film of lubricant. The effects of such a film on life can be treated on two levels: the macroscopic redistribution of pressure brought about by the elastohydrodynamic film can be used to correct the macroscopic strain field in the matrix which defines the function γ . Microscopically, it is possible to describe the influence of a partial elastohydrodynamic film on asperity interactions. The result is a modification of the microstress field in the vicinity of the surface, and a new distribution of cell life in the population of cells containing surface defects. In principle, it is thus possible to calculate these elastohydrodynamic effects on rolling body life.

1. Size Effects

Because the present study is ultimately aimed at improvements in radar antenna bearings, and radar antenna bearings are among the largest made, size effects are of great importance. The model presented offers several clues to size effects.

The self-propagating crack size A_p is defined as the smallest crack which progresses at a rate independent of the original defect. Clearly, the magnitude of this crack must depend on the "effective radius" of the most severe original defect, i.e. it must be related to the volume within which the stress raising effect of the defect is felt. This volume almost certainly is proportional to defect diameter. Accordingly, A_p depends on the size distribution of the original defects. Since A_p enters into the determination of the Phase I life of a defect, it produces a size effect on this life. Absolute bearing size will enter into this effect inasmuch as it influences steel processing practices, and thereby the size distribution of defects.

The other Phase I size effect concerns the cell size introduced in connection with Equation (2.25). Since cells should contain only one defect, their size is related to the defect spacing. This spacing is, again, influenced by absolute bearing

size through the correlation between reduction in cross sectional area undergone by the steel during rolling or forging and defect distribution. Cell size indirectly enters Equation (2.28) because the number of cells present in a given rolling body depends on their size. Of course, the distribution $F(d)$ in Equation (2.25) also depends on cell size so that the effect is not a simple one.

Given a defect size distribution and a cell size, the stressed volume of the rolling body, in terms of multiples of unit cell volume, introduces the volume effect of life, represented by Equation (2.30).

Turning now to size effect on Phase II life, Equation (2.24b) shows two size effects. One is represented by A_c . It is the (hypothetical) effect of absolute crack size on propagation rate. The other, more obvious effect, is represented by the ratio A_c/S and represents the fact that a crack must grow to a certain size with reference to the size of the highly stressed zone before a spall can form (the crack must propagate from the depth of maximum shear stress to the surface, or vice versa, since spalls are typically of a depth comparable to that of the location of maximum shear stresses). This effect is related to absolute bearing size through S .

There are, of course, numerous size effects implied in Equation (2.15) relating total macrostrain to external load and contact geometry. Most of these effects are accounted for in the Lundberg-Palmgren theory. Steel processing effects are also implied in Equation (2.15) through the fact that the yield strength σ_y may well depend on the absolute size of the part. The same applies to the ductility parameter D in Equation (2.14). Clearly, the proposed model offers ample room for the exploration of size effects, providing, of course, that the necessary experimental evidence can be secured.

9. THE STRESS-STRAIN RELATIONSHIP

The relationships between external load, resulting elastic stresses, total strain ϵ_0 defined in Equation (2.15) and microplastic strain ϵ_c defined in Equation (2.9) are key elements of the model. In the present state of the work, no choice has been made among the possible approaches by which ϵ_c can be related to a calculated stress field. However, it seems obvious that high plastic strains will be associated with high shear stresses, and considerable effort was spent in identifying sources of high shear stress under more general conditions than previously available. Three stress analyses, described in subsequent sections, are of importance.

a. Contours of Equal Shear Stress Range in Elliptical Hertz Contact

The proposed model requires a definition of S , the highly stressed zone. A logical definition may be the zone bounded in each cross section by a suitably selected line of equal shear stress range, and extending around the rolling body. Since calculations for such lines are not available from the literature, they have been computed and are shown in this report. Numerical equations and graphical relationships have been developed from which areas S , bounded by selected contours of equal shear stress range, can be obtained. One way in which these contours may be used is by observing that the distorted tubular annulus between any two contours is a zone of roughly equal shear stress range, and therefore presumably equal total macrostrain ϵ_0 . It is, therefore, an area in which the functions γ_I and γ_{II} of Equations (2.24a) and (2.24b) can be considered constant (with, perhaps, a secondary influence of the variable hydrostatic compression on ductility.) A correct way of evaluating Equation (2.27) to obtain rolling body life might be to consider the distribution functions $G_i(N_i)$ identical for cells within each of these annuli of equal shear stress range; to apply Equation (2.28) to one annular set of cells at a time, and then to compute $H(N_i)$ from Equation (2.27) by multiplying the survival probability functions for each annulus.

b. Near-Surface Stresses Due to Asperities

A major result of research into rolling contact fatigue, referenced later in this report, has been the recognition that destructive spalling fatigue failure of rolling bearings is often preceded and precipitated by "surface fatigue", a sequence of events starting with plastic flow in the immediate subsurface layers of material, followed by profuse microcracking at the surface, and eventually leading to spall formation for surface originated cracks. This series of phenomena was found to be lubrication related and has been associated with the thickness of the elastohydrodynamic lubricant film, compared to the roughness of the contacting surfaces. Until recently, it was thought that elastohydrodynamic films prevent surface fatigue by reducing tractive (tangential) forces at the surfaces which, in the absence of a full elastohydrodynamic film, are transmitted between contacting asperities. Because the magnitude of these hypothetical tractive forces cannot be determined generally, and because several pieces of contradictory evidence could not be accommodated, mathematical modeling of surface originated fatigue was at a standstill.

In the last months, it has been proposed that surface fatigue occurrences be explained by purely normal Hertzian pressure, acting in asperity dimensions. The role of the elastohydrodynamic film is, in this view, that it prevents sharp pressure gradients from arising at each separate asperity, whereas these pressure gradients do occur whenever two asperities come into contact through a boundary lubricant film. This hypothesis, if proven correct, can serve as a basis of mathematical modeling for surfaces. To explore its workability, stress analysis was conducted on a model asperity, represented by a prismatic ridge topped by a small radius. It is shown in subsequent sections of this report that high near-surface shear stresses arise under the model asperity when it is pressed against a flat surface (or opposing asperity), and that the magnitude of this shear stress depends on two parameters: the degree of depression of the asperity (in partial elastohydrodynamic contact this is determined by the film thickness to roughness ratio) and by the typical slope of the asperity side. Experimentally, it was shown that asperities on relatively rough, e.g. as-ground, surfaces have steep slopes of the order of 30° , asperities on finely honed surfaces have only slopes of the order of 4° , whereas the finest achievable lapped surfaces of bearing balls have asperity slopes of less than 1° . Calculation shows that the shear stresses under ground asperities reach the yield strength of hard steel for relatively little asperity depression, those under asperities of honed surfaces break into the plastic range at substantial depression values (low film thicknesses) only, whereas those on lapped surfaces should not become plastic under any conditions.

In principle, it should be possible, based on this analysis, to calculate the life of cells immediately adjacent to a rough surface, based on Equations (2.24a) to (2.24c).

The sequence of events leading to surface originated failure is this: Near-surface plasticity occurs under contacting asperities, and eventually causes microcracks at the surface to a shallow depth of several hundred microinches. The life prior to the formation of these cracks can be calculated from Equations (2.24a) to (2.24b). The formation of surface fatigue cracks (not to be confused with deep cracks which may originate at the surface and will be discussed subsequently) represents a failure phase prior to Phase I. Experimentally, it is a distinct phase both more widespread and more rapid of progression than the subsequent spalling. The surface equipped with microcracks can be treated as a source of Phase I crack generation because the microcracks are so small and shallow that they do not show an apparent tendency

to interact. The near-surface volume of material thus acquires a large population of new defects, viz. the surface fatigue cracks. One can thus enter Equations (2.24a) to (2.24c) again and calculate spalling fatigue life for this population of defects.

c. Stresses Under a Surface Imperfection

Aside from the generalized roughness of the surface, there are localized imperfections, generally of the scratch or furrow type, on every practical bearing surface. It has been recognized recently that such furrows are points of origin for spalls. They represent a second population of defects (alluded to previously) which competes with the subsurface defects of the inclusion type to generate spalling failures. Whereas the Lundberg-Palagren theory is capable of predicting life for failures originating at inclusions by using the macroscopic shear stress range in lieu of the total strain parameter ϵ_0 , and by assuming that the stress raising properties of the inclusions (their severity) are a material constant, the same approach is not feasible for surface originated failures because there is no high shear stress at the surface of a Hertzian contact between ideally smooth surfaces in the absence of traction. This difficulty disappears, however, if it can be shown that high shear stresses arise in the vicinity of localized surface imperfections when they enter the Hertzian contact area. The stress analysis given in a later section of this report has shown that this is the case. The surface stress field was calculated under a furrow type imperfection, represented by a long prismatic depression in a plane. The depression has two rounded edges and sufficient depth to prevent contact at the bottom below the rounded edges. High shear stresses were found under the rounded edges, and their relationship to the geometric parameters of the defect were determined. This calculation permits the assignment of a severity value to a surface furrow.

d. Plastic Strain Determination

According to Equation (2.8), the plastic strain physically relevant to crack propagation rate is ϵ_a , the strain at the site of the crack front. It is convenient and simple to visualize this strain as the result of a pre-existing total strain ϵ_0 in the undisturbed matrix on which operates the strain raising effect of the defect. This description has been used in the preceding discussion. However, the critical (maximum) plastic strain actually existing at a crack front does not, in general, depend on a single strain scalar of the macroscopic strain field.

The actual strain field in the vicinity of an imperfection is quite complex and, even if the macrostrain in the vicinity of the inclusion can be considered uniform, depends on all three principal strains. Thus, generally

$$\epsilon_c = f(\epsilon_{1.0}, \sigma_y, \theta_{1.0}) \quad (2.36)$$

with different defect severity factors applicable along each principal strain axis.

The preceding paragraphs contain two examples for the calculation of a defect-induced maximum shear stress, based on defect geometry (i.e. severity) and assuming a simple macrostress field. Given this maximum shear stress at the defect, it may be possible to put forth a simple model of plastic flow at the defect, yielding ϵ_c .

Whether it will be necessary for the application of the model to proceed to the actual determination of the plastic strain magnitude at defects, or whether calculation of a maximum shear stress, attached to known defects, is a sufficient refinement of the Lundberg-Palmgren macroshear stress criterion to accommodate currently available experimental evidence is a subject for future investigation.

10. OUTLOOK

A mathematical model for the prediction of the life of rolling contacts has been proposed, based on a concept of crack propagation from pre-existing defects. Numerous parameters characterizing the material and geometry of the contact and the operating conditions are incorporated in the model, including load, lubrication, matrix strength, defect population, and defect severity.

Further work now underway is aimed at closer definition of as many of the influential variables as feasible, considering currently available information. Starting with the Lundberg-Palmgren prediction of bearing life, the incorporation of experimentally documented effects not previously accounted for, will be attempted, and prediction will be correlated with life test results.

While an all-inclusive and completely verified life prediction formula is not in sight, it appears likely that a worthwhile improvement over past prediction methods can be achieved for those design, manufacturing, and operating conditions of rolling bearings where experimental data suffice for the determination of parameter values required in the formula.

SECTION III

SYNOPSIS OF LUNDBERG-PALMGREN THEORY

The first successful systematic attempt to treat rolling bearing fatigue life analytically was made by G. Lundberg and A. Palmgren in 1947, under sponsorship of AB & SKF, Gothenburg, Sweden (1)*. This work was further pursued with a special view towards roller bearing fatigue life and reported in (2).

This work of Lundberg and Palmgren is the basis for the life prediction method standardized by the Anti-Friction Bearing Manufacturers Association (AFBMA) (2), the ASA (4) and ISO (5) for computing rolling bearing life.

1. FAILURE PROBABILITY DISTRIBUTION

The following is Lundberg-Palmgren's development of the probability $S(L)$ that a bearing ring will survive to life L . S , in today's parlance is termed a reliability function. Its arithmetic complement $F(L) = 1 - S(L)$ is the cumulative form of a failure probability function.

Let $\lambda(N)$ be a hypothetical function which describes the fatigue "condition" of a differential volume ΔV of a ring or rolling element material at a depth z below the rolling surface after N cycles of stress are endured by that volume. Let $\Delta\lambda(N)$ represent the change in the material fatigue condition within a small number of additional cycles ΔN .

The probability that the volume develops a "failure" (i.e., a crack) in the interval $(N, N+\Delta N)$ is taken to be

$$f(\lambda(N)) \cdot \Delta\lambda(N) \cdot \Delta V \quad (3.1)$$

The probability of not developing a failure (crack) in this interval is the arithmetic complement, of Equation (3.1), i.e.

$$1 - f(\lambda(N)) \cdot \Delta\lambda(N) \cdot \Delta V$$

*Numbers in parentheses refer to the References at the end of this report

If $\Delta S(N)$ denotes the probability that the volume element endures N cycles without cracking, then by the product law of probability theory,

$$\Delta S(N+\Delta N) = \Delta S(N) \cdot [1 - f(\lambda(N)) \Delta \lambda(N) \cdot \Delta V] \quad (3.2)$$

Dividing by ΔN and re-arranging gives

$$\frac{\Delta S(N+\Delta N) - \Delta S(N)}{\Delta N} = -f(\lambda(N)) \cdot \Delta S(N) \cdot \frac{\Delta \lambda(N)}{\Delta N} \cdot \Delta V \quad (3.3)$$

Taking limits as $\Delta N \rightarrow 0$ gives, by definition of the derivative,

$$\frac{d \Delta S(N)}{\Delta S(N) dN} = -f(\lambda(N)) \cdot \frac{d \lambda(N)}{dN} \cdot \Delta V \quad (3.4)$$

Or since $-\frac{d \Delta S(N)}{dN \Delta S(N)} = \frac{d}{dN} \log \frac{1}{\Delta S(N)}$ one has:

$$\frac{d}{dN} \left(\log \frac{1}{\Delta S(N)} \right) = \Delta V \cdot f(\lambda(N)) \frac{d \lambda(N)}{dN} \quad (3.5)$$

Integrating both sides from 0 to N and remembering that $\Delta S(0) = 1$ gives:

$$\log \frac{1}{\Delta S(N)} = \Delta V \cdot \int_0^N f(\lambda(N)) \frac{d \lambda(N)}{dN} dN = \Delta V \cdot G(\lambda(N)) \quad (3.6)$$

where

$$G(\lambda(N)) = \int_0^N f(\lambda(N)) \frac{d \lambda(N)}{dN} dN$$

The probability that the entire volume V will endure N cycles is, (assuming independence of the volume elements), the product of the probability that the individual elements will endure.

In equation form,

$$S(N) = \Delta S_1(N) \cdot \Delta S_2(N) \cdot \dots \cdot \Delta S_p(N) \quad (3.7)$$

so that:

$$\log \frac{1}{S(N)} = \sum_{j=1}^p \log \frac{1}{\Delta S_j(N)} \quad (3.8)$$

Using Equation (3.6) gives:

$$\log \frac{1}{S(N)} = \sum G(\lambda(N)) \Delta V \quad (3.9)$$

As $\Delta V \rightarrow 0$ the summation becomes an integral and one has

$$\log \frac{1}{S(N)} = \int_V G(\lambda(N)) dV \quad (3.10)$$

Changes in the material condition at a depth z are taken to depend on:

- a) macrostress $\tau(N)$ which is most dangerous from the point of view of material fatigue. This stress is, based on observations, taken to be the alternating shear stress in the direction parallel to rolling.
- b) the material condition $\tau(N)$.
- c) the depth z

Thus the change in material condition $d\lambda(N)$ for a small number of load cycles dN may be written:

$$\frac{d\lambda(N)}{dN} = J(\lambda(N), \tau(N), z) \quad (3.11)$$

The reliability function $S(L)$ is uniquely determined from Equations (3.10) and (3.11) if the functions G and J are known.

It is postulated that these functions are power functions as follows:

$$G(\lambda(N)) = k(\lambda(N))^g \quad (3.12)$$

$$J(\lambda(N), \tau(N), z) = (\lambda(N))^{-j} \cdot K(\tau(N), z) \quad (3.13)$$

Where g and j are constants.

From Equations (3.11) and (3.13)

$$(\lambda(N))^j d\lambda(N) = K(\tau(N), z) dN \quad (3.14)$$

From Equations (3.10) and (3.12), the condition $S(N=0) = 1$ requires that $\lambda(N=0) = 0$

Integration of Equation (3.14) between the limits of 0 and N therefore gives:

$$\frac{(\lambda(N))^{j+1}}{j+1} = \int_0^N K(\tau(N), z) dN \quad (3.15)$$

Introducing $j+1 = \frac{q}{\theta}$ gives:

$$\lambda(N) = \left[\frac{q}{\theta} \cdot \int_0^N K(\tau(N), z) dN \right]^{\frac{\theta}{q}} \quad (3.16)$$

Using Equations (3.12) and (3.16) in Equation (3.10) gives:

$$\log \frac{1}{S(N)} = k \left(\frac{q}{\theta} \right) \cdot \int_V \left[\left(\int_0^N K(\tau(N), z) dN \right)^{\frac{\theta}{q}} \right] dV \quad (3.17)$$

If the amplitude $\tau(N)$ is independent of N, equation (3.17) becomes:

$$\log \frac{1}{S(N)} \sim N^{\frac{q}{\theta}} \int_V K^{\frac{\theta}{q}}(\tau, z) dV \quad (3.18)$$

In a Hertz stress field of given geometry, the value of $\tau(z, x)$ at depth z and coordinate position x measured perpendicular to the rolling direction from the midpoint of the contact ellipse is given by:

$$\tau = \tau_0 \cdot f \left(\frac{x}{a}, \frac{z}{a} \right) \quad (3.19)$$

where τ_0 and z_0 represent the maximum shear stress amplitude and a is the half axis of the contact ellipse in the direction across the raceway.

Using Equation (3.19) in Equation (3.16) gives:

$$\log \frac{1}{S(N)} \sim N^c \int_0^{l_0} \int_0^\infty \int_0^\infty K^e \left(\tau_0 f \left(\frac{x}{a}, \frac{z}{z_0} \right), z \right) dx dz dy$$

where l_0 is the length of the raceway

Introducing the change of variables $u = \frac{x}{a}$, $v = \frac{z}{z_0}$

$$\text{gives: } \int_V K^e(\tau, z) dV = a \cdot z_0 \cdot l_0 \cdot \varphi(\tau_0, z_0) \quad (3.20)$$

where

$$\varphi(\tau_0, z_0) = \int_0^\infty \int_0^\infty K^e \tau_0 f(u, v), z_0 v \, du \, dv$$

Using Equation (3.20) in (3.18) gives:

$$\log \frac{1}{S(N)} \sim N^c \cdot a \cdot z_0 \cdot l_0 \cdot \varphi(\tau_0, z_0) \quad (3.21)$$

The relationship $\log \frac{1}{S(N)}$ or:

$$F(N) = 1 - S(N) = 1 - \exp - \left(\frac{N}{N_0} \right)^c$$

defines the probability distribution of bearing lives when many identical bearings are operated under identical conditions. This distribution is today known as a (two parameter) Weibull distribution.

It is postulated that:

$$\varphi(\tau_0, z_0) = \tau_0^c \cdot z_0^{-h} \quad (3.22)$$

where c and h are unknown positive constants which satisfy the intuitively reasonable relationship that:

$$S(N) = 1 \quad \text{if} \quad \tau_0 = 0 \quad \text{or} \quad z_0 = \infty.$$

Using Equation (3.22) in Equation (3.21) gives:

$$\log \frac{1}{S(N)} \sim \frac{\tau_0 c N^{\frac{c}{h-1}} a l_0}{z_0} \quad (3.23)$$

At this point in the development, the quantities τ_0 and z_0 are related through Hertz's equations for the contact of elastic bodies, to the bearing geometry (rolling body diameter D_R , raceway diameter D_n and raceway curvature in a plane perpendicular to the rolling direction) the elastic constants of the material (Young's modulus and Poisson's ratio) and the contact force Q . The contact force is here assumed constant and independent of position on the ring. The notion of an equivalent load is later introduced to account for those cases (e.g. the stationary outer ring of a radially loaded bearing) where load varies with ring position.

Also introduced is the number of contact cycles per revolution u defined through the relationship,

$$N = uL \quad (3.24)$$

where L is bearing life in millions of revolutions.

For point contact, the half-width of the contact ellipse is replaced by its expression in terms of load from the Hertzian equations.

In the 1947 treatment, the half-width a for line contact is taken equal to three quarters of the roller length. The 1952 extension of the work deals specifically with roller bearings and introduces a new expression relating stress and effective roller length.

For point contact, these substitutions lead (for the contact at either bearing ring) to:

$$\log \frac{1}{S} \sim \phi \cdot D_R^{\frac{5c+h-5}{c-h+5}} \cdot Q^{\frac{c-h+1}{c}} \cdot L^c \quad (3.25)$$

where ϕ contains bearing geometrical parameters to powers which are linear combinations of the exponents c , h and e . The rolling body diameter appears, in the function ϕ , only in the form D_R/d_m , where d_m is the pitch diameter of the rolling elements. The survival probability given by Equation (3.25) is that of the ring-rolling element contact. Both the ring and the rolling element are considered equally likely to fail.

Q_c is defined to be the ring load Q for which the life for the fixed value $S = 0.9$ is one million revolutions.

The life L_{10} for the same survival probability under a load Q is found from:

$$Q^{\frac{c-h+1}{s}} \cdot L_{10}^s = Q_c^{\frac{c-h+1}{s}}$$

or

$$L_{10} = \left(\frac{Q_c}{Q}\right)^{\frac{s-h+1}{s}} = \left(\frac{Q_c}{Q}\right)^p \quad (3.26)$$

where $p = \frac{c-h+1}{s}$

The rolling element load Q is proportional to the bearing load F , and hence, C_i or C_o the dynamic capacity of the (inner or outer ring) contact, defined as the bearing load for which the ring will endure one million revolutions with a survival probability of 0.90, is proportional to Q_c . Accordingly, from Equation (3.26):

$$L_{10} = \left(\frac{C}{F}\right)^p \quad (3.27)$$

2. EQUIVALENT LOAD

From Equation (3.25) the logarithm of the reciprocal of the survival probability is proportional to Q^W , where $W = \frac{c-h+1}{2}$ and Q may vary with angular position ψ on the bearing ring, i.e. $Q = Q(\psi)$.

In view of Equation (3.25), the summing over the complete ring, of probabilities that individual ring segments of angular length will survive, gives rise to:

$$\log \frac{1}{S} \sim \int_0^{2\pi} Q^W(\psi) L^s \frac{Dn}{s} d\psi = Q_0^W \pi Dn L^s \quad (3.28)$$

wherein

$$Q_0 = \left(\frac{1}{2\pi} \int_0^{2\pi} Q^W(\psi) d\psi\right)^{\frac{1}{W}} \quad (3.29)$$

Q_e is thus the equivalent constant ring load for which the ring contact survival probability is the same as it is under the spatially variable loading.

A continuous load distribution, the components of which integrate to equal the applied radial and axial load, is introduced as an approximation of the discontinuous rolling element loads and the integral of Equation (3.29) is evaluated as a function of the ratio of radial and axial load.

3. CAPACITY OF A COMPLETE BEARING

The probabilities S_i and S_o that the bearing inner and outer ring contacts, respectively, will endure beyond a life L under a bearing load F are given by :

$$\log \frac{1}{S_i} = k_i F^W L^e \quad (3.30)$$

$$\log \frac{1}{S_o} = k_o F^W L^e \quad (3.31)$$

where k_i and k_o are constants of proportionality.

The probability S that the complete bearing survives to life L is the product of S_o and S_i hence :

$$\log \frac{1}{S} = (k_i + k_o) \cdot F^W \cdot L^e \quad (3.32)$$

By definition when in Equations (3.30) to (3.32) the survival probability is taken equal to 0.9 and the life $L = 1$ the loads will be equal to the respective dynamic capacities. Thus,

$$\log_{0.9} \frac{1}{S} = k_i C_i^W = k_o C_o^W = (k_i + k_o) \cdot C^W$$

from which it is found that :

$$\frac{1}{C^W} = \frac{1}{C_i^W} + \frac{1}{C_o^W} \quad (3.33)$$

4. DETERMINATION OF CONSTANTS IN THE LIFE FORMULA

a. Point Contact

Taking logarithms of Equation (3.25) gives $\log \log \frac{1}{L} = e \log L$ plus terms not containing L. e is the parameter characterizing the dispersion of lives L of a sample of identical bearings operated under identical conditions. e is today generally designated as the "Weibull slope" of the life distribution. Bearing life test results yield an estimate of e as the slope of the straight line which is obtained when the percentage of unfailed bearings is plotted against life on paper so ruled that the ordinate is proportional to $\log \log \frac{1}{L}$ and the abscissa to $\log L$ (such a diagram is called a Weibull plot).

From Equation (3.27) it is seen that the exponent $p = \frac{c-h+1}{e}$ may be determined from tests conducted under various loads F as the slope of the line obtained when L and F are plotted on log-log paper.

Finally, from Equation (3.28) the exponent $\frac{c+h-6}{c-h+1}$ may be found as the slope of the line on logarithmic coordinates of the value of Qc (or C) plotted against roller diameter Da. The tests must, in view of the fact that Da/dm appears in the function Φ , be run for constant values of Da/dm.

The results of these three test series are then solved simultaneously to give c, h and e.

b. Line Contact

The treatment of the line contact problem given in the 1947 work (1) is amplified and revised in the 1952 publication, (2) which is devoted exclusively to roller bearings. In this treatment, the contact load is taken to be proportional to the 1.1 power of the deformation; in the earlier work a linear relationship was assumed for line contacts.

It was found that for line contact, the exponent p in the load-life relationship is 4 rather than 3 as in the point contact case.

It is possible for some roller bearings to have point contact within one range of loads and line contact within another. It is even possible, under some conditions in a roller bearing, for the roller to make line contact with one raceway and point contact with the other.

As a compromise made for simplicity's sake a conservative compromise exponent of $p = 10/3$ is assumed. The capacity is thereby modified by a factor which is calculated so that when line contact prevails at both contacts the error in using the exponent $10/3$ rather than 4 is made as small as possible over the most frequently used life range.

In computing the capacity of roller bearings, a reduction factor ($\lambda < 1$) is introduced as an attempt to account for the stress concentration which may occur at roller ends as well as the effect of inexactlly aligned rollers.

Another reduction factor ($\beta < 1$) is introduced for thrust loaded bearings to account for the effect of the greater sliding undergone by the rolling bodies. In thrust loaded bearings, roller loading is not cyclic, but remains virtually constant, exacerbating the problem of rolling element sliding and traction forces.

5. FACTORS OMITTED FROM LUNDBERG-PALMGREN THEORY

Lundberg and Palmgren, in their preface, acknowledge the absence, from their development, of several factors known to affect endurance life. Specifically cited as areas for future investigation are:

- a) Effect of stress history
- b) Work hardening
- c) Lubricant effect on stress distribution
- d) Effect of residual stresses (set up in the rings by interference fits)
- e) Effect of edge loading in line contact.(2)
- f) Effect of radial load on contact angle in ball bearings.
- g) Effect of indexing of the ball rotational axis which, since any given point on the ball is cyclically stressed for only part of the bearing life, results in a lesser number of failures initiating in the ball material than in the ring material
- h) Effect of surface traction (in thrust loaded bearings)
- i) Effect of geometrical imperfections on load distribution.

Many other life factors are, of course, absent from the Lundberg and Palmgren treatment without having been specifically enumerated by these authors.

SECTION IV

VARIABLES AND MECHANISM OF ROLLING CONTACT FATIGUE

1. FAILURE MECHANISMS (GENERAL)

The definition of functional failure in rolling bearings depends on the application. Except in instrument bearings in which torque is often the main criterion in defining 'failure', the most common definitions of functional failure involve visible damage to the rolling tract. As an example of specific interest, the failure of many large radar antenna pedestal bearings has been found to fail under this definition involving smearing, spalling and surface distress in the rolling track. From these broad concepts of failure, several groups of reasonably well defined changes in the rolling bearings can be identified which represent failure modes (6).

The failure of a particular bearing is a consequence of several competing modes of failures classified in Table 1. Among the modes of bearing failure, the present study pertains to the prediction of contact fatigue life (mode 3 of Table 1). This is justified since wear and plastic flow (failure modes 1 and 2 of Table 1) can be eliminated in most rolling bearings by suitable design and operation controls whereas cracking (failure mode No. 4) is not, specifically, a rolling contact failure. On the other hand, all loaded rolling contacts are eventually subject to fatigue failure.

The functional failure due to contact fatigue is characterized by the emergence of a fatigue crack causing removal of a sizable piece of metal from the rolling surface (spalling). The spalling itself is preceded by the initiation and propagation of one crack out of possibly several, which arise in the rolling element through stress cycling.

The present study considers that the fatigue process in rolling contact is associated with the generation and propagation of fatigue cracks in the bearing material and the fatigue life is taken as the number of cycles at which fracture (spalling) occurs. This is a concept taken from numerous fatigue studies most of which are not for rolling contacts. Fatigue theories related to fatigue crack behavior in materials can be divided into two distinct approaches: one is the so-called "engineering" (or "phenomenological") approach (e.g., Manson (7), Coffin (8) and Morrow (9) on low cycle fatigue and Dugdale (10), Paris (11) on sheet specimens) which is interested in quantitative treatment

Table 1

FAILURE MODES OF ROLLING CONTACTS

- | | |
|-----------------------|---|
| 1. Wear type failures | 1.1 Surface removal |
| | 1.1.1 Removal of loose particles ("Wear") |
| | 1.1.2 Chemical or electrical surface removal |
| | 1.2 Cumulative material transfer between surfaces ("Smearing") |
| 2. Plastic flow | 2.1 Loss of contact geometry due to cold flow |
| | 2.2 Destruction by material softening due to unstable overheating |
| 3. Contact fatigue | 3.1 Spalling |
| | 3.2 Surface distress |
| 4. Bulk failures | 4.1 Overload cracking |
| | 4.2 Overheat cracking |
| | 4.3 Bulk fatigue |
| | 4.4 Fretting of fitted surfaces |
| | 4.5 Permanent dimensional changes |

of macroscopic variables; and the other is the metallurgical (or microscopic) approach (e.g. Laird (12), Wood (13) and Grosskreutz (14)) which deals with the basic microscopic phenomena such as dislocations, slipbands and microcrack initiation in material under cyclic strain.

Both approaches teach that plastic occurrences, either macroscopic (such as in low cycle fatigue) or microscopic (in high cycle fatigue (19)) are a source of crack generation and propagation under fatigue loading.

Many attempts have been made to relate quantitatively the fatigue life to the magnitude of cyclic plastic strain. In published literature, quantitative treatment has been made by Manson et. al. regarding crack generation and propagation in notched specimens (15) without taking into account the microscopic plastic behavior. The prediction of the life for a complete machine component has been found very difficult because of the complexities in geometry and loading. Peterson in (16) has studied quantitatively the fatigue problem in turbine components associated with cracking due to thermal cycling by determining cyclic plastic strain corresponding to the condition of operation, taking account of stress concentrations. Rice and Brown (41) have attempted to interpret the fatigue fracture of loaded structural elements in terms of fatigue crack propagation and to formulate the fundamentals of a statistical theory in these terms.

Quantitative treatment as described above has to rely on many assumptions and phenomenological conclusions from specimen testing. The basic physical mechanisms of crack generation and propagation in materials under fatigue loading is of course a subject of much fundamental research in metallurgy. More specifically, cracking at non-metallic inclusions has been a subject of considerable interest in basic metallurgical studies but quantitative treatment correlating various macroscopic quantities has not been available for engineering design application.

From the reviewed literature, the following can be gleaned:

In spite of the great physical complexity involved in the process of crack initiation from inclusions, it is well recognized that non-metallic inclusions are stress (or strain) raisers and potential sites of plastic deformation. The above described literature points to a close association between fatigue life and the magnitude of the cyclic plastic strain. It may be possible to build a fatigue life model by predicting

the number of stress cycles needed to generate a self-propagating crack as a function of the magnitude of (cyclic) plastic strain at the stress raisers. The plastic strain is, in turn, dependent on the macro-strain field and the inherent characteristics of the stress raiser. Such an engineering approach has been used by Manson & Hirschberg (15) and by Peterson (16) to relate crack initiation life to the stress concentration factor at a notch.

The quantitative dependence of fatigue life on fatigue ductility and plastic strain can be described as follows:

Coffin (8) and Manson (26) has investigated the fatigue behavior in a wide range of materials by strain cycling at constant strain amplitude about zero mean strain. In the low cycle region where failure occurs in 10^5 cycles or less, the fatigue life is found to be related to the plastic strain amplitude by an equation of the following form for all materials:

$$\Delta \epsilon_p = H (2N_f)^z \quad (4.1)$$

where $\Delta \epsilon_p$ = total plastic strain amplitude

N_f = number of cycles to failure

H and z = constants

Manson (15) has extended the treatment of strain cycling data to cover the entire range of fatigue lives, from the low cycle region where strains are predominantly plastic to the high cycle regions where strains are predominantly elastic. Manson's equation relating total strain amplitude and cycles to failure is given by:

$$\Delta \epsilon_T = \Delta \epsilon_p + \frac{\Delta \sigma}{E} = H (2N_f)^z + L (2N_f)^W \quad (4.2)$$

where $\Delta \epsilon_T$ = total strain amplitude

$\Delta \epsilon_p$ = plastic strain amplitude

$\frac{\Delta \sigma}{E} = \Delta \epsilon_e$ = elastic strain amplitude

H and z = the same constants given in Equation (4.1)

L = a constant, (see Figure 33)

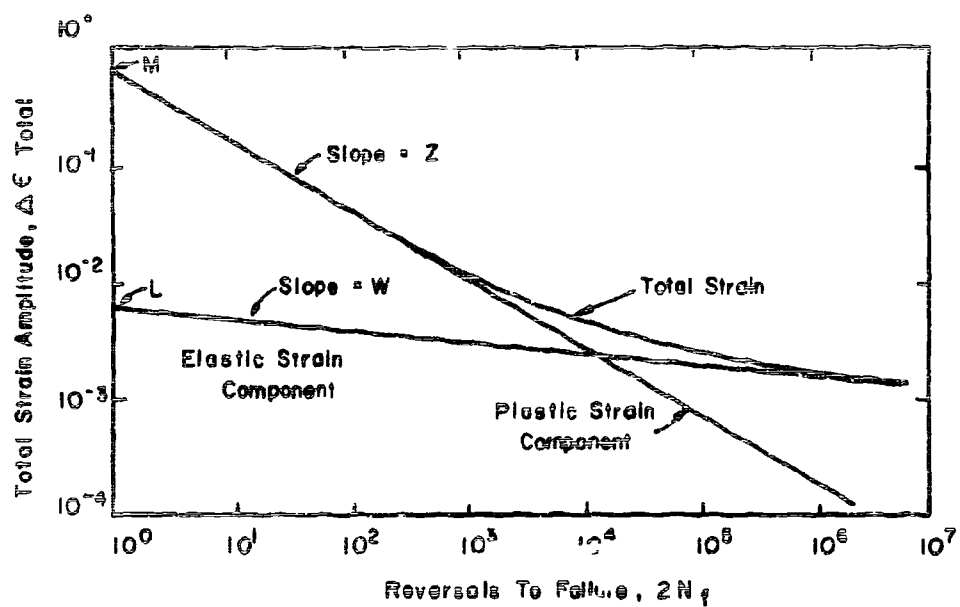


Figure 33. Plastic, Elastic and total Strain vs. Fatigue Life

where E = Young's modulus

and w = a constant, (see Figure 33)

This general equation covering the entire range of cycling lives is shown in Figure 33 and represents a curve which can be expressed as the sum of the two straight line components. The steeper component to the left represents the plastic strain component while the elastic strain component is represented by the line with shallow slope.

The fatigue behavior of a material can therefore be characterized by the four constants of Equation (4.2) i.e. M , z , L , and w .

The plastic portion of the curve reflects the matrix properties of the material while the fact that the elastic portion of the curve slopes at all is due to the presence of defects. This latter conclusion follows when one considers that, in purely elastic reversals of strain, nothing changes in the metal to cause failures. Since failure does occur in the nominally elastic stress range, it must be due to localized yielding, probably in the vicinity of defects.

Thus, failures in the elastic region are defect dominated. The plastic region of the curve, however, is derived from tests where large plastic strains develop throughout the bulk material. While it is true that defects will exert some influence on this portion of the curve, properties of the matrix, e.g. strain hardening and ductility, will tend to dominate the results. The predominance of matrix properties in the plastic region has been confirmed by matrix properties in the plastic region and has been confirmed by Morrow (49). He states that little, if any, difference can be found between clean and dirty steels of similar processing and composition when tested in the plastic strain region.

Thus the plastic portion of the curve is suited to describe the ductility properties of candidate materials. The intercept of the plastic strain line at $2N_f = 1$ (M of Equation (4.2)) and the slope of the line (z) are considered fundamental fatigue properties by Manson and Morrow. They have been given the names "fatigue ductility coefficient" and "fatigue ductility exponent" by Morrow (16).

It is apparent from inspection of Figure 33 that an increase in the intercept value (M), for a fixed slope (z), will translate the line upwards, giving increased fatigue lives. Similarly a decrease in the slope at a fixed intercept will rotate the line upwards with the same effect on life.

2. VARIABLES AFFECTING FATIGUE IN ROLLING CONTACT

Many variables have been found to affect fatigue life of rolling bearings. Based on the current knowledge of the operation, lubrication, metallurgy and theory of fatigue fracture of rolling bearings, Table 2 presents an extensive list of "external" variables which control the fatigue life of a rolling contact. These variables can be grouped into categories as shown below:

a. Material Variables

- 1) Factors that effect the material yield strength and ductility: material analysis, hardness, soft constituents (retained austenite, ferrite, bainite), grain size, alloy segregation.
- 2) Factors that modify the applied stress field: residual stresses originating from heat treatment, grinding and plastic flow during operations.
- 3) Material imperfections acting as stress raisers such as non-metallic inclusions or imbedded micro-cracks. (Lenticular carbides which develop during bearing operation may also serve as localized stress raisers.)
- 4) Modulus of elasticity and Poisson ratio, as they affect material rigidity and control the maximum stress level.

b. Surface Micro-Geometry Variables

- 1) Surface imperfections such as grinding furrows, scratches and dents which serve as surface stress raisers.
- 2) General surface roughness as induced by methods such as grinding, honing and lapping, and as characterized by a) amplitude, e.g., the composite r.m.s. surface roughness defined as the square root of the sum of the squares of the r.m.s. roughnesses of two surfaces rolling together and b) a plasticity parameter of the asperities, such as their typical slope.
- 3) Compositional and hardness properties of the surface.
- 4) Surface coatings.

<u>Material Variables</u>	<u>Surface Micro-geometry Variables</u>	<u>Design Variables</u>	<u>Operating Variables</u>
1. Material Strength & Ductility a. hardness b. alloy segregation c. grain size and orientation d. soft constituent 2. Residual Stress 3. Material Impurities a. inclusions (non-metallic) b. carbides c. imbedded micro-cracks 4. Elastic Modulus	1. Surface Imperfections a. grinding furrows b. sc. scratches c. dents 2. Surface Finishing a. composite r.m.s. b. asperity slope 3. Surface Coating 4. Material Composition (hardness at surface)	1. Rolling Track Length (pitch dia) 2. Groove Conformity 3. Ball Diameter 4. Roller Diameter 5. Roller Length 6. Crown & Blending	1. Load 2. Speed 3. Temperature 4. Spin-roll Ratio 5. Lubricant a. viscosity b. boundary lubricating ability

c. Design Variables Related to Design Dimensions

1) Rolling body design such as rolling track length (pitch diameter), groove conformity, ball (roller) diameter and number of rolling elements, contact angle, roller length and crowning and blending radii, accuracy parameters influencing load distribution, dimensions controlling sliding, etc. These variables affect stress distribution, stressed volume and number of stress cycles.

2) Cage and auxiliary parts design as it affects rolling element forces.

d. Operating variables such as:

1) Load magnitude and direction as it affects stress level and stressed volume.

2) Speed, as it affects lubrication (e.g., EHD film thickness).

3) Temperature, as it affects lubrication and material strength.

4) Lubricant properties, a) rheological properties, e.g. (1) viscosity and (2) boundary lubricating ability (chemistry, additives, etc.)

5) Atmospheric conditions, contaminants, etc.

The above listed variables are believed to be applicable for rolling bearings of all types and sizes, including large radar antenna pedestal bearings.

3. MECHANISMS OF FAILURE IN ROLLING CONTACT

The present study is based on the concept that there are two competing fatigue mechanisms operating to promote spalling failure in rolling contact, namely, sub-surface initiated failure and surface initiated failure. In the former the cracks are generated from highly stressed sub-surface regions, whereas in the latter, fatigue cracks are generated from the rolling surface. Recent findings of metallurgical investigations (6, 17, 18) support this classification of failure mechanisms. Structural changes have been observed in bearing material under

repeated loading (17, 18). These structural changes encompass (1) plastic deformation around sub-surface weak points (e.g., "butterflies"), which indicates that the micro-defects act as stress raisers, (2) the formation of plastic deformation bands ("white etching areas") in the sub-surface high shear stress zone signifying the existence of a region within which the yield limit is exceeded for some volumes and (3) near-surface micro-plastic occurrences at asperities and at surface defects.

Among the structural changes described, the formation of "butterflies" around inclusions has long been closely associated with sub-surface crack generation (1, 18). A similar association between structural changes and cracking has recently been demonstrated for the near surface changes (3 above) and with some indirection, for the generalized changes (2 above).

The following summarizes current knowledge of the two fatigue failure mechanisms described:

a. Sub-surface Failure

The concept of sub-surface failure is well covered in the Lundberg-Palmgren theory (1) in which the fatigue crack is assumed to start from weak points, e.g., slag inclusions. In current terminology, these weak points give rise to the local stress concentrations and plastic flow in the surrounding matrix material. According to Lundberg-Palmgren the site of crack generation is the zone of high shear stress in the sub-surface region of a rolling element. Fatigue cracks are found to start at weak volumes and to grow under repeated loads until an advanced stage of fatigue cracking is reached. This stage is characterized by the distortion of the macroscopic stress field due to the cracks. Eventually the destructive process of spalling sets in at one (or more) locations causing removal of a sizeable piece of metal from the surface.

It is recognized that the formation of localized plastic deformation around inclusions (i.e., the "Butterfly" structure) is stress dependent. In (1), Lundberg and Palmgren hypothesized that the actual stresses at sub-surface "weak" points of the material are proportional to the magnitude of the macroscopic stress and are modified by factors dependent on the size and microscopic shape of each of the "weak points" (inclusions). Recent metallurgical investigation has

supported this hypothesis, e.g., Littman and Widner (16) pointed out that each non-metallic inclusion has its own stress concentration factor which depends on the size, shape, physical and mechanical properties of the inclusion; Martin and Eberhardt (17) have shown that "Butterfly" structure can occur at locations where the calculated "macro" stresses are below the threshold level of plastic flow indicating that inclusions are stress raisers.

b. Surface Distress and Surface Initiated Fatigue

The above described sub-surface failure will occur for any lubrication and for smooth surfaces (thick elasto-hydrodynamic films). Under such conditions, life to failure appears to be independent of both lubricant and surface texture. There is, however, an altogether different mode of rolling contact fatigue referred to as surface distress (6, 20). Although surface distress does not imply raceway destruction, it is a precursor and often a precipitator of spalling failure, apparently by generating severe surface-adjacent defects (about 100 μ in. depth) which then serve as crack initiation points (6). Metallurgical evidence has shown that this "surface distress" involves a type of plastic working and subsequent fatigue microcracking of the immediate surface-adjacent layers of the metal. The occurrences of near-surface plastic deformation are of the following two kinds:

- a) Wide-spread microplastic flow at surface asperities caused by asperity interaction;
- b) localized plastic flow at surface imperfections, such as grinding furrows, scratches and debris dents.

Regarding the occurrence of near-surface plastic deformation, Tallian (6) has suggested that it is controlled by the elastohydrodynamic lubricant film and by surface roughness parameters, e.g., the composite surface roughness r.m.s. and a typical asperity slope angle. In (6) it is hypothesized that the near surface plastic occurrences are a consequence of severe interaction of asperities. The approach of the contacting asperities depends on the mean EHD film thickness defined as the distance between the mean line of the two asperity profiles. The degree of approach determines the severity of the plasticity in the asperity.

It has also been reported, based on recent investigations of failed bearings (17), that for rolling bearings made of improved clean steel (obtainable by the vacuum melting process), a larger fraction of failures is found to be associated with the surface defects. This is because of

THE INTERDEPENDENCE OF VARIABLES AFFECTING CONTACT FATIGUE LIFE

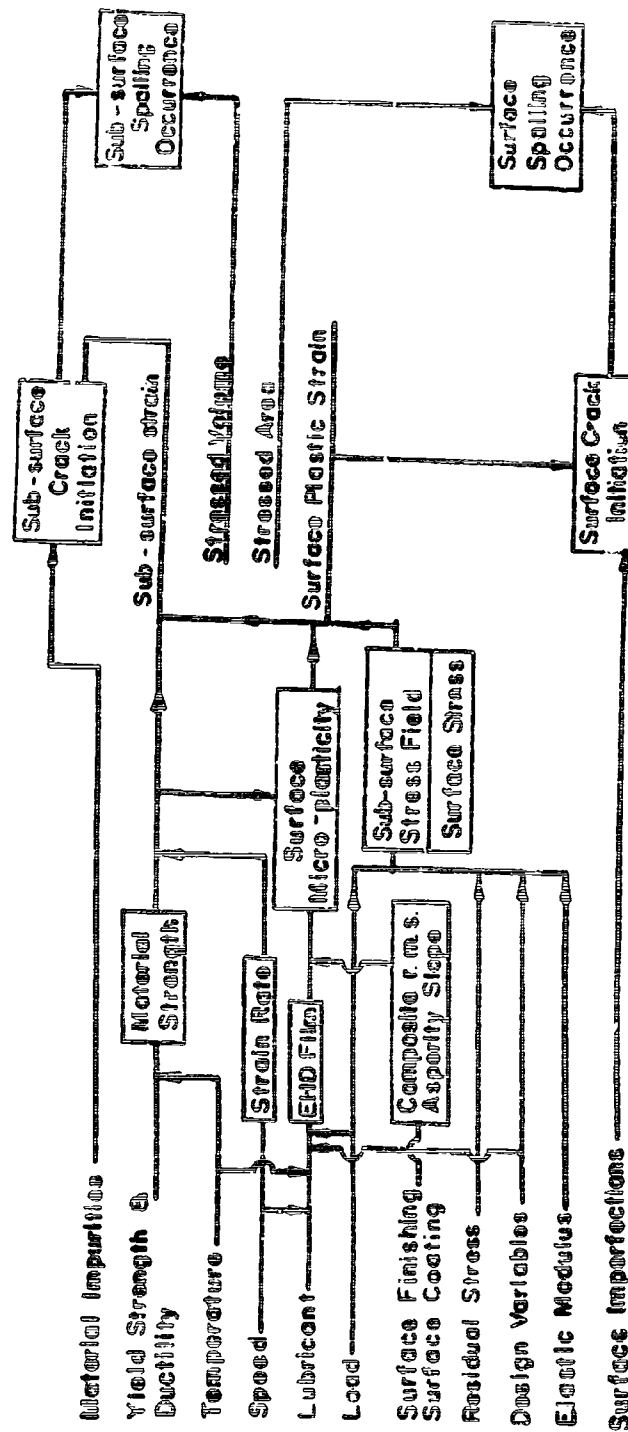


Figure 1. The Interdependence of Variables Affecting Contact Fatigue Life

fewer and smaller subsurface non-metallic inclusions being available to initiate fatigue cracks in clean steel; thus the role of surface defects becomes more important.

4. FAILURE PROCESS DIAGRAM (ROLLING CONTACT)

Taking into account the previously described failure variables and failure mechanisms, it is possible to draw a diagram of the rolling contact failure process, Figure 1 shows this diagram which depicts the interdependence of the variables and their effects on subsurface and surface initiated spalling occurrences. This chart is a brief, self-explanatory, illustration of the previously described failure process.

The interacting effects of the variables are shown in captions enclosed in the rectangular blocks of Figure 1. The arrows mounted on the lines signify an effect.

The "terminations" of this flow chart, shown on the right hand side of Figure 1, represent 1) subsurface spalling occurrences controlled by subsurface crack initiation within a stressed volume and 2) subsurface spalling occurrences controlled by the surface crack initiation within a stressed area.

The stressed area and volume are determined from the quasi-elastic subsurface and surface stress field. Subsurface (or surface) crack initiation is affected by subsurface (or surface) weak points and the magnitude of subsurface (or surface) plastic strain. This, in turn, depends on the quasi-elastic stress field and the material yield strength which have the effect of controlling the amount of plastic deformation for a given stress field.

Ductility is believed to have the ability to suppress crack initiation and propagation in a material matrix under a given cyclic macrostrain field. It is, of course, a matrix parameter.

In Figure 1, the variables listed on the left side consist of two kinds, i.e., 1) material impurities and surface imperfections acting as weak points and 2) the variables affecting the macrostress field in the material.

The following describes the external variables which affect the macrostress field in bearing material:

a. The effect of design variables (dimensions), elastic modulus, and load on the elastic macrostress field has been well covered in the Lundberg-Palmgren theory (1) based on the application of Hertzian theory (32).

b. Residual stress serves as a stress modifier.

c. The lubricant properties, speed and temperature at the rolling contacts are parameters which have been found in recently developed elastohydrodynamic theory to affect EHD film thickness. Rolling contact experiments have shown (46) that the EHD film thickness and surface finish (characterized by surface roughness RMS and asperity slope angle) have significant influence on the occurrence of surface distress or near-surface micro-plasticity (6,44).

d. The effect of speed i.e. strain rate (or the time a material element is under stress) and temperature in the highly stressed volume (or surface) on material strength is hypothesized on the basis of non-rolling contact experience.

SECTION V

FORMULAS FOR FATIGUE CRACK GROWTH

Both sub-surface and surface failure mechanisms require a crack initiation phase followed by a crack propagation phase to reach the point of a functional failure (spalling) of a rolling element (6).

The crack initiation in material under cycling stressing is, in general (except in sharply notched specimens), a process requiring many cycles of stressing. The necessary conditions for crack initiation are generated by plastic deformation around material inhomogeneities or stress raisers as a result of cyclic stressing. The crack initiation phase can be defined as that involving the formation of cracks on the microscopic scale. It is usual to consider that this phase is ended when a self-propagative crack emerges at the site of crack initiation, i.e., one which propagates without further influence of the original defect.

An analysis by Manson and Hirschberg (15) postulates that the time to the initiation of a fatigue crack depends on the magnitude of the cyclic plastic strain at the strain raiser. The strain raiser considered in (15) is a macroscopic one, i.e., a notch in a specimen. As usual, the fatigue process is divided into two stages, i.e., crack initiation and crack propagation. A certain crack size, called "engineering size", is used to demarcate the two stages. It was found that when an "engineering size" crack emerges, at the notch the controlling strain for further crack growth no longer depends on the strain concentration factor at the notch but on the nominal strain in the specimen. Because the strain range is relatively high at the notch, the role of micro-defects can be neglected. The concept in this analysis is useful in understanding crack initiation in rolling contact, although it is realized that in rolling contact material the strain raisers are actually all microscopic.

In rolling contact the micro-defects can be treated, for convenience, as strain raisers. Little is known about the cyclic strain concentration near a microstress raiser, that is, for a given real inclusion, there is no adequate means to know either by observation or by analysis the strain raising factor. Metallurgical experience (3) has shown that the size of an inclusion has a large effect on crack initiation life; other variables such as shape and content may also have an effect on crack initiation. The size and shape of the defects present in

in any volume of metal are stochastic quantities. This calls for a statistical treatment in which the cyclic strain raising factor or severity (designated as d) of a defect, as a function of its size, shape and content, is introduced as a random variable.

In formulating mathematical models of rolling contact fatigue for the surface and sub-surface failure mechanisms, it is hypothesized that the fatigue cracks in both mechanisms initiate from micro-defects or "weak points". In the sub-surface failure mechanisms, the micro-defects are embedded in the matrix material whereas in surface failures pre-existing surface micro-defects and surface fatigue cracks are considered as the weak points of a surface element. This "pre-existing micro-defect hypothesis" is postulated, of course, only for hard steels as used in bearings.

Current concepts of metal fatigue assert (22) that the controlling parameters of fatigue life are the applied cyclic strain and material ductility as they influence crack growth. The rate of growth of a crack is denoted by dA/dN where A is the instantaneous crack area and N is the number of stress cycles. It is hypothesized that the growth rate dA/dN of a crack is dependent on the plastic microstrain ϵ_0 and ductility, D , prevailing in the vicinity of a "defect" consisting of a "pre-existing original" defect and of the crack initiated therefrom.

Thus,

$$\frac{dA}{dN} = A (\epsilon_0, D) \quad (5.1)$$

The ductility D has been defined in static test as (21):

$$D = \log(A_0/A_f) = \log \frac{1}{1-R.A.} \quad (5.2)$$

where A_0 and A_f are the initial and final area of the fracture cross section in the tensile test, and R.A. is the conventional reduction in area.

However, there exist other ductility related quantities which can be defined from fatigue tests. It was shown by Coffin (9) that a straight line results when plotting the logarithm of cycles to failure against the logarithm of cyclic reversing plastic strain in specimen fatigue tests, i.e.,

$$\Delta \epsilon_p = M (2 N_f)^2 \quad (5.3)$$

where $\Delta\epsilon_p$: the reversed cyclic plastic strain
 N_f : number of cycles to failure
 M : a material constant called the fatigue ductility coefficient (9)
 z : a constant

Thus, M and z are ductility related constants which can be determined from specimen tests of material. The above introduced quantity D can be considered as a function of M and z .

It is known (23) that increasing hydrostatic compression increases material ductility. In the above formula, (Equation 5.1), the beneficial effect of compressive hydrostatic stress can be taken into account by the use of the ductility D existing under the given stress conditions.

The hypothesis expressed in Equation (5.1) takes account of the facts that (1) microplastic deformation is a criterion for fatigue crack generation; and for a given ductility, the crack generation rate increases with the magnitude of cyclic plastic strain amplitude at the stress raiser and (2) the generation or the growth of a crack is controlled by the material ductility, i.e., highly ductile material cracks less rapidly.

It is assumed that there exists a macroscopic strain ϵ_0 at the location of the defect if the defect were absent so that a quantity called "defect severity" can be defined as the strain raising factor operating on ϵ_0 to give the strain at the (sub-surface or surface) defect. Θ , the severity of a combined defect and crack is hypothesized as a function of the original defect severity, d , the instantaneous crack area A and the size of the stressed volume (for sub-surface fatigue) or area (for surface fatigue) S , i.e.:

$$\Theta = \Theta(d, A(N), S) \quad (5.4)$$

The severity of localized plastic strain concentration around the defect (as demonstrated e.g. by "Butterfly" structure) is dependent on: 1) the strain raising properties of the defect, Θ 2) the macrostrain ϵ_0 assuming the strain raiser is absent and 3) the yield strength σ_y of the matrix material, i. e.

$$\epsilon_c = \epsilon_c(\epsilon_0, \sigma_y, \Theta) \quad (5.5)$$

* ϵ_0 is defined as total (elastic plus plastic) macrostrain.

Introducing Equations (5.4) and (5.5) into Equation (5.1) one has:

$$\frac{dA}{dN} = \Lambda (\Theta, \epsilon_0, \sigma_y, D) \quad (5.6)$$

It is convenient to separate the variables influential in crack growth into two groups: variables related to defects and matrix variables. In Equation (5.6), Θ is the variable related to defects whereas ϵ_0, σ_y and D are related to the matrix. For simplicity, the matrix effects are consolidated into a single function γ , i.e.:

$$\begin{aligned} \frac{dA}{dN} &= \Lambda (\Theta, \gamma), \\ \gamma &= (\epsilon_0, \sigma_y, D) \end{aligned} \quad (5.7)$$

Assume that the crack growth rate dA/dN can be expressed as a product of the severity function and the matrix function, i.e.:

$$\frac{dA}{dN} = \Theta \cdot \gamma \quad (5.8)$$

Using this designation, the defect function plays the role of a "proportionality constant" between the crack rate and the matrix function of the macrostrain ϵ_0 , the ductility D and the micro-yield strength σ_y . This proportionality constant is, of course, an inherent quality of a given defect.

In the crack initiation phase, the effect of stressed area S on Θ can be neglected. The function Θ_I can be written as a product of a function of d designated by $\Gamma(d)$ a function of A , designated by $f_I(A)$. Thus one has:

$$\Theta_I = f_I(A) \cdot \Gamma(d) \quad (5.9)$$

and

$$\frac{dA}{dN} = f_I(A) \cdot \Gamma(d) \cdot \gamma_I(\sigma_y, D, \epsilon_0)$$

In the crack initiation period, the matrix material parameters σ_y, D and macro-strain ϵ_0 can be considered to be constants with cycling. Letting γ_I be independent of N , Equation (5.9) can be further written as follows after integration with respect to N :

$$f_I(A) = N_I \cdot \gamma_I \cdot \Gamma(d) \quad (5.10)$$

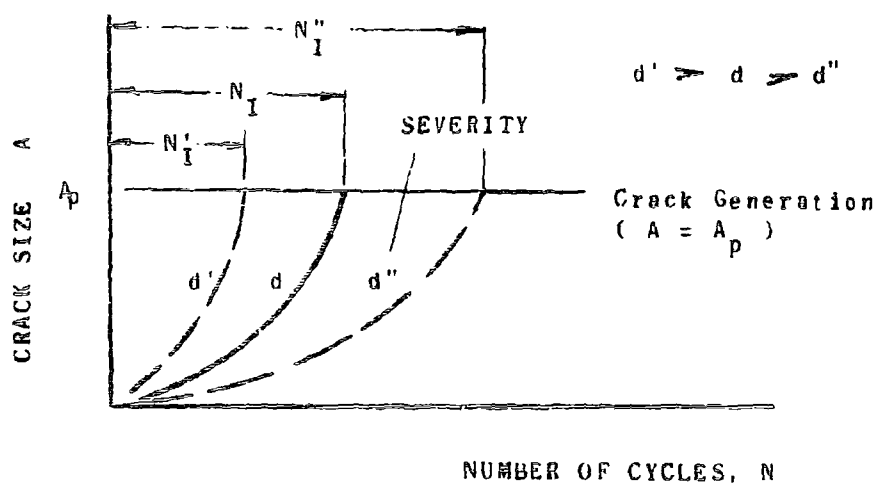


Figure 2. Schematic Representation of Growth of Micro-Cracks

A quantity N_I representing the number of stress cycles required to reach a crack size A_p may be defined from:

$$f_I(A_p) = N_I \gamma_I \Gamma(d) \quad (5.11)$$

where f_I represents a function.

Since A_p , the self propagating crack size, is a constant, Equation (5.10) implies that for a given defect with severity "d" existing in the highly stressed volume, there is a corresponding life N_I to crack generation in numbers of cycles, associated with this defect. Thus, Equation (5.6) can be solved for N_I :

$$N_I = \frac{f_I(A_p)}{\gamma_I \Gamma(d)} \quad (5.12)$$

The function $\Gamma(d)$ can be assumed to be an increasing function of d, thus increasing d corresponds to decreasing N_I . Equation (5.12) implies that every defect has a life associated with its defect severity for a given matrix factor γ_I . Figure 2 shows a schematic representation of growth of micro-cracks from defects of various severities and the dependence of cycles to self-propagating crack generation, on original defect severity d as defined in Equation (5.12), provided that the matrix parameter is the same for each case. The graph shows that increasing d corresponds to increasing slopes of the curves plotting A against N (i.e., the crack rate) and shorter life to crack generation.

The degree of concentration of the plastic strain around an original defect diminishes the distance from the original defect. One can define a macroscopic size of the crack beyond which the crack propagation rate is dependent only on the plastic strain at the crack tip and no longer on the plastic strain around the original defect. (An analogy of this case is given in Manson's fatigue life analysis of notched specimens in (15)). Thus, for large A, it can be assumed that the severity of a "defect after N cycles" is dependent on the instantaneous crack area A and the applied strain ϵ_0 but is independent of the original defect severity d. The propagation of a crack of this size constitutes Phase II of the fatigue process. For a large crack area, an increasing ratio A/S corresponds to an increase of stress concentration at the crack tip. For such cracks dA/dN increases with A/S. Taking into account this size effect, we assume that the severity

function Θ_I in Phase II is a product of a function of A and a function of $\frac{A}{S}$ as follows:

$$\frac{dA}{dN} = \Theta_{II}(A, S) \cdot \gamma_{II}(q, B, \sigma_y) \quad (5.13)$$

$$\Theta_{II} = f_{II}(A/S) \cdot f_3(A) \quad (5.14)$$

where the use of the function γ_{II} denotes the fact that this function may be different for the propagation and initiation phases.

Integrating Equation (5.13) gives:

$$\int_{A_p}^{A_c} \frac{dA}{f_3(A) \cdot f_2(A/S)} = \gamma_{II} \int_{N_I}^N dN \quad (5.15)$$

where N_{II} now represents the number of cycles measured from the starting time when $A = A_p$. Equation (5.15) becomes upon integration:

$$f_{II}(A, A/S) = \gamma_{II} N \quad (5.16)$$

The crack area reaches critical size A_c after N_{II} cycles in Phase II where, from Equation (5.16):

$$N_{II} = f_{II}(A, A_c/S) / \gamma_{II} \quad (5.17)$$

The total number of cycles N_L required to produce a crack of size A_c is:

$$N_L = N_I + N_{II} \quad (5.18)$$

Neglecting any additional cycling requires in Phase III for a crack to grow from size A_c to produce failure, N_L represents the life in cycles associated with a defect of severity d . It is not yet known whether the amount of life thus neglected may be appreciable.

Using Equations (5.12), (5.17), Equation (5.18) becomes

$$N_L = \frac{f_I(A_p)}{\gamma_I \cdot \Gamma(d)} + \frac{f_{II}(A_c, A_c/A_s)}{\gamma_I} \quad (5.19)$$

Equation (5.19) is purely deterministic in nature, that is, if the functions are all known and the ductility and stressed area are prescribed, then Equation (5.19) will yield the number of cycles that elapse until a crack emanating from a defect of designated severity d situated in the stress field so as to be subjected to a strain ϵ_0 , achieves the area A_c .

It has been shown in non-rolling fatigue testing (22) that the crack initiation phase has higher dependence on applied stress (or strain) than the crack propagation phase, which implies that the increase of stress level will shorten the crack initiation phase more rapidly than the crack propagation phase ((24), (25)). Thus the Phase I/Phase II ratio decreases with increasing stress level.

The formation of a spall from a critical sized crack is considered a rapid process associated with cleavage and dimple rupture (27). It is recognized (26) that final fracture is a complex subject in itself, especially, when large plastic strain is involved. It is assumed that the critical crack size A_c is related to the size of the stressed area S in a rolling element cross section:

$$A_c = k_1 S, \quad k_1 = \text{const.} \quad (5.20)$$

It is assumed that after reaching a critical crack size the fracture enters a new phase, i.e., Phase III, which involves a relatively rapid cracking (spalling) process, resulting in a visible spall on rolling surface.

SECTION VI

STATISTICAL THEORY OF ROLLING ELEMENT FAILURE

1. GENERAL

Figure 29 illustrate, more fully the meaning that attaches to some of the terms introduced previously and required in the subsequent development.

View A in Figure 29 is an arbitrary cross section through a rolling body showing the area S of the highly stressed volume defined to be the volume within which the stress critical for fatigue exceeds a designated magnitude.

The highly stressed zone is divided into a large number of cubical cells of small volume. The volume is so selected that within it can be contained an "engineering size" crack as it exists at the end of failure Phase I. Each of these cells is considered to contain a defect from which a crack may initiate.

Defect severity, as defined in Section V, is a function of a defect's size, shape and composition.

In effect, under this model, the ring is assumed to be built solely of cells containing defects, each having different severity. Naturally for clean steels a large proportion of the cells will be occupied by defects having a severity close to unity, (i.e. with no effect on the strain at the cell).

View B of Figure 20 is a side view of the rolling body stretched out into a rectangle. The length denoted by l_s is the circumference at the neutral axis of the rolling body cross section. View B also shows how the stressed area S varies with the coordinate y for a general loading condition.

Equation (5.12) of Section V states that each cell in the highly stressed zone of the rolling body has associated with it a number of stress cycles required for a crack originating in that cell to become of critical size. Equation (5.17) defines an additional life increment N_{II} for a cell's crack to grow to critical size.

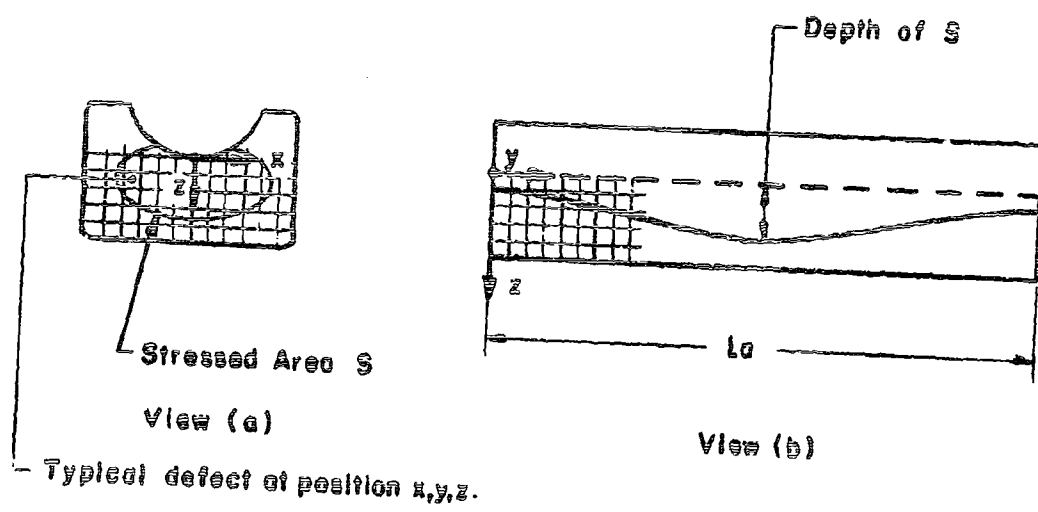


Figure 29. Coordinate System and Defect Cells

The lives N_i in any set of cells are statistically independent.

The Phase II cell lives N_{ij} are not necessarily independent of each other. It is highly likely that the N_{ij} life of one cell will be altered if a crack in a nearby cell completes its Phase I life and the crack extends beyond the initial cell boundaries.

Without the independence assumption the problem of finding the probability distribution of the rolling body life is quite complex.

The treatment which follows will focus on Phase I life under the assumption that Phase II life is either i) negligible compared to Phase I life or ii) constant for every cell.

The Phase I life predicted by equation (5.12) is a deterministic function of the defect severity and the operating variables.

Stochastic considerations enter when one considers that the defects in a steel matrix vary in their severity.

Every defect in a rolling body has a potential Phase I life. The rolling body life is the smallest of these. The defects in any one rolling body constitute a sample from a population of defects with varying severities. The defect producing the smallest Phase I life will thus vary randomly between rolling bodies so that the life of a randomly selected rolling body is a random variable.

We seek the probability distribution of rolling body Phase I life over the population of rolling bodies which are subject to identical material, geometrical and environmental conditions. At this stage in theory, we ignore the (relatively small) variability between and within components, of the matrix properties such as yield strength, and ascribe the variation in life between rolling bodies to the variability in defect severity.

2. DEFECT SEVERITY DISTRIBUTION

The severity of the defect, situated at coordinates x , y and z , is assumed to be a continuous random variable independent of the defect location coordinates. That is, a defect having given severity is not predisposed to be located near any particular set of x , y and z coordinates (group of volume elements in the rolling body).

Let $F(d)$ denote the cumulative distribution function for defect severity. Then, the probability that the severity d_1 of the defect located at coordinates x , y and z is less than a value d , is given by the continuous function $F(d)$.

$$\begin{aligned} \text{Prob} [d_1 \leq d | (x, y, z)] &= F(d) & ; d \geq 1 \\ &= 0 & ; d < 1 \end{aligned} \quad (6.1)$$

Where d_1 represents a realized value of severity and where the left hand side in the above statement is read: "the probability of $d_1 \leq d$ given a set of values x , y , and z is".

The fact that the given coordinates x , y and z do not appear as parameters in the expression for $F(d)$ states what was said above about the independence of the severity distribution from the defect coordinates.

Surface, as opposed to subsurface, defects pose no special problem at this point. They simply represent the special case $z=0$, where the cells on the surface degenerate into square platelets of negligible depth. No information presently exists concerning the functional form for $F(d)$. Two forms will be advanced in Appendix VI which satisfy some plausible general conditions that any defect severity distribution must possess. Other distributions just as plausible may be found however, and experimental information is ultimately required to choose between candidate severity distributions. Nonetheless, valuable insight can be gained regarding the statistics of possible failure, without precise knowledge of this distribution. The discussion will therefore proceed in terms of a general, unspecified, severity distribution $F(d)$.

3. DISTRIBUTION OF "DEFECT LIFE"

From Equation (5.12) one may find the relationship between a defect's severity d and its phase I life N_I as follows:

$$d = \Gamma^{-1} \left(\frac{B}{N_I} \right) \quad (6.2)$$

where

Γ^{-1} denotes the inverse of the function $\Gamma(d)$ defined in Section V,
 $B = \frac{f_I(A_D)}{V_I}$ is a non-random quantity depending on A_D , ϵ_0 , σ_y and η .

Equation (6.2) may be used to transform the distribution of defect severity into a distribution of life associated with the defect population.

Assuming that the function Γ^{-1} is a single valued and monotonically increasing function of d , one may write:

$$G(N_I | x, y, z) = \text{Prob} \left[N_I < N_I \right] = \text{Prob} \left[d_1 > \Gamma^{-1} \left(\frac{B}{N_I} \right) \right] \quad (6.3)$$

$$= 1 - \text{Prob} \left[d_1 < \Gamma^{-1} \left(\frac{B}{N_I} \right) \right]$$

From equation (6.1)

$$G(N_I | x, y, z) = 1 - F \left[\Gamma^{-1} \left(\frac{B}{N_I} \right) \right] \quad (6.4)$$

Equation (6.4) is the distribution of life at coordinate position x , y and z associated with the population of defects having the severity distribution of Equation (6.1).

Of course in any given rolling body, the volume element at x , y , z will have a specific defect severity and hence a determined life calculable from Equation (5.12). It is when one considers the varying values of the severity of the defect that can occupy the cell located at coordinates x , y and z in different rolling bodies from the population of such bodies, that a distribution of cell life results.

Every cell has, associated with it, a distribution function of the form of Equation (6.4), wherein the parameters (since they depend on D , e_0 , and S) will depend upon the location of the cell.

The life of a complete rolling element is identical to the life of that cell which, of all the cells, has the shortest Phase I life according to Equation (5.12).

From probability theory, the probability that at least one cell produces failure before N_I cycles is the complement of the probability that all cells survive beyond N_I cycles.

Assuming independence of the cell lives, (as one may, for Phase I life considered here, based on the choice of cell size to include completely a crack of "engineering size"), the probability that all cells survive is the product of the probability that each survives. Thus $H(N_I)$ is given by:

$$H(N_I) = 1 - \prod_i \left[1 - G_i(N_I) \right] \quad (6.5)$$

where G_i denotes the distribution of Equation (6.4) at the i -th cell in the structure. The symbol w_i denotes that the terms evaluated at each cell in the structure including the surface cells are to be multiplied together.

Equation (6.5) is of great generality and could be evaluated numerically for any general shape and loading condition if the functions in Equation (6.5) and the severity distribution of Equation (6.1) were known.

These functions are not presently known however, and may not be practically determinable in the near future. One has as a recourse, the option of making plausible assumptions of the unknown function forms.

In view of the uncertainty with which any function can be assumed, it is preferable in this regard to limit the number and restrictiveness of such assumptions. In particular it is preferred to assume, if possible, only the asymptotic behavior of a given function rather than its specific form since this merely limits the possible functions to an admissible class.

At the present state of theory, we will restrict the discussion to applications wherein the distributions $G_i(N_i)$ may be considered identical for all values of i .

This corresponds to the Lundberg-Palmgren assumption on failure locations for the case of thrust loaded bearings. The Lundberg-Palmgren assumption is, broadly, that all points inside a highly stressed zone are equally likely to fail and no point outside this zone will fail. For a thrust loaded bearing, the highly stressed area is the same in all cross sections, and thus the above stated case applies.

4. ASYMPTOTIC DISTRIBUTION OF SMALLEST DEFECT LIFE IN A ROLLING BODY

If a random sample of size m is drawn from a distribution having a distribution function $F(x)$ the ordered values $x_1, x_2, \dots, x_m$ (where $x_1 < x_2 < \dots < x_m$) are themselves random variables having a distribution, over repeated samples, of a form which depends upon order number i , sample size m and parent distribution $F(X)$.

In particular for $i = 1$, that is for the smallest member of the sample, the distribution $F_1(X_1)$ is given by, (45):

$$F_1(x_1) = 1 - [1 - F(x_1)]^m \quad (6.6)$$

As the sample size m increases, the distribution $F_1(X_1)$ under rather general conditions converges to a specific form. In the cases where it does converge, it will converge to one of three distributional forms, depending upon the behavior of the parent distribution $F(x)$ in the vicinity of the smallest admissible value for x (See Gumbel, (47) p. 162).

The Weibull distribution is one such limiting form for the distribution of smallest values and it is applicable to parent distributions which have finite admissible values. In order for the distribution of x_1 , to converge to the Weibull distribution, the parent distribution $F(x)$ must behave like a power function in x in the vicinity of the minimum admissible value, that is:

$$F(x) \rightarrow \beta(x-x_0)^k \quad (6.7)$$

Where β and k are positive constants and x_0 is the smallest admissible value. If Equation (6.7) is satisfied, it has been shown (Epstein (2)) that for $x < x_0$, $F_1(x_1)$ has the following form:

$$F_1(x_1) = 1 - \exp \left[-\beta(x_1 - x_0)^k \right] \quad (6.8)$$

Thus the distribution of rolling body lives (shortest defect lives) will be Weibull distributed if the parent distribution of cell Phase I life $G(N_I)$ satisfies the condition of Equation (6.7) wherein the minimum Phase I life corresponding to infinite defect severity is ∞ .

$$\lim_{N \rightarrow N_0} G(N_I) = \lim_{N \rightarrow N_0} \left[1 - F \left[\Gamma^{-1} \left(\frac{N}{N_I} \right) \right] \right] = \beta N_I^k \quad (6.9)$$

If Equation (6.9) is satisfied for the actually applicable functions F and Γ , the distribution $H(N_I)$ becomes, by Equation (6.8):

$$H(N_I) = 1 - \exp \left[-\beta(N_I)^k \right] = 1 - \exp \left[-\left(\frac{N_I}{N^*} \right)^k \right] \quad (6.10)$$

where $N^* = 1/(\beta\beta)^{1/k}$

Sample size m is the total number of defects (cells in a rolling body).

Denoting the number of defects per unit volume by η , one has:

$$N = \eta S l_a = \eta V \quad (6.11)$$

Substituting (6.11) in (6.10) gives

$$H(N) = 1 - \exp - \left(\frac{N_I}{(\beta \eta V)^{-1/k}} \right)^k \quad (6.12)$$

Equation (6.12) shows the stressed volume appearing as a scale parameter of the Weibull distribution. This is the volume effect of the Lundberg-Palmgren formula, which is believed to be in good agreement with existing ball bearing fatigue data.

We have now shown that the Lundberg-Palmgren formula is compatible with the fatigue failure model developed above under the restrictions on the two functions Γ and F embodied in Equation (6.9) above.

SECTION VII

MACROSCOPIC STRESS CALCULATION FOR THE HIGHLY STRESSED VOLUMES IN A HERTZIAN CONTACT

In rolling contact, there exists a "reversing shear stress" (1) which has long been regarded as the best criterion of rolling contact fatigue strength; usually on the ground that its range of variation is always equal to or greater than that of the maximum shear stress and that there is some evidence suggesting that total amplitude of shear stress correlates with life (1,29, 43). This assumption was used by Lundberg and Palmgren (1) who stated that most subsurface fatigue cracks are generated at a depth equal to that of the reversing shear (at the central circumferential plane). Later, Greenert (28), in testing toroids of various curvatures, found that the reversing shear range gives good correlation with fatigue life. In available literature, the shear stress has been determined only for the central circumferential plane of a rolling element on which the maximum shear range is equal to that of orthogonal shear stress (τ_{yz}). Since the fatigue life is closely related to the highly stressed volume, it is necessary to know the volume of material subject to a certain level of shear range for a given applied Hertzian load. The determination of this volume requires a knowledge of the stress distribution under a Hertzian load.

A computation of (reversing) shear stress range in a Hertzian elliptical stress field has been performed. Previous workers have studied the maximum alternating shear stress on the center plane of a contact $x = 0$, (see Figure 3). At $x = 0$, the shear stress is maximum in a plane at a known distance below the surface and parallel to the XY plane. As shown below, the alternating shear stress for other planes, ($x \neq 0$) is maximum on planes inclined to the XY-plane, and therefore is not orthogonal.

The computation presented here is a generalization of Lundberg's solution which was confined to a central circumferential plane only. The present solution provides a means to determine the stressed volume at various stress levels, considering the shear range as the critical stress.

The basic mathematical formulation for the subsurface stresses at all locations in a Hertzian stress field is presented in Appendix I. These formulas were developed by Lundberg and Sjövall in 1951 (34).

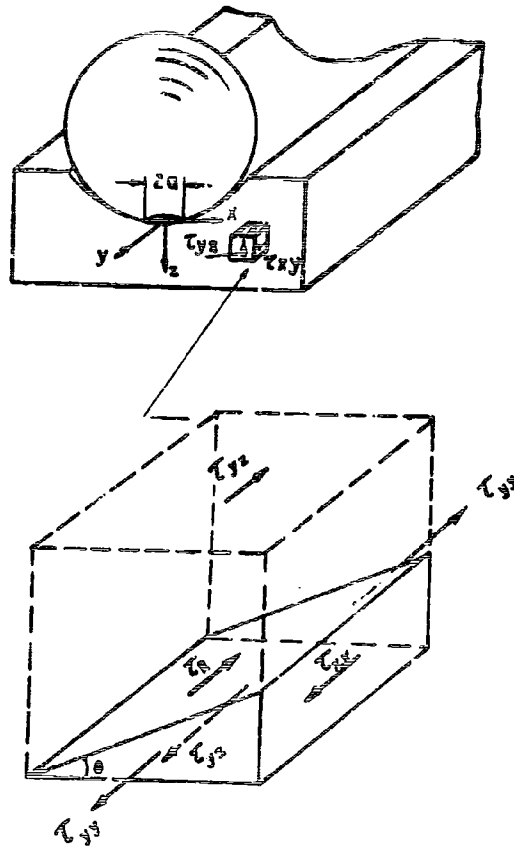


Figure 3. Equilibrium of Stresses Acting in Y-Direction

The current computation involves the determination, for any point in a cross section (perpendicular to rolling direction), of the range of maximum shear stress (τ_R) acting in the direction of rolling at a certain value of y (where y is the distance from the central plane of symmetry in the direction of rolling) but on any plane inclined to the surface (with an angle θ_1). The determination of the values of y and θ_1 , at which τ_θ is maximum, requires an iterative maximization process.

The following stress analysis was conducted:

- a) Based on the equilibrium of stresses acting on a sub-surface element shown in Figure 3, the shear stress acting on an arbitrarily inclined plane can be expressed as the combination of two orthogonal shear stresses τ_{zy} and τ_{xy} i.e.

$$\tau_\theta = \tau_{zy} \cos \theta + \tau_{xy} \sin \theta \quad (7.1)$$

where τ_{zy} and τ_{xy} can be computed from formulae available in (34) for given values of x , y , z and a/b (the ratio of major axis and minor axis of a contact ellipse).

Since both τ_{xy} and τ_{zy} are odd functions of y , τ_θ is also an odd function of y . Both τ_{xy} and τ_{zy} vanish when $y = 0$. Furthermore, since $\tau_{xy} = 0$ for all y when $x = 0$, the amplitude of τ_θ is equal to that of τ_{zy} at the central plane in the direction of rolling (where $x = 0$). This justifies Lundberg's analysis (1) of setting the amplitude of τ_{zy} equal to the maximum shear amplitude for $x = 0$.

- b) For given values of x and z the maximum value of τ_θ can be found by setting:

$$\frac{\partial \tau_\theta}{\partial \theta} = \tau_{xy} \cdot \cos \theta - \tau_{zy} \cdot \sin \theta = 0 \quad (7.2)$$

and

$$\frac{\partial \tau_\theta}{\partial y} = \frac{d}{dy} (\tau_{zy}) \cos \theta + \frac{d}{dy} (\tau_{xy}) \sin \theta = 0 \quad (7.3)$$

These two differential equations can be reduced to:

$$\frac{d}{dy} (\tau_{xy}^2 + \tau_{zy}^2) = 0$$

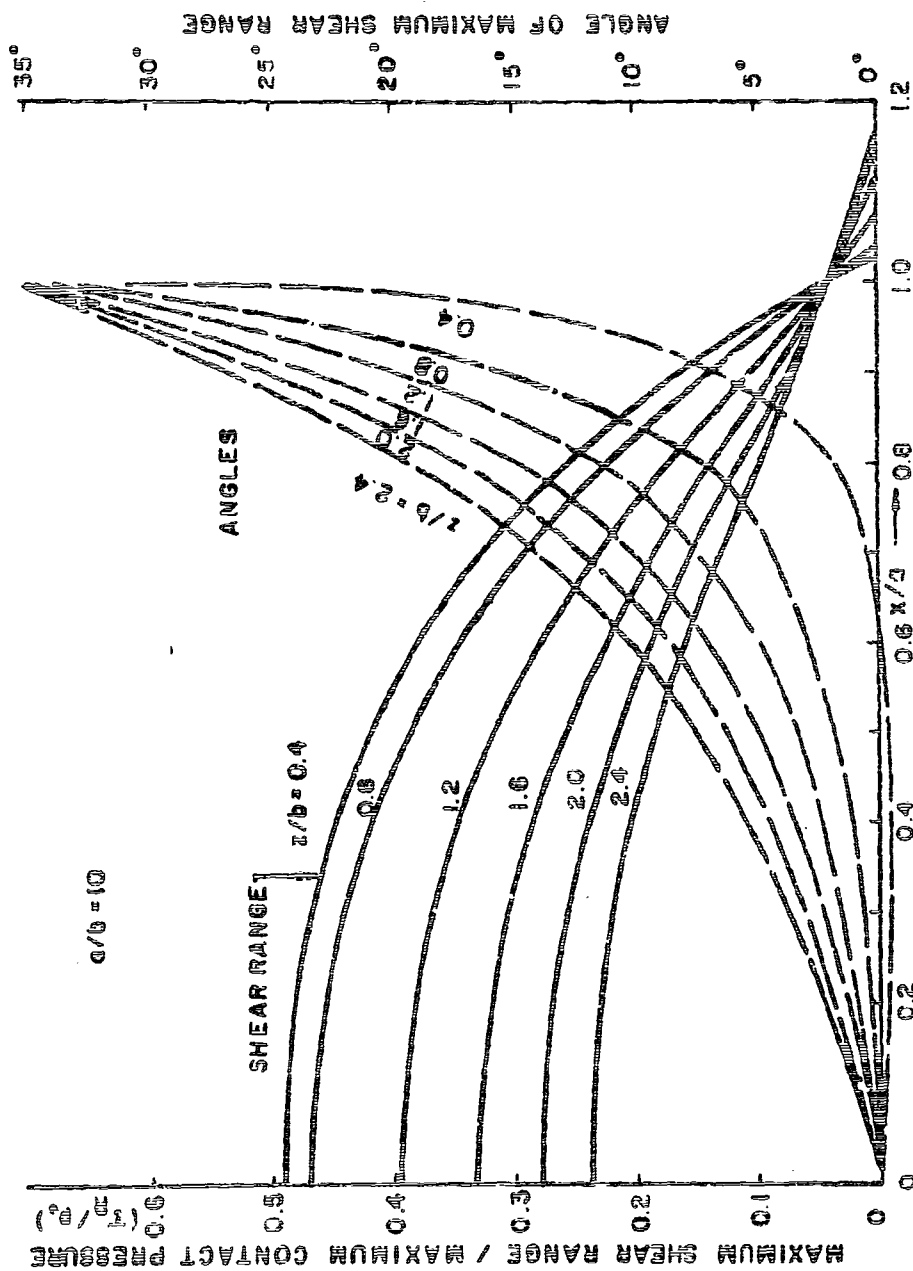


Figure 4. Variation of τ_R and θ_1 with x and z

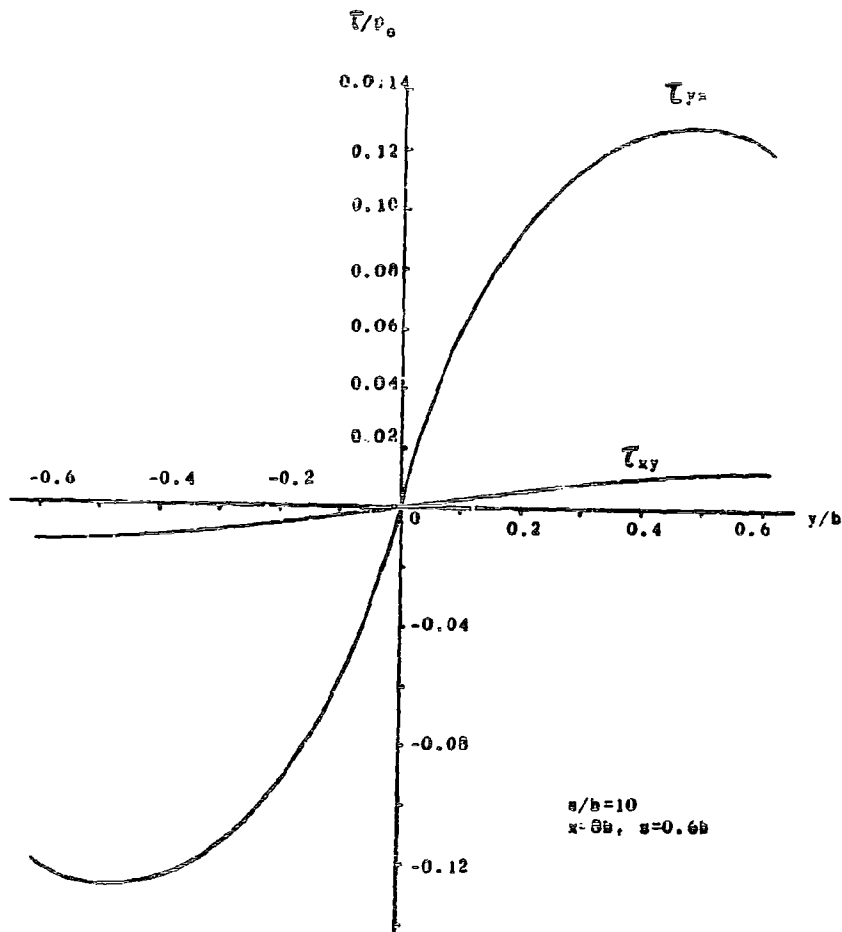


Figure 5. Typical Variation of Orthogonal Shear Stresses with y for given x and z

which implies that when $\tau_{zy}^2 + \tau_{xy}^2$ is maximum, τ_R is also a maximum. Therefore, a value of y_1 , corresponding to maximum $\tau_{zy}^2 + \tau_{xy}^2$ can be found from the stress calculation and using the values of τ_{xy} and τ_{zy} at $y = y_1$, the value $0 = \theta_1$ for which τ_θ is maximum, can be derived from Equation (7.2)

$$\text{i.e.}; \quad \tan \theta_1 = (\tau_{xy}/\tau_{zy})_{y=y_1} \quad (7.4)$$

The maximum shear range τ_R is equal to:

$$2(\tau_\theta)_{\max.} = 2(\tau_\theta)_{y=y_1}$$

Using the values of θ_1 and y_1 thus obtained, the maximum value of τ_θ can be obtained from Equation (7.1) after substitution.

Using a digital computer, numerical computation has been made for values of τ_{xy} , τ_{zy} and τ_θ (in terms of the maximum contact pressure p_0) as functions of y/a and θ .

Figure 4 plots, from these computations, the variation of τ_R as a function of x and z for a contact with $a/b = 10$. The variation of θ_1 as a function of x and z is shown by dotted curves in Figure 4.

At $x = a$, all of the τ_R and θ_1 curves very nearly intersect. This indicates that at $x/a = 1.0$ the maximum stress range is equal to $0.04p_0$ independent of depth z for the z values shown in Figure 4. Further, the plane on which the maximum occurs is defined by $\theta_1 = 35^\circ$ independent of z .

Figure 5 shows the variation of τ_{yz} and τ_{xy} with y at a fixed location in the ring cross section, i.e. at $x = 8b$ and $z = 0.6b$. Figure 4 shows that for a given depth z , the shear stress range τ_R reaches a maximum at $x = 0$ and decreases monotonically with x .

Using the method described above, the contours of equal shear range on a ring cross-section for a given major axis / minor axis ratio (characterizing the contact ellipse) can be drawn, based on the values of shear range computed for a large number of grid points equally spaced in the highly stressed zone of a ring cross section.

Figure 6 shows contours of equal shear range for the eccentricity ratio $a/b=10$ representing a typical ball-race contact of

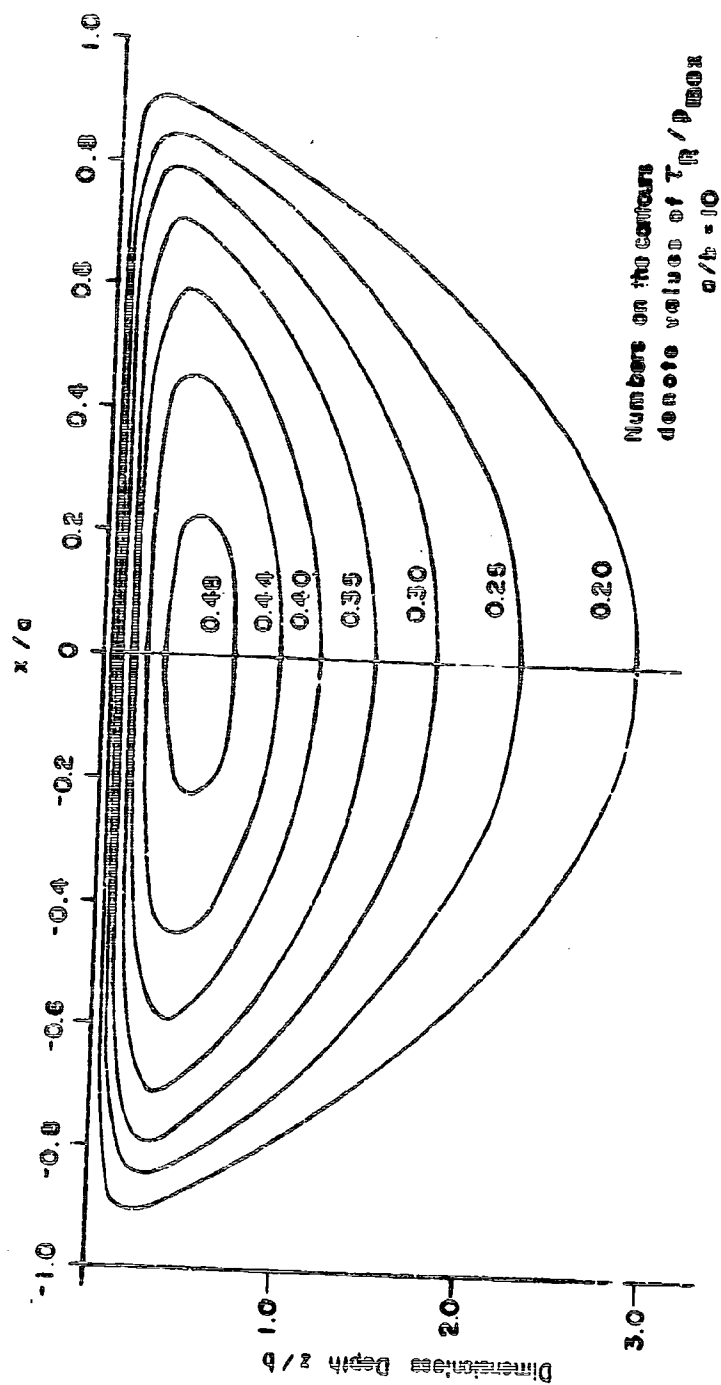


Figure 6. Contours of Equal Shear Range τ_R in a Hertzian Elliptical Contact ($a/b = 10$)

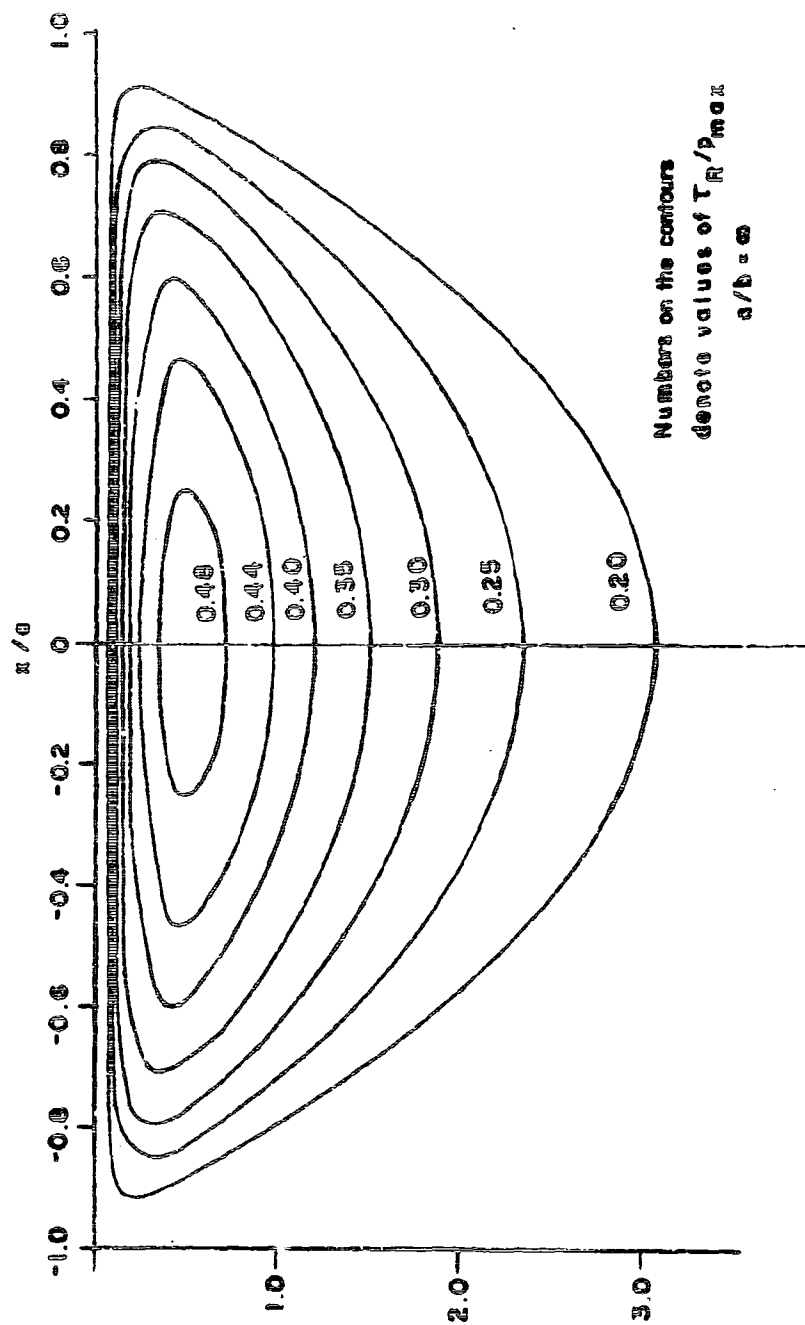


Figure 7. Contours of Equal Shear Range τ_R in a Hertzian Elliptical Contact ($a/b = \infty$)

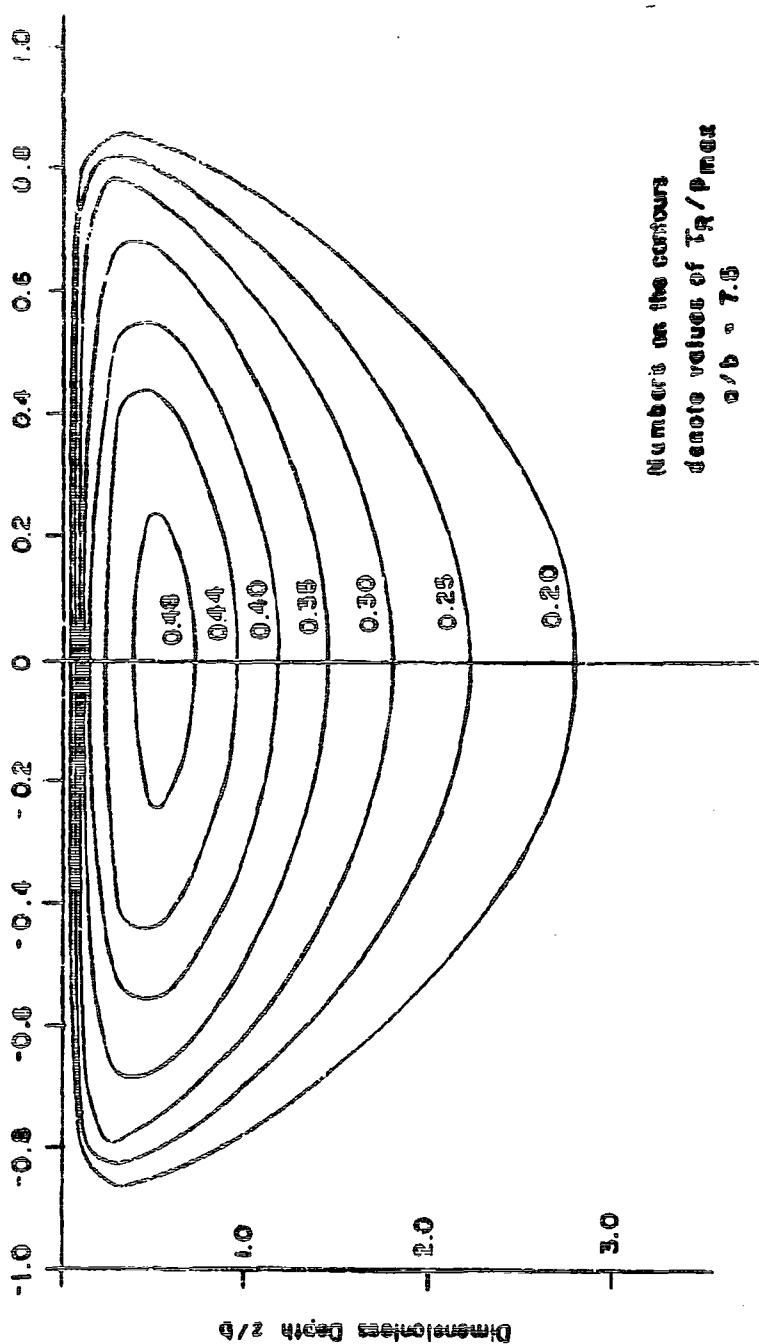


Figure 8. Contours of Equal Shear Range τ_R in a Hertzian Elliptical Contact ($a/b = 7.5$)

a deep groove ball bearing. These contours are drawn based on a 500 grid point system in a quadrant of the $x-z$ plane containing the major axis of the contact ellipse. It is noticed that the shear range along $x = 0$ is very close to that in a contact of two infinitely long cylinders.

A limiting case for $a/b = \infty$ has been deduced from formulas in (34) which corresponds to an infinitely narrow contact ellipse (see Appendix II.) In this case, the stress distribution in the central circumferential plane ($x = 0$) is identical to the plane stress solution for a Hertzian two-dimensional contact. Using the results of the limiting case, it is possible to plot the values of shear range on a plane with coordinates x/a and z/b .

Figure 7 shows the contours of equal shear range for $a/b = \infty$ i.e. an infinitely narrow contact ellipse with its major axis lying on the X -axis, perpendicular to the direction of rolling. By comparison with Figure 6, it is seen that the contours of equal shear range for the two cases are nearly identical except that in the case $a/b = \infty$, the level of τ_R lies slightly deeper under the surface than for $a/b = 10$.

Contours of equal shear range τ_R for the case $a/b = 7.5$ have also been obtained and these are shown in Figure 8.

Figure 9 plots the area S expressed as multiples of ab and enclosed by contours of equal τ_R as a function of $\tau_R/p_{\max}$ for values of $a/b = 7.5, 10$, and ∞ . The points where these curves intersect the abscissa correspond to the maximum values of τ_R throughout the stressed region, i.e. $2\tau_0$. It is noted that the curves lie very close to each other. The three curves coincide even more closely if one plots S/az_0 vs $\tau_R/2\tau_0$, by multiplying the vertical coordinate of Figure 9 by b/z_0 and the horizontal coordinate by $p_{\max}/\tau_0$, where z_0 is the depth of the point below the surface which is subject to the maximum shear stress range $2\tau_0$. The values of b/z_0 and $p_{\max}/\tau_0$ as functions of a/b are obtained from (34) and are tabulated in Table 3. Figure 10 is a plot of S/az_0 vs $\tau_R/2\tau_0$ and shows that the curves having parameters $a/b = 7.5, 10$ and ∞ nearly coincide. This means that the stressed area S at any given τ_R value is proportional to the product of a and z_0 where a is the semi-major axis and z_0 is the depth location coordinate of τ_0 . This approximate relationship appears to be valid for $7.5 < a/b < \infty$; a range encompassing the usual dimensions of a ball-raceway contact ellipse, including line contact. This finding supports Lundberg-Palmgren's analysis (1) in which the stressed area in the ring cross-section is assumed to be proportional to az_0 .

Table III

VARIATION OF z_0/b AND τ_0/p_{max} WITH a/b IN HERTZIAN CONTACT

a/b	5	7.5	10	∞
z_0/b	0.4861	0.4925	0.4963	0.5
τ_0/p_{max}	0.2476	0.2480	0.2494	0.2500

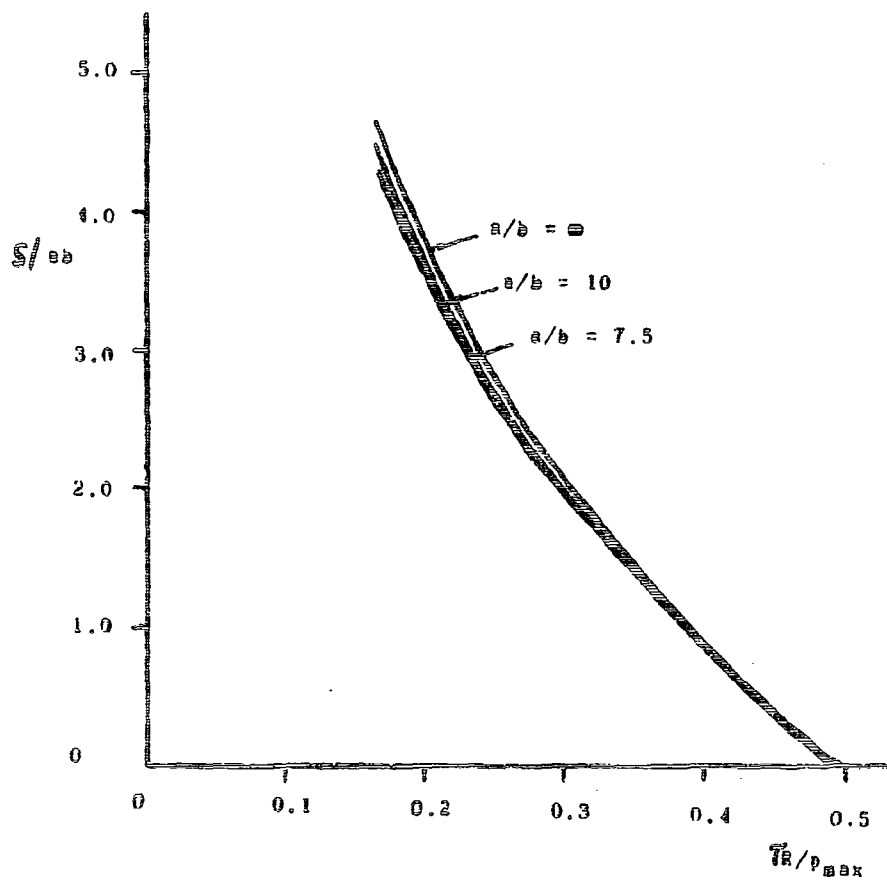


Figure 9. Variation of S/ab with a/b and τ_R/p_{max} .

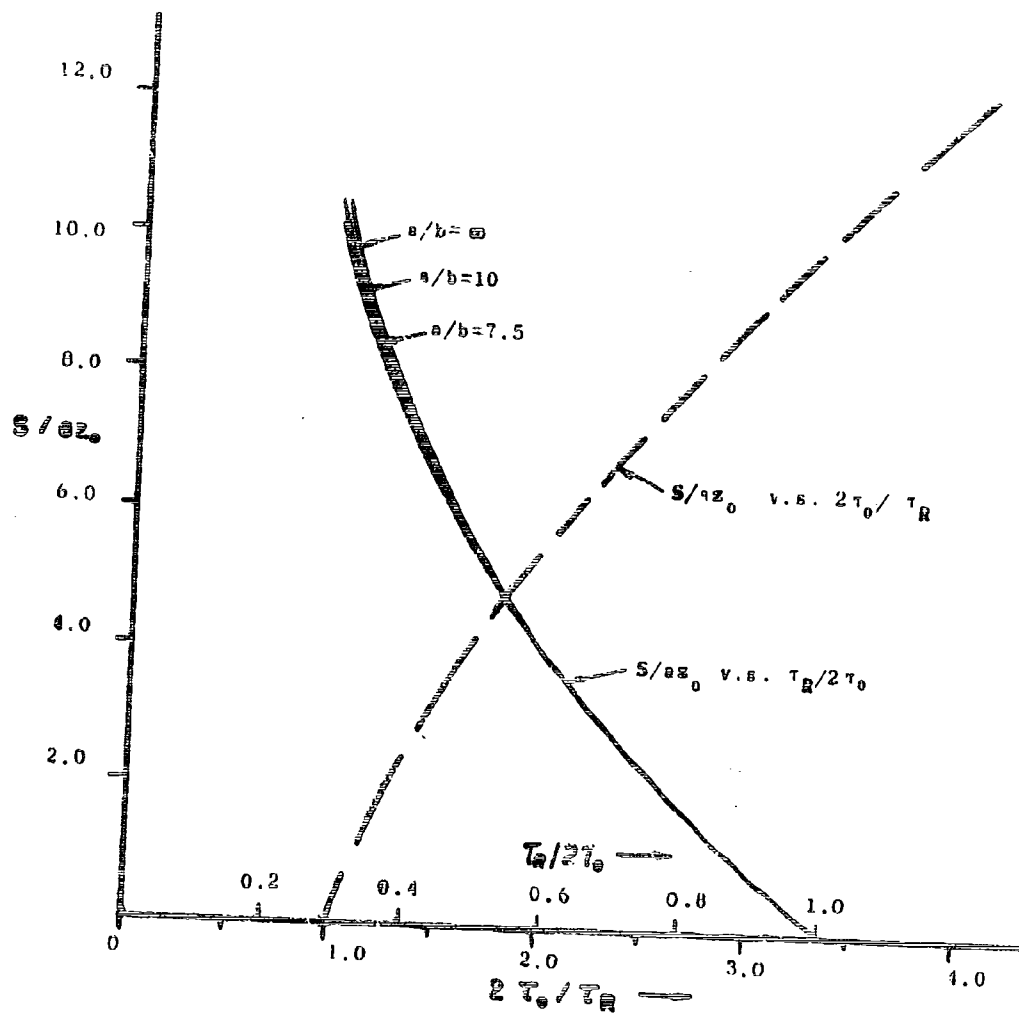


Figure 10. Variation of S/az_0 with a/b , $\tau_R/2\tau_0$ and $2\tau_0/\tau_R$

A selection of the values to be used in defining the stressed area can be made by setting τ_R equal to a threshold stress level for plastic deformation for the given material composition and hardness. If selected in this manner, τ_R is a constant for a given material. For future use, it is more convenient to plot S/az as a function of $2\tau_0/\tau_R$ shown as dotted line in Figure 10.

This curve shows that plastically stressed (macroscopic) area exists only for $2\tau_0/\tau_R > 1$. The size of the plastically stressed area can be approximated by :

$$S \sim \pi \cdot z_0 \cdot \left[\frac{2\tau_0 - \tau_R}{\tau_R} \right]^{0.75} \quad (7.5)$$

This formula shows that the area enclosed by the contour of equal τ_R can be expressed approximately as a function of τ_R . It is also of interest to know the average shear range, $(\tau_R)_{ave}$, in the stressed area, S , enclosed by the contour of equal τ_R . Appendix IV gives the details of the computation using Equation (7.5). The result yields the following approximation:

$$\frac{S}{az_0} = -6.45 \left(\frac{(\tau_R)_{ave}}{\tau_R} \right) + 12.9 \quad (7.6)$$

where $(\tau_R)_{ave}$ = average magnitude of a maximum reversing shear range in a closed contour of equal τ_R

S = area enclosed by the closed contour of equal τ_R

a = major axis of the contact ellipse

τ_0 = maximum reversing shear stress

z_0 = depth where τ_0 exists

SECTION VIII

DETERMINATION OF SHEAR STRESS NEAR ASPERITIES

To demonstrate a possible mechanism for the generation of plastic deformation in asperities, an elastic analysis is presented below for the stress distribution in a contacting asperity using an idealized asperity profile. The result of this analysis may enable one to predict the location and severity of a plastic occurrence and take into account the relevant variables such as lubricant film thickness and surface roughness. It should also be noted that in the following analysis the friction at the contact surface is neglected.

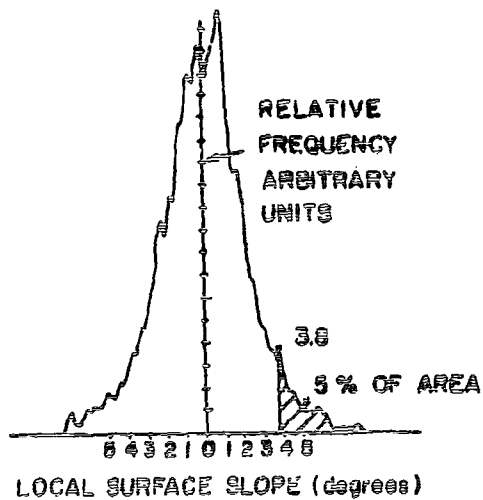
Figure 11 shows the distribution of slopes for three typical, abrasively finished surfaces. From these distributions, it is seen that the typical slope of abrasively finished hard steel surfaces varies widely. For a ground surface, the 95th percentile slope is $\theta_{95} = 29^\circ$; for a much smoother honed surface, the 95th percentile is $\theta_{95} = 3.8^\circ$, while for the still smoother, lapped surface, it is only $\theta_{95} = 0.7^\circ$. The RMS surface roughness values of these three surface finishes were found to be $\sigma = 13, 1.8$ and $0.4 \mu\text{in}$, respectively.

Quantitative insight into the plastic occurrences at asperities, as a function of parameters listed above, can be achieved by introducing a simple mathematical model in which the contact of a single asperity having an idealized profile with an elastic half plane is considered. This model contains no assumption regarding the height distribution of a population of asperities.

Figure 12 shows an idealized two dimensional asperity shape, similar to what one would expect to find on ground, honed, or lapped surfaces, formed by a multitude of cuts by sharp and straight-edged abrasive grains. The asperity is a plane-sided ridge with a curved tip of radius R . The slope angle of the sides is θ which varies with the process of surface finishing and is accessible to experimental determination (6).

The plane contact problem with the above described profile and a straight-edged half plane can be solved based on Muskhelishvili's method of singular integral equations (36), provided that there exist no sharp curves or corners along the entire profile. The derivation of the solution for this contact problem is given in detail in Appendix III.

TRANSVERSE ROUGHNESS OF
HONED RING SURFACE



GROUND FINISH OF
1/4" DIA. BALL

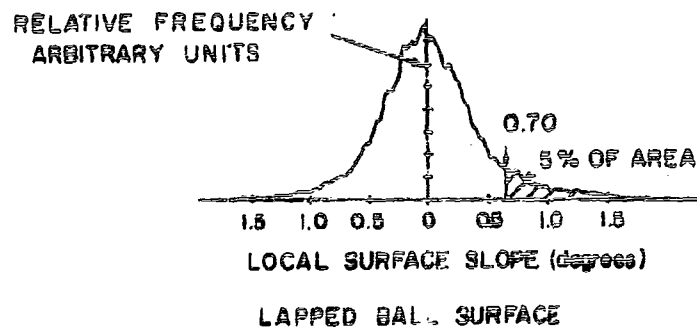
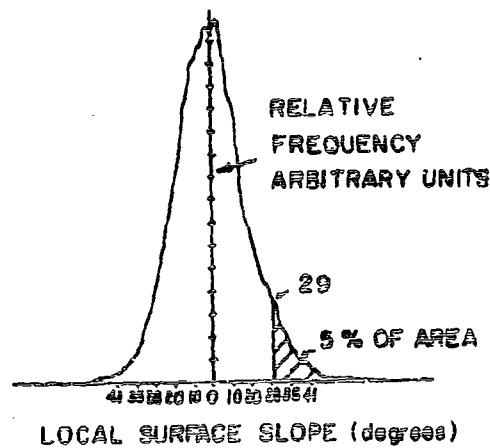


Figure 11. Asperity Slope Distribution on Ground, Honed and Lapped Surfaces

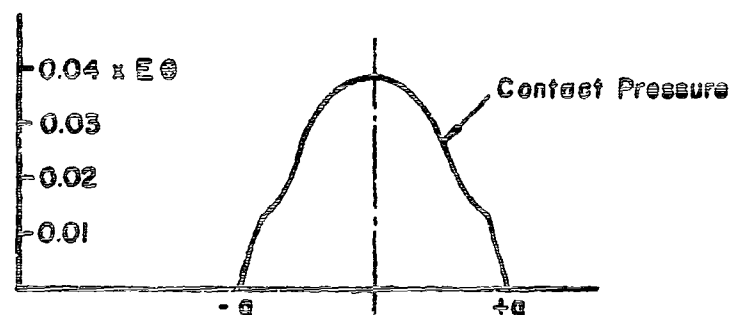
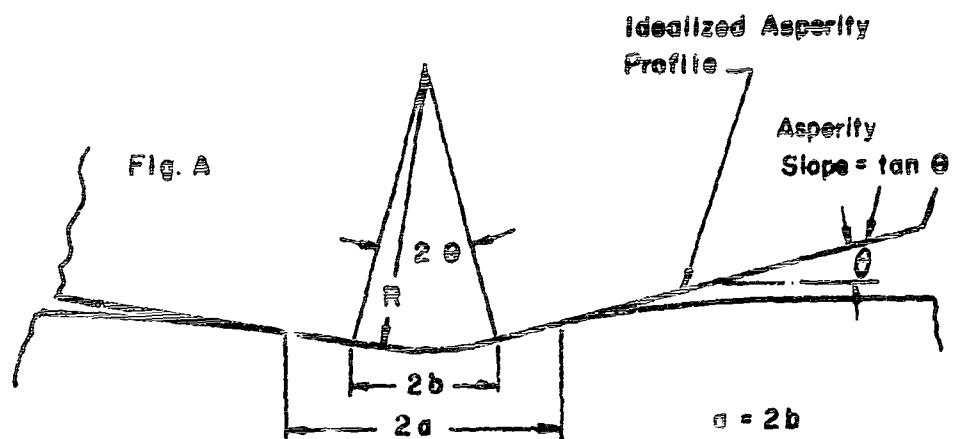


Fig. B

Figure 12. The Contact of an Idealized Surface Asperity

A numerical computation of surface pressure was performed and is shown in Figure 12, based on the case $a = 2b$ wherein the width of the contact region is equal to twice the width of the curved base. Computation for other values of a/b is equally feasible.

The surface pressure distribution plotted in Figure 12 was computed based on a closed form solution derived from (36). However, for the sub-surface stress distribution there is no closed form solution available for computation and a numerical technique is required to obtain the sub-surface stresses.

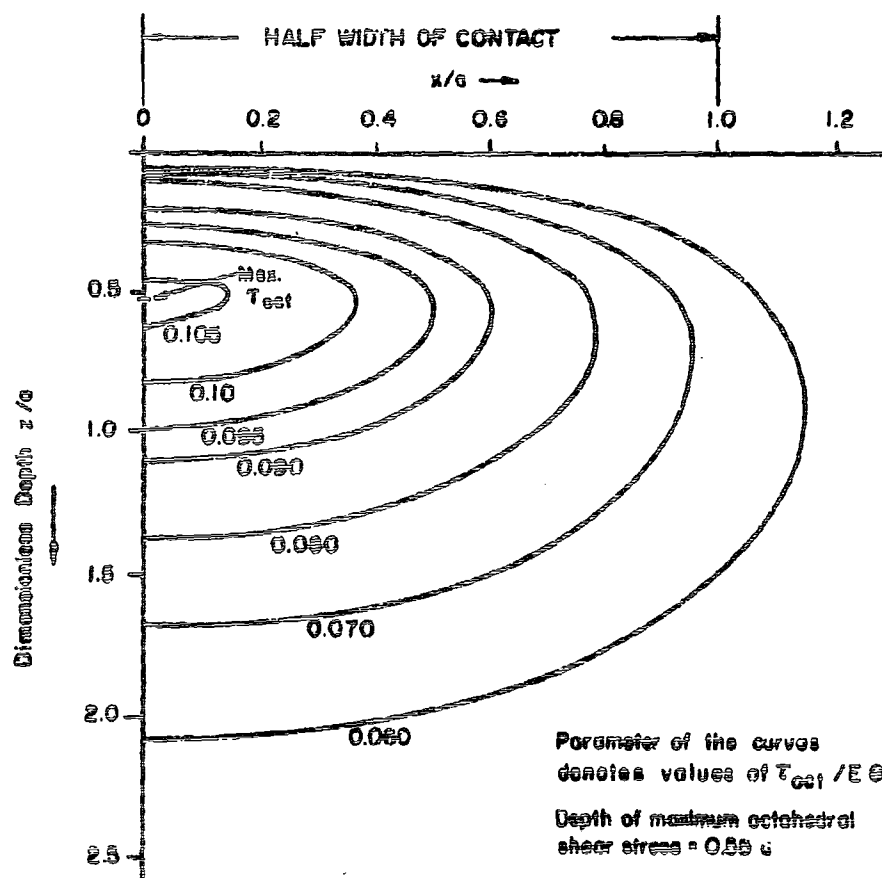


Figure 13 shows the contours of equal octahedral shear stress^{*} in terms of the product of Young's modulus and asperity slope, $E'\theta$ in the cross section at the plane of the contact depicted in Figure 12. It can be seen that the depth where the maximum value of octahedral shear stress occurs is about 0.55 of the half-width of the asperity contact region on the surface. For a typical asperity contact width on a finely finished surface (1 μ in. RMS) (6), this depth is of the order of tens to hundreds of microinches. Thus, high shear stresses are indeed generated close to the surface.

The following relationship has been derived for the magnitude of the shear stress along the axis of a symmetry: (Appendix III)

$$\tau_{45^\circ} = - \frac{E'\theta}{\pi^2} \frac{s}{b} - \frac{E'\theta}{\pi^2} \left[\frac{\pi}{2} - \tan^{-1}(\gamma \tan u) - \gamma(\pi/2 - u) \right] \quad (8.1)$$

where $\xi = z/a$
 $\gamma = \xi / (1 + \xi^2)^{1/2}$
 and $u = \cos^{-1}(b/a)$

Figure 14 shows a plot of the maximum value of τ_{45° computed from Equation (6.1) against b/a . The maximum value of τ_{45° is seen to be proportional to θ and is a function of the ratio a/b . When $b/a \rightarrow 0$, i.e. when the contact area is wide in comparison with the asperity tip width, the following limiting case exists:

$$(\tau_{45^\circ})_{\max} \rightarrow E'\theta / \pi^2 \quad (8.2)$$

$$z_{\max} = 0.904 \sqrt{a b} \quad (8.3)$$

*The octahedral shear stress is defined as $\tau_{\text{oct}} = \sqrt{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2} / 3$, where $\sigma_1, \sigma_2, \sigma_3$ are the three principal stresses, whereas the Von Mises yield stress is $\sigma_D = 3^{1/2} \tau_{\text{oct}} / 2$. Note that τ_{oct} acts on one of the faces of a regular octahedron with vertices on the axes X, Y and Z.

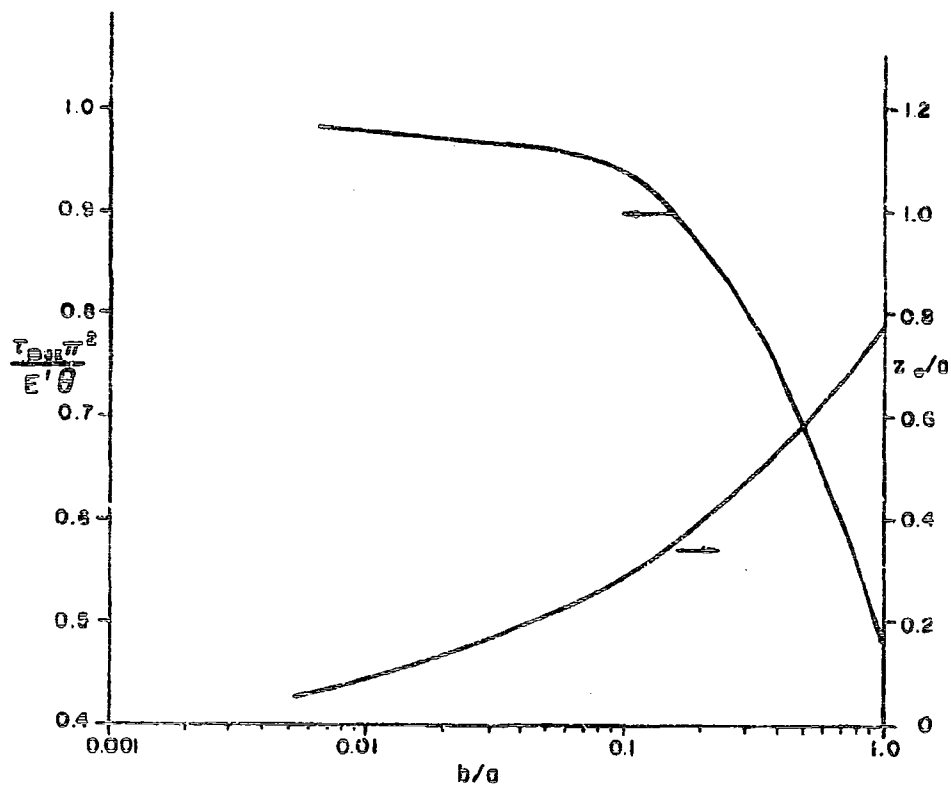


Figure 14. Variation of $(\tau_{45^\circ})_{\text{max.}}$ and its Depth z_0 with b/a for the Simple Asperity Model

and the contact force is:

$$P \rightarrow E' \frac{2}{3} a / \pi$$

These relationships imply the following interesting facts when $b/a \rightarrow 0$, i.e. the elastic deformation of an asperity is large in comparison with its tip width (for example, in the case of a sharply tipped asperity or heavy elastic depression of an asperity). This is the case for low h/σ (the minimum EHD film thickness/composite surface roughness RMS ratio) values:

1. The maximum shear stress, $\tau_{\max}$, approaches a constant value proportional to the asperity slope angle, but independent of the load.
2. The depth, $y_{\max}$, of the point at which $\tau_{\max}$ occurs increases with the 1/2-th power of the semi-width "a" of the asperity contact.
3. The load carried by the individual asperity contact is proportional to the contact width "a" and the angle θ . This single asperity model yields, asymptotically, a proportionality between P and a, in agreement with the Archard's postulate (35), which states that in dry contact the real asperity contact area increases linearly with the load.

Of course, there must be cases in which the degree of indentation or a/b is not large. This occurs, generally, when
a) there is a thick lubricant film separating the two rough surfaces, i.e. h/σ is large or, specifically, when
b) the tip radius of the asperity on real surfaces is large in comparison with the asperity spacing S.

The following presents a simple model for asperity interaction for the purpose of relating near-surface plastic occurrences to the minimum EHD film thickness and surface roughness parameters.

Figure 15 (A) shows the two dimensional contact of a single asperity assumed to be rigid, with an elastic straight edged half plane; the degree of approach of the two bodies is controlled by the EHD minimum film thickness h. For convenience, the vertical distance between the outer line of the asperity profile and the asperity tip is set equal to 3σ where σ is the RMS value of the asperity profile and is of an acceptable order of magnitude approximation for real profiles. For the hypothetical

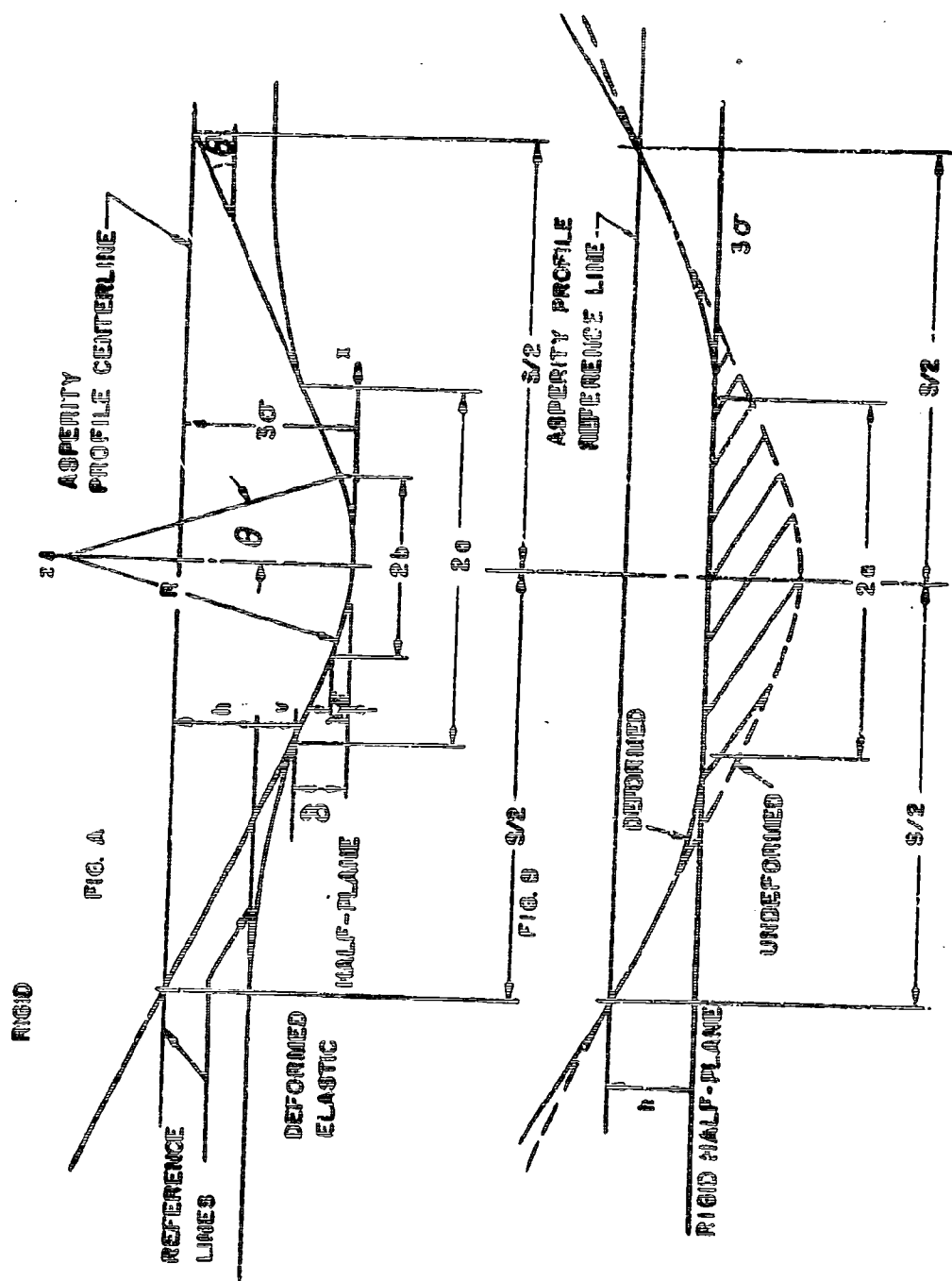


Figure 15. A Simple Model of Asperity Contact

profile, the assumption is arbitrary because the valleys of the asperity, are not defined.

The vertical distance between the asperity tip and the contact edge δ is obtained from the geometry of Figure 15, which is

$$\delta = \delta_0 + (a-b) \tan \theta \quad (8.4)$$

where $\delta_0 = R (1 - \cos \theta)$, the depth of the asperity tip and

$$b = R \sin \theta \quad (8.5)$$

One may consider a multitude of rough surfaces with asperities of the shape shown in Figure 15 with different values of σ . A reasonable assumption for the degree of their rounding at the tip is that $\delta_0 = \lambda \sigma$ where λ is a constant, proportional to R , i.e. that the rounding occupies a constant fraction of the asperity height between tip and profile center-line.

From Equation (8.5) one obtains, using $\delta_0 = \lambda \sigma$

$$R = \lambda \sigma \cdot (1 - \cos \theta)^{-1} \quad (8.6)$$

$$b = \lambda \cdot \sigma \cdot \sin \theta \cdot (1 - \cos \theta)^{-1} \quad (8.7)$$

The deformed profile of the elastic half plane in contact with the rigid asperity is given by the following formula (4)

$$\begin{aligned} v(x) &= \frac{1}{\pi} \int_a^x \frac{\sqrt{x^2 - a^2}}{t-x} \cdot \int_a^b \frac{f'(t) dt}{\sqrt{a^2 - t^2}} dx \\ &= \frac{1}{\pi} \int_a^b \frac{f'(x)}{\sqrt{a^2 - x^2}} \int_a^x \frac{\sqrt{x^2 - a^2}}{t-x} dx dt \end{aligned} \quad (8.8)$$

$$\begin{aligned} \text{where } f'(t) &= t/R & |t| < b \\ &= b/R, & b < |t| < a \end{aligned}$$

$$\text{or } v(x)/a = a/R \cdot f(x/a, b/a) \quad (8.9)$$

The above formula can be integrated after expanding the integrand $\sqrt{a^2 - x^2}/(t-x)$ into an ascending power series in a/x . The computed dimensionless deformation outside the contact region is plotted against x/a in Figure 16 using the ratio b/a as the parameter.

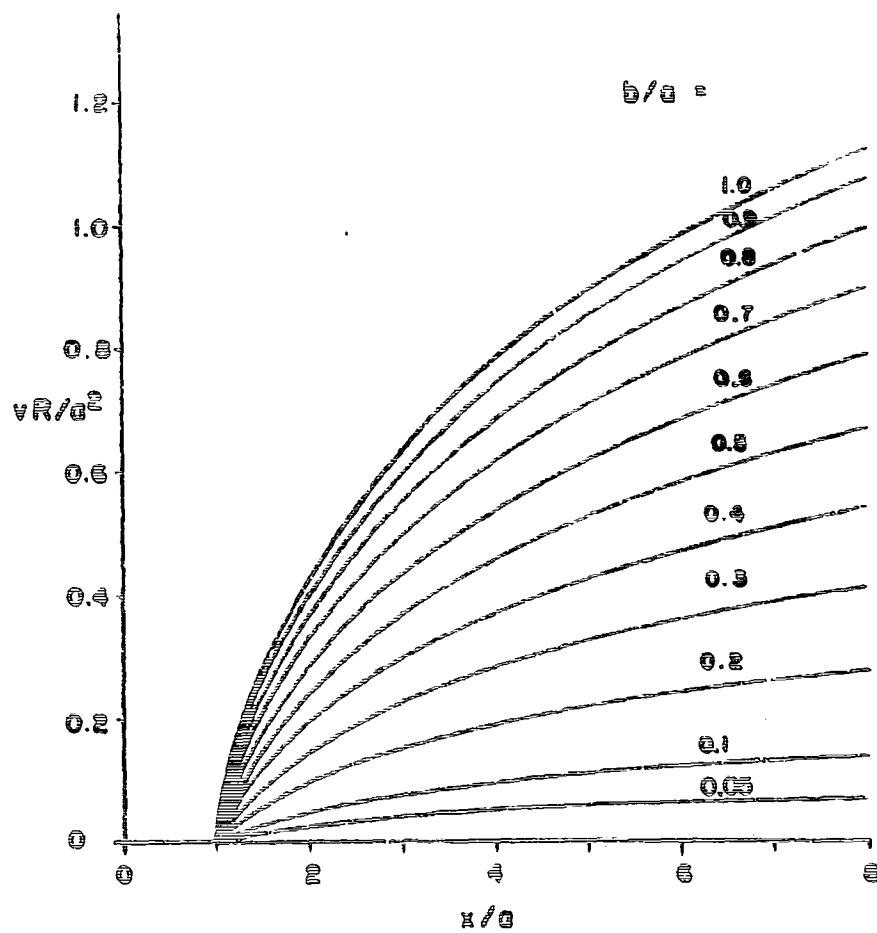


Figure 16. Variation of Surface Deformation Outside the Contact Zone of the Simple Asperity Model

In case of a real asperity contact between two steel surfaces, the asperity is not rigid. It has the same elastic properties as the half plane. In addition, both surfaces have asperities. For (39), a convenient approximation of the contact phenomena is arrived at by considering one surface having a roughness profile equal to the composite of the two actual surface roughnesses, while the other surface is flat. (37) shows how the composite roughness can be calculated for undeformed asperities (among other formulas, one finds $\sigma^2 = \sigma_1^2 + \sigma_2^2$ as quoted before). Assuming that the same composite forming procedure is acceptable for deformation calculations, Figure 15 (A) represents an asperity of the composite roughness. In order to calculate deflection, this composite asperity must be considered elastic and the smooth half plane rigid. Assuming low slopes and large radii for the asperity, the deflections shown in Figure 15 for the elastic half-plane will still apply but must now be measured separately from the undeformed asperity profile. In order to be able to do so, horizontal reference lines must be established on the deformed half plane (the entire surface of the infinite half plane in contact is, strictly speaking, deformed and the magnitude of the deflection does not converge). A reasonable reference line is the horizontal line drawn through two symmetrical points on the profile at a distance of $S/2$ from the line of symmetry of the asperity, where S is the asperity spacing. Deformation of the surface beyond these points is then to be neglected. For a symmetrical asperity profile, the profile points at a distance of $S/2$ from the asperity tip lie on the profile centerline. The use of these points as a deformation reference implies that all asperity deflection occurs in the portion from the centerline to the tip.

With these reference points established, one obtains the deformed shape of the asperity by subtracting the shaded area from its undeformed shape. (See Figure 15 (B).)

The average EHD film thickness, h , between two rough surfaces has been defined (37) as the distance between their profile centerlines. If the composite roughness profile is used, then h is the distance between its centerline and the smooth reference surface. By this definition one obtains from Figure 10:

$$h + \delta + v(S/2) = 3\sigma \quad (0.10)$$

where 3σ is the undeformed asperity height from the profile centerline to the tip.

In dimensionless form, this equation can be written by means of Equations (8.4) and (8.7) as

$$\begin{aligned}\frac{h}{\sigma} &= 3 - \delta/\sigma - v(S/2)/\sigma \\ &= 3 - \delta b/\sigma - (a-b) \cdot \tan \theta/\sigma - v(S/2)/\sigma \\ &= 3 - \lambda - \lambda (a/b-1) \cdot (1+\cos \theta) / \cos \theta - v(S/2)/\sigma \quad (8.11)\end{aligned}$$

The known function $v(x)$ in Equation (8.9) can be rearranged in the following form:

$$\frac{v}{\sigma} \cdot \left(\frac{\sigma}{R}\right) \cdot \left(\frac{b}{a}\right) \cdot \left(\frac{R}{b}\right) = \left(\frac{R}{b}\right) \cdot \left(\frac{b}{R}\right) \cdot f\left(\frac{x}{b}, \frac{b}{a}, \frac{b}{a}\right) \quad (8.12)$$

By means of Equations (8.5) and (8.6) one has

$$v(S/2)/\sigma = \lambda(1+\cos \theta) \cdot \left(\frac{a}{b}\right)^2 f\left(\frac{S}{2b}, \frac{b}{a}, \frac{b}{a}\right) \quad (8.13)$$

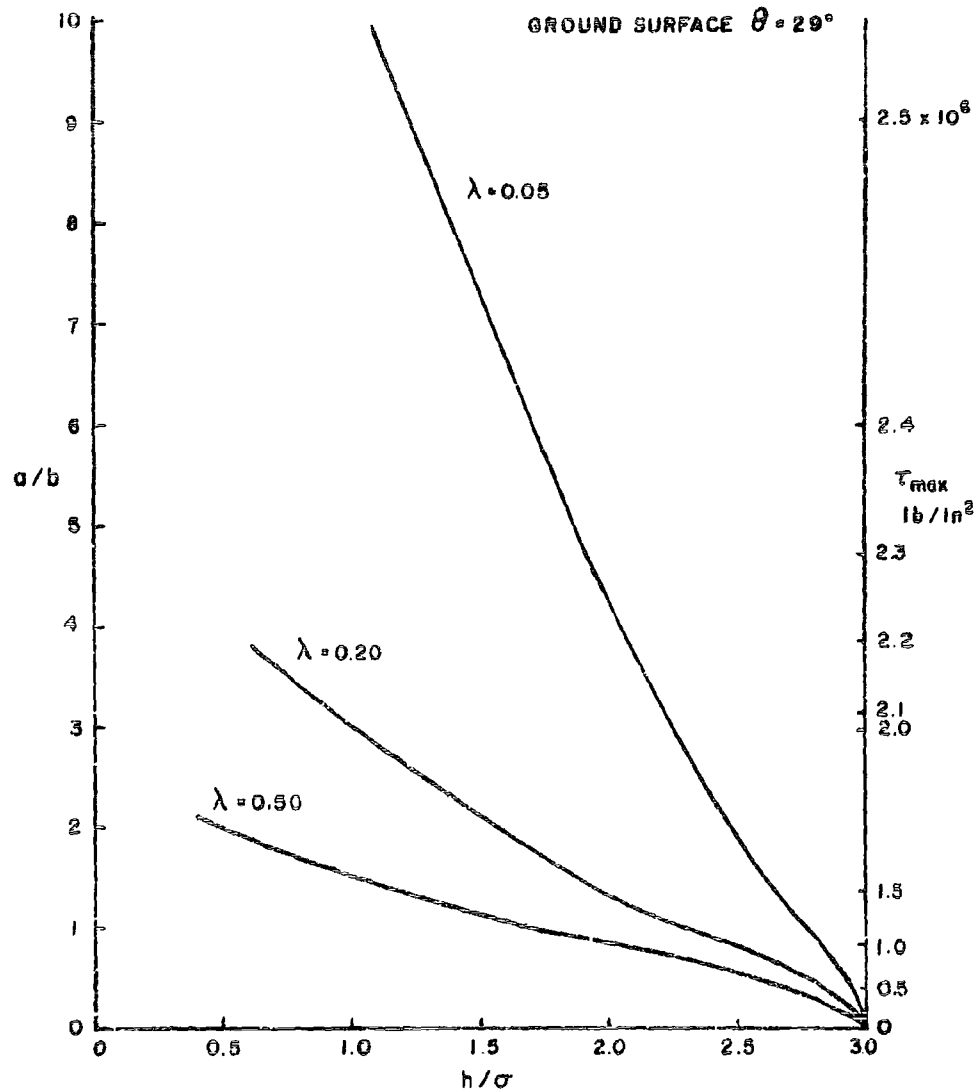
For given values of λ , θ and σ , one can determine h/σ as a function of a/b (or vice-versa) by using Equations (8.11) and (8.13).

Computation of h/σ has also been performed for the case that $a/b \leq 1$, i.e. the region of contact of the asperity falls within the curved tip without touching the straight sides of the asperity profile. This case occurs only when h/σ is relatively large but less than 3. (It is noted that in the present model, there is no asperity contact for $h/\sigma \geq 3$).

Figures 17, 18 and 19 plot, for the ground, honed and lapped surface finishes discussed above, a/b as a function of h/σ . The parameter for each curve is the relative tip width, which is a function of the tip radius R .

The curves show the general trend that b/a decreases with decreasing h/σ . Since b is a constant, decreasing b/a means increasing the contact width a . Since the maximum shear stress increases with increasing a (or decreasing b/a), one also sees that for a given asperity profile the maximum shear stress in the asperity increases with decreasing h/σ .

ASYMPTOTIC VALUE
 $= 2.74 \times 10^6 \text{ PSI}$

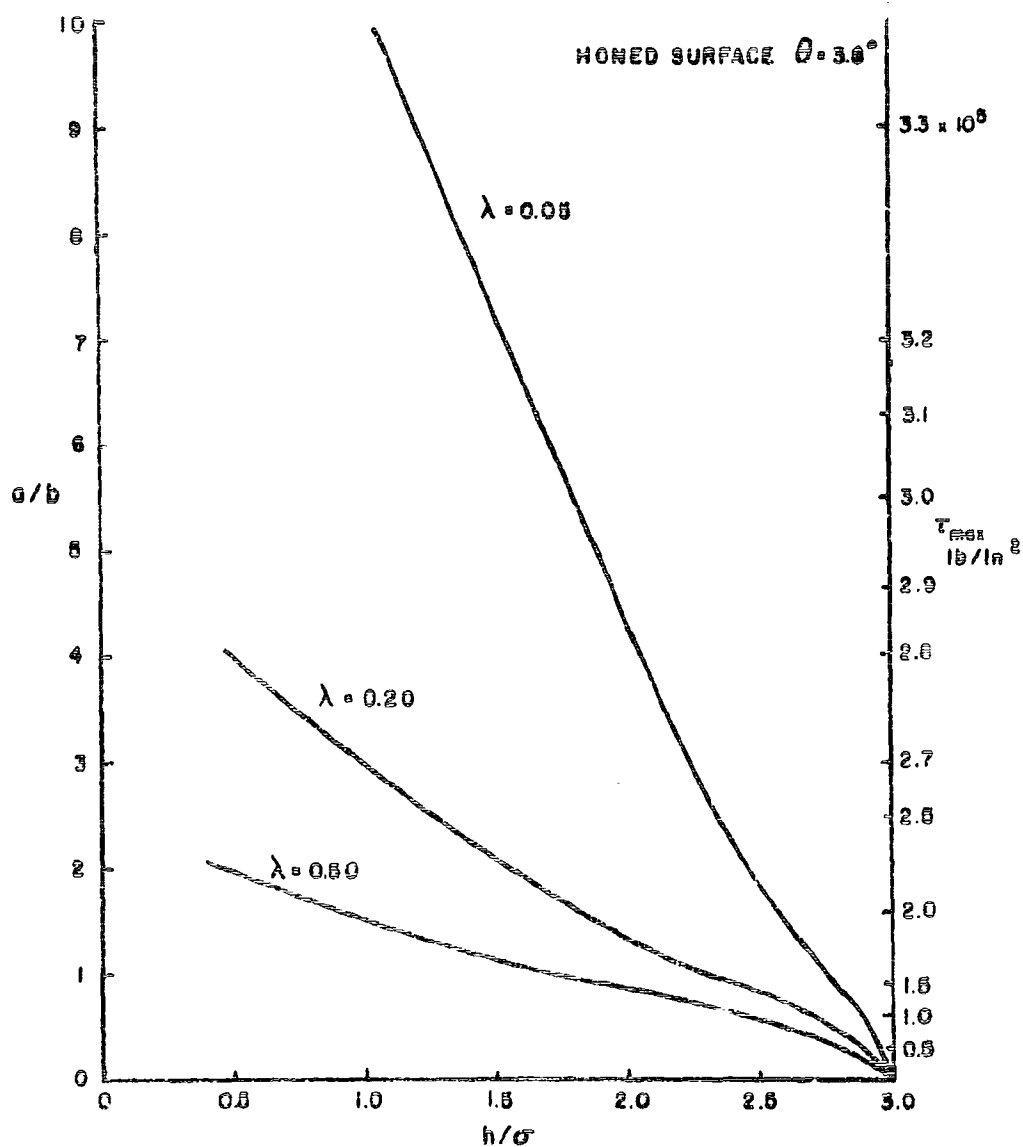


FILM THICKNESS/COMPOSITE SURFACE ROUGHNESS RMS

Figure 17. Variation of $\tau_{\max}$ and a/b with h/σ for a Ground Surface

ASYMPTOTIC VALUE

$$= 3.0 \times 10^6 \text{ PSI}$$



FILM THICKNESS / COMPOSITE SURFACE ROUGHNESS RMS

Figure 18. Variation of $\tau_{\max.}$ and a/b with h/σ for a Honed Surface

ASYMPTOTIC VALUE
 $\approx 9.61 \times 10^4 \text{ PSI}$

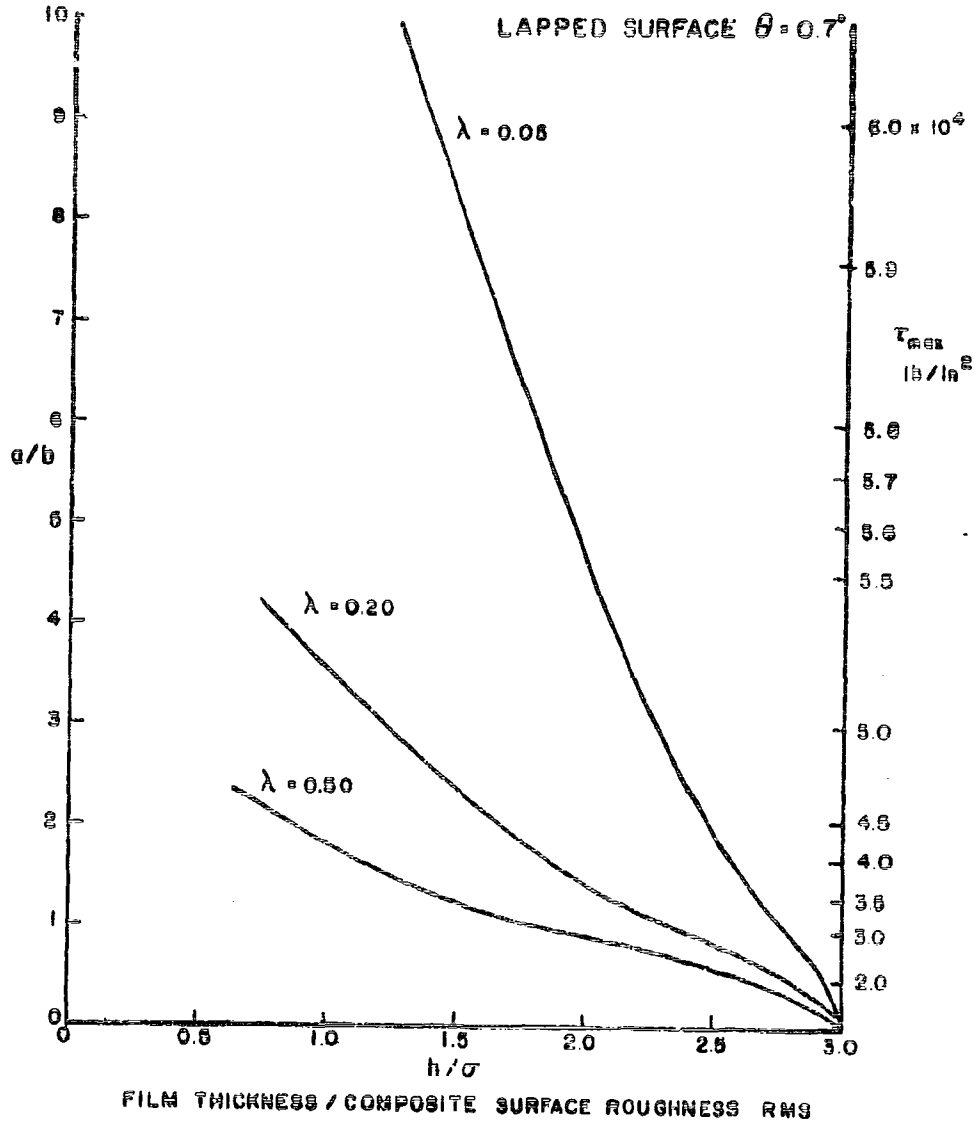


Figure 19. Variation of τ_{max} and a/b with h/σ for a Lapped Surface

It is also shown in Figures 17 - 19 that for a given h/σ value, decreasing values of λ or R (sharper tipped asperities) correspond to increasing a/b or the contact width ' a ' and shear stress τ_{max} . However, it is recalled that for large values of ' a ' (or $b/a \rightarrow 0$), τ_{max} will reach an asymptotic value independent of h . The relationship between b/a and τ_{max} for the three surface finishes discussed here is plotted in Figure 20, which permits the scaling of the ordinates of Figures 17-19 in terms of τ_{max} , in addition to h/a . This has been done by showing a nonlinear ordinate scale on these plots.

From the above results based on a simple plane asperity interaction model, it is possible to relate the maximum near surface shear stress to three parameters explicitly given for a given lubrication condition and surface finish, namely h/σ , θ and R/σ where h = minimum film thickness, σ = combined surface roughness (rms), θ = asperity slope angle and R = asperity tip radius.

It can be seen from Figure 17 that for ground surfaces with a typical asperity slope angle $\theta = 29^\circ$, the maximum shear stress level is quite high, i.e. at $h/\sigma = 2$ and $\lambda = 0.5$, the maximum shear stress is $\sim 10^6$ psi which is considerably higher than the magnitude of maximum Hertzian shear stress usual in rolling contacts. It is expected that plastic flow will occur causing a "blunting" of the asperity tip. For a smoother surface finish, e.g., a honed surface with $\theta \approx 3.8^\circ$ and $\sigma \approx 1.0$ μ in., it can be seen from Figure 18 that the maximum shear stress is considerably lower than that of a ground surface. For example, at $h/\sigma = 2$ and $\lambda = 0.5$, the maximum shear stress $\sim 1.5 \times 10^5$ psi, which is of the same order of magnitude of maximum Hertzian shear stress. Thus severe plastic deformation is not expected in the asperity.

It is seen (by using Figure 19) that still lower maximum shear stress will occur in lapped surface, e.g. at $h/\sigma = 2.0$, the maximum shear stress in this case is only 3×10^4 psi (which is lower than the yield strength of hard steel) and no plastic deformation (or surface distress) should occur.

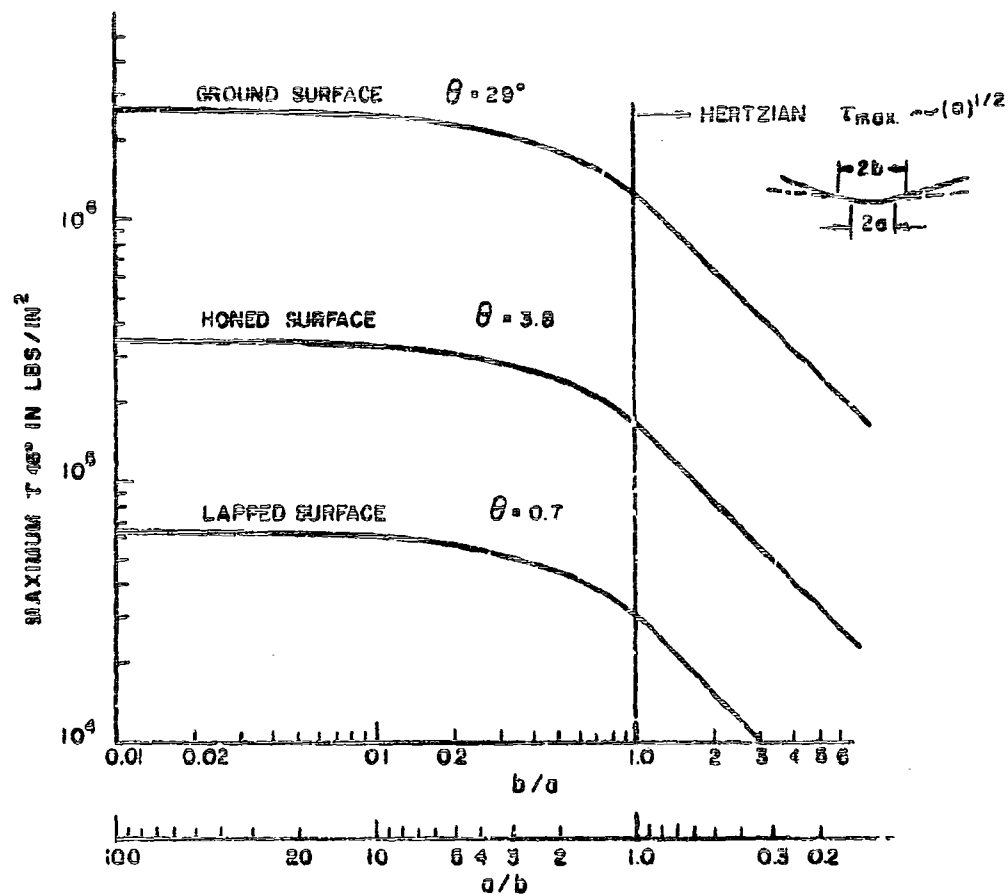


Figure 20. Variation of τ_{max} with a/b

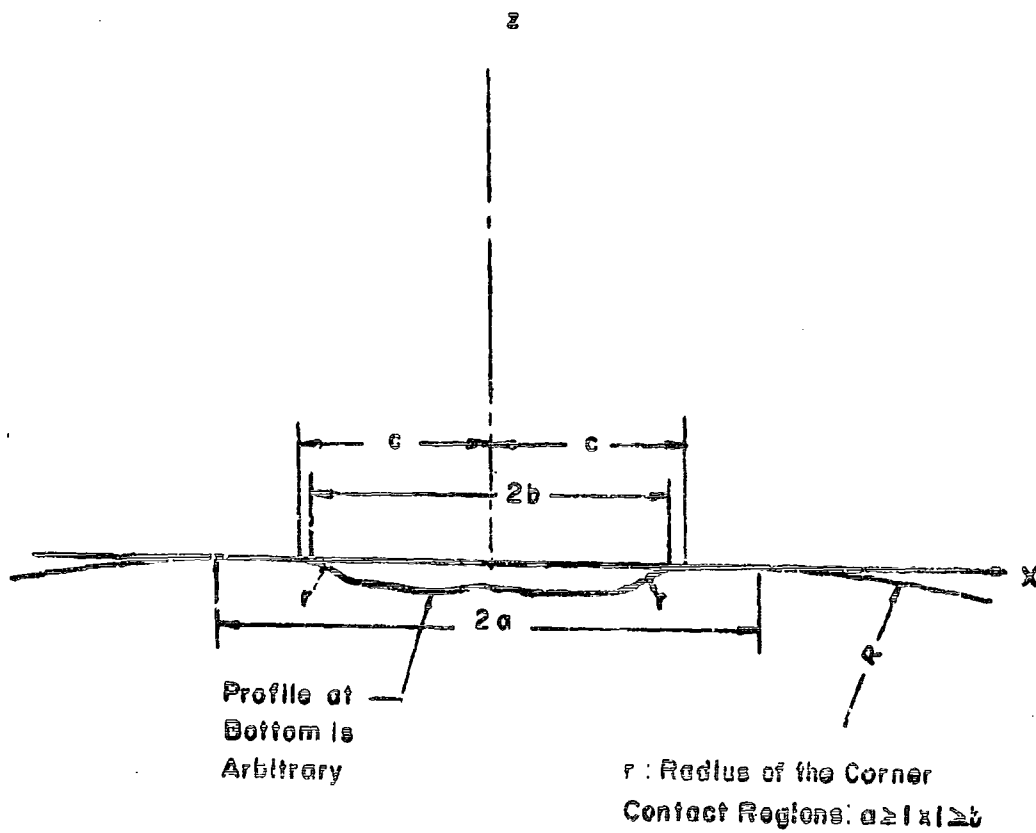


Figure 21. Schematic Representation of an Idealized Surface Defect

SECTION IX

DETERMINATION OF SHEAR STRESS BENEATH A FURROW

A two dimensional, exploratory analysis of the contact stress of bodies containing an idealized surface micro-defect as shown in Figure 21 has been conducted.

The defect in Figure 21 is a two dimensional "depression", i.e. infinitely long and Figure 21 shows its cross-section, the rims of which are formed by two radii, tangential to the surrounding (original) surface at points $2c$ apart. The Hertz area is $2a$ wide where $a > c$, and the contact does not extend closer to the bottom of the defect but rather, there is a free surface of width $2b$ in the defect, whereby $c > b > 0$, i.e. the free surface at the bottom of the defect is free, its shape is irrelevant provided that it is sufficiently depressed not to contact the opposite body.

In addition, the profile is assumed smooth, having radii of curvature at all points considerably greater than the characteristic dimensions of the defect. The contact region consists of two portions due to the presence of the open cavity, covering the cross-sectional co-ordinates $b < x < a$ and $-a < x < -b$ where the values of a and b ($< c$) are determined by the defect geometry parameters, i.e. r and c , and by the load.

This contact problem can be solved by applying the Muskhelishvili theory of complex variables to the mixed boundary value problem of an elastic half plane with multi-contact zones (36).

In rolling elements, the surface defects are small such that the characteristic dimensions, c and r , of the defects are of a smaller order than the linear dimensions of the rolling elements. Therefore, it is justified to consider a limiting case wherein $c/R \rightarrow 0$ and $r/R \rightarrow 0$. Using this assumption, the problem reduces to the compression of two straight-edged bodies, one of which contains a shallow surface defect, as shown in Figure 22. Frictional traction on the surface is neglected. The contact pressure at the interface when the defect is absent is assumed to be p_0 . In the presence of a defect, the contact pressure at points removed from the defect is expected to approach asymptotically the undistorted value p_0 as x approaches infinity.

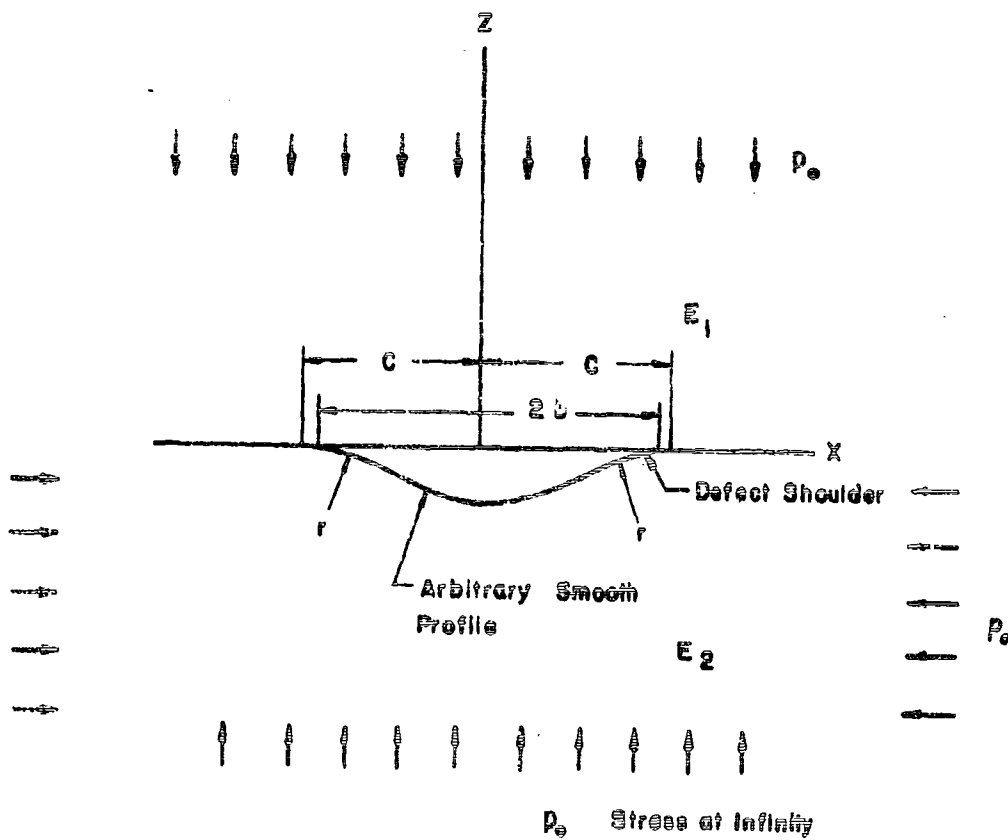


Figure 22. Schematic Representation of Contacting Bodies Containing Surface Defects (Limiting Case $R \rightarrow \infty$)

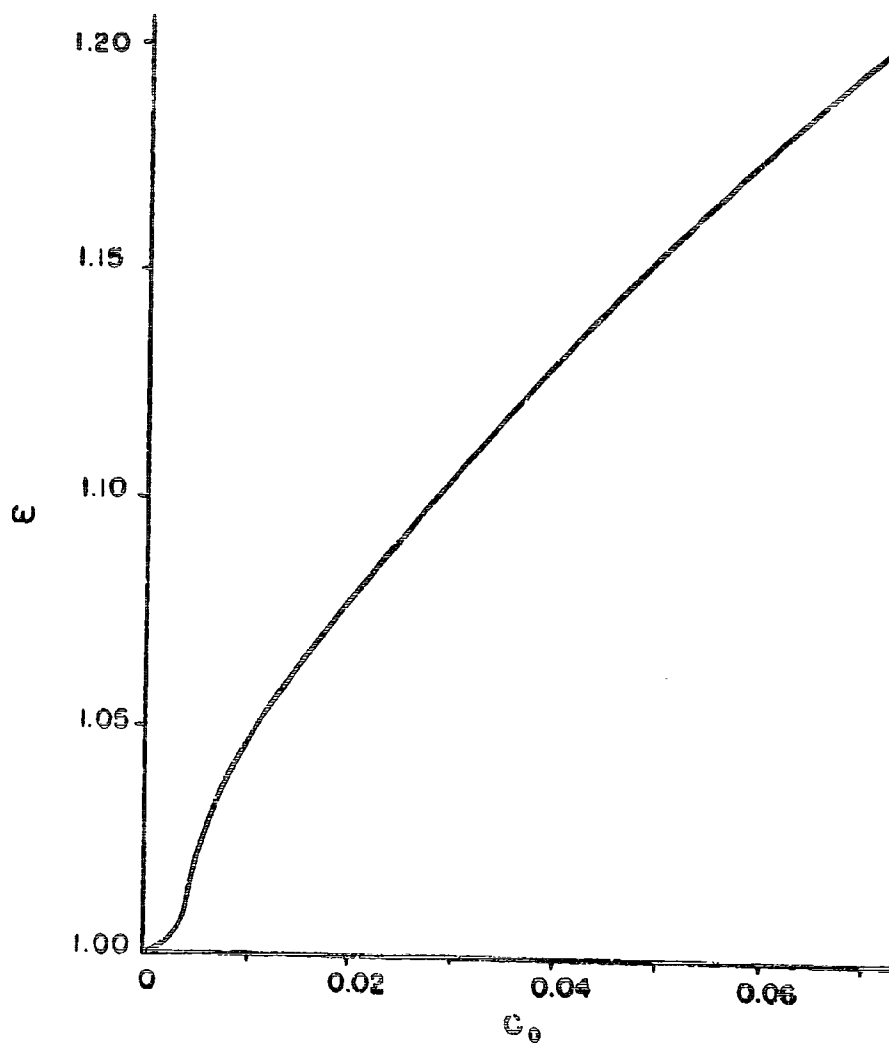


Figure 23. Variation of ϵ ($= c/b$, or defect width/distance between two contact edges) with C_0

1. COMPUTATION OF CONTACT PRESSURE NEAR A SURFACE DEFECT

A closed form solution has been derived in Appendix IV for this limiting case. The expression for the contact pressure is as follows:

$$p(x) = (E'/2\pi^2 r) \left[I(x) - I(-x) \right] \quad (9.1)$$

where

$$I(x) = (x-c) \cdot \log \left| \frac{(x-b) \cdot \eta - \sqrt{x^2 - b^2}}{(x-b) \cdot \eta + \sqrt{x^2 - b^2}} \right|$$

and $E' = \left[(1-\nu_1^2)/\pi E_1 + (1-\nu_2^2)/\pi E_2 \right]^{-1}$ the reduced Young's modulus

(For steels, $E' = 51.8 \text{ lb/in}^2$)

$$\eta = \tan \left(0.5 \cdot \cos^{-1} \frac{b}{c} \right)$$

The value of b is determined from the following formula:

$$\log \left(\sqrt{w^2 - 1} + w \right) - \sqrt{w^2 - 1}/w = C_0 \quad (9.2)$$

where

$$w = c/b$$

and $C_0 = \pi^2 p_0 r / c E'$, a dimensionless parameter.

Figure 23 plots the relationship between w and C_0 . Since C_0 can be computed from the known variables p_0 and r/c , it is possible to obtain $b = c/w$ where w is obtained for a given value of C_0 from Figure 23.

A numerical example has been computed for the distribution of contact pressure on the surface, assuming $c/b = 1.20$, corresponding to $C_0 = 0.069$. The values can be obtained by setting (1) $r/c = 1.8$ and $p_0 = 2 \times 10^5 \text{ psi}$ or (2) $r/c = 0.9$ and $p_0 = 4 \times 10^5 \text{ psi}$. Figure 24 plots the dimensionless pressure p/p_0 as a function of x . The results show that there is a pressure rise in the vicinity of the defect edge, reaching a value of 3.06, i.e. there is a significant concentration of pressure at the defect edge.

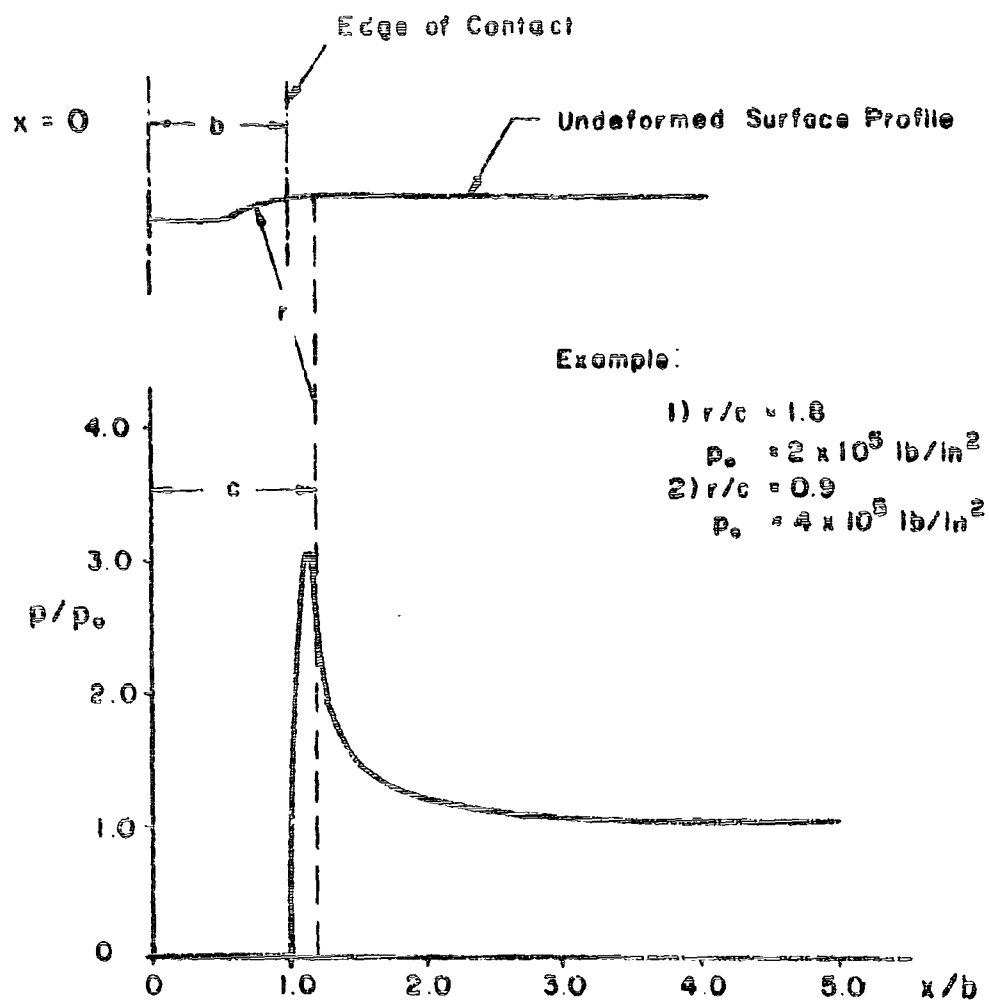


Figure 24. Pressure Distribution Near a Surface Defect
 (for $c/b = 1.2$)

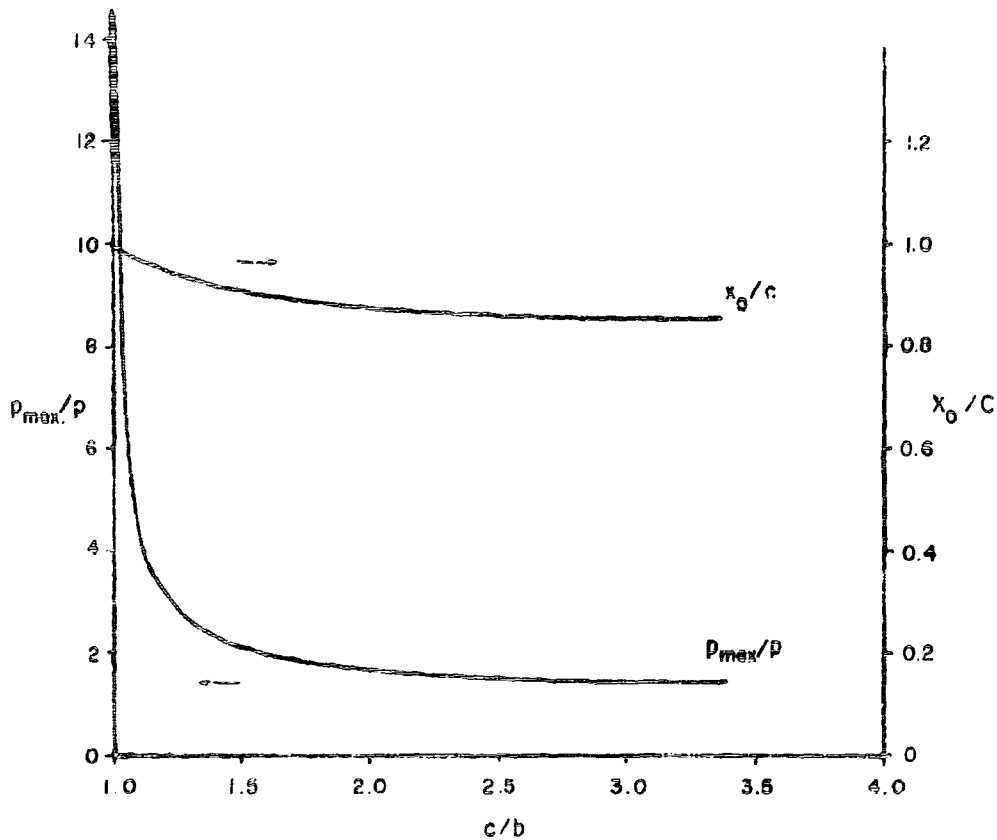


Figure 25. Variation of Maximum Contact Pressure and its Location Coordinate with Dimensionless Parameter c/b for an Idealized Defect

Figure 25 shows the distance x_0 of the pressure peak arising as a result of the defect as measured from the defect center and the magnitude of the maximum contact pressure in terms of p_0 plotted as a function of c/b . The result shows the high pressure peaks occur within the area of radius r . Based on Figures 24 and 25, Figure 26 plots the variation of p_{max} as a function of r/c for steel, using four values of p_0 , i.e., 1×10^5 , 2×10^5 , 3×10^5 and 4×10^5 psi as parameters. It can be readily seen that the maximum contact pressure increases with decreasing values of r/c .

It can be seen from Figure 26 that all the curves approach constant values of p_{max}/p_0 (>2.5) when r/c is greater than 2. It can be concluded that for all values of r relatively small compared with the rolling element size and $p_0 < 4 \times 10^5$ psi, we have $p_{max}/p_0 > 2.5$.

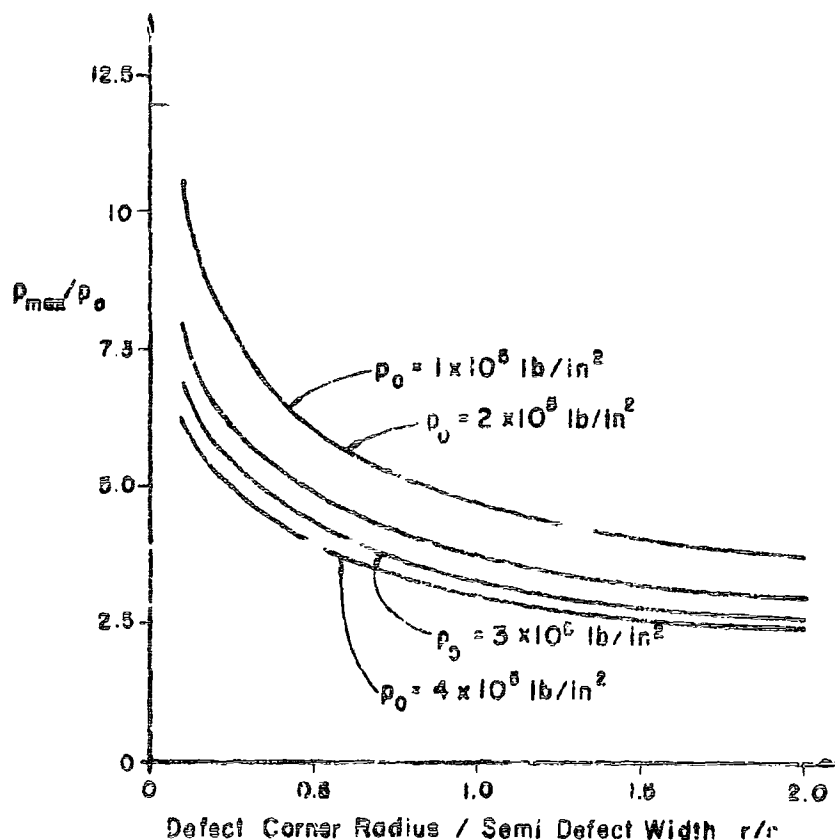


Figure 26. Variation of Maximum Contact Pressure as a function of Defect Geometric Parameter r/c and Nominal Pressure p_0

2. SUBSURFACE STRESS DISTRIBUTION IN THE VICINITY OF A SURFACE DEFECT

It is of interest to know the stresses existing near the high pressure peak at $x \neq 0$. The sub-surface stress distribution can be computed using a numerical integration technique based on the solution for the stress field on a half plane under a concentrated normal load. The numerical method requires that the region of surface loading be finite in width along the x axis. This can be arranged by resolving the surface pressure into one component which occupies a finite width and a uniform pressure acting on the surface of the entire half plane as shown in Figure 27.

Figure 28 plots the contours of equal von Mises yield stress σ_D for a typical surface defect with parameters $c/b = 1.2$ corresponding to $r/c = 0.9$ for $p_0 = 4 \times 10^5$ psi or $r/c = 1.8$ for $p_0 = 2 \times 10^5$ psi in steel. It can be seen that the maximum value of τ_{oct} or octahedral shear stress occurs under the location of maximum surface pressure. Thus $(\sigma_D)_{max}$ is about $1.36 p_0$ which is considerably higher than $0.7 p_0$ occurring at the axis of symmetry ($x = 0$) shown in Figure 28. Furthermore, as shown in Figure 28, the depth of $(\sigma_D)_{max}$ is $\sim 0.13 c$ which is considerably smaller than $0.9 c$ which occurs along $x = 0$.

In the above defect originated stress analysis, the determining parameters are found to be the ratio of defect width $2c$ and corner radius r and the undisturbed surface pressure p_0 . The degree of stress concentration at the shoulder increases with increasing defect width and decreasing defect shoulder radius. Although subsurface maximum shear stress has not been computed for a wide range of r/c values, an example has been computed, corresponding to a typical realistic defect size. The result shows that the contact pressure reaches a peak value of 3.1 times p_0 whereas the maximum value of the quantity giving the von-Mises yield criterion is $1.3 p_0$ (compared with $0.32 p_0$ in the Hertzian contact). It is expected that significant plastic deformation will occur at the defect shoulder.

Surface plastic deformation has been observed at asperities (on ground and at times on honed surfaces) at low h/c values, and at many surface defect shoulders. It is believed that this is due to the high shear stress predicted above for both of these surface failure origination points. In the case of severe asperity interaction the plastic deformation occurs at the tip causing a decrease of asperity slope. In the case of surface

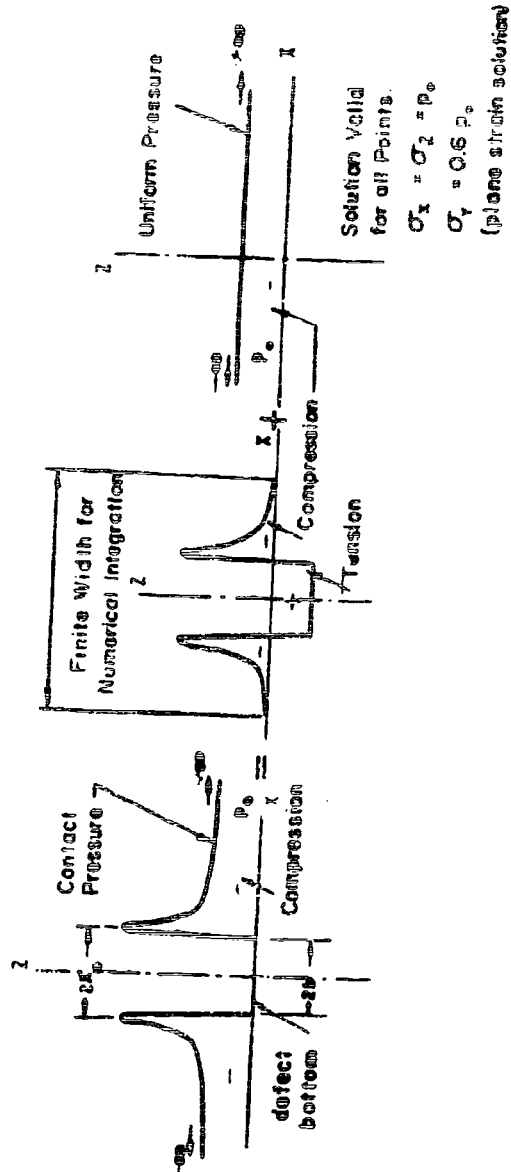


Figure 27. Application of Method of Superposition for Numerical Integration to Obtain Sub-surface Stress Distribution under a Surface Defect

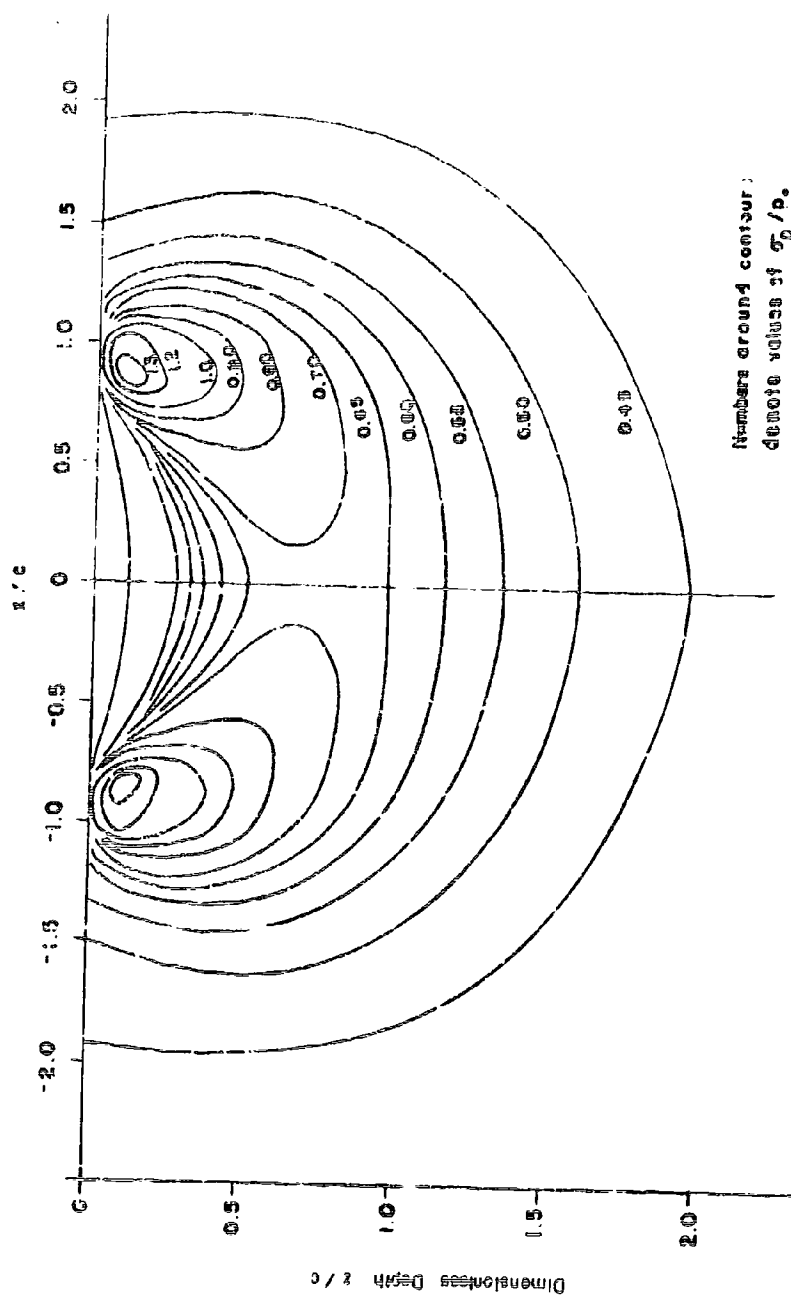


Figure 28. Contours of Equal Von-Mises Yield Criterion in Material around an Ice-cracked Surface Defect

defects, sharp corner radii (compared with defect width) have been observed in new surfaces. After many cycles of rolling, it has been observed by surface tracing that the corner radius increases considerably as a result of "rounding off".

It is plausible to assume that the surface profile of asperities and defect shoulders stabilizes, i.e. the plastic deformation ceases to grow after a certain number of stress cycles. This is a kind of "shakedown" process on the microscopic scale. Eldredge and Tabor (38) and Johnson (32) in studying macroscopic shakedown of ball tracks have concluded that after a certain shakedown is reached the material will behave elastically. In spite of the large difference in scale, the basic mechanism involving shakedown in bearing rolling tracks and/or asperity tips (or defect shoulders) can be similar. Using this argument it may be assumed that the asperity tip and defect shoulder will behave elastically after a stabilized surface is achieved. Based on Figures 17 - 19 and 25 it is seen that a run-in asperity (or defect surface) having acquired a smaller asperity slope (or a larger defect corner radius) suffers a lower degree of stress concentration than when it is new. From this fact it can be said that surface plastic deformation flows in the direction required to reduce the degree of stress concentration. In hard steel, work hardening is high and the amount of deformation at "shakedown" will be limited. It is possible that the run-in shape of asperity tips and defect shoulders will continue to have some stress concentration.

APPENDIX I

FORMULAS FOR STRESSES IN A HERTZIAN STRESS FIELD

The stresses at an arbitrary point below the surface of a bearing ring induced by contact of a ball or roller are computed under the assumption that areas of the ring and rolling body are large enough compared to the size of the contacting area that the contacting bodies may be considered infinite in extent.

In the vicinity of the contact the surfaces are assumed to be describable by second degree polynomials.

All of the stresses concerned are referred to a rectangular Cartesian coordinate system with the xy-plane fixed on the boundary surface of the semi-infinite body with the z-axis directed into the body. The xz and yz planes coincide with the symmetry planes of the contact ellipse, the equation of which is,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (A1-1)$$

where a is the major semi-axis lying on the x-axis and b the minor semi-axis lying on the y axis.

The stresses are given by (34).

$$\frac{\sigma_x}{\sigma_0} = Q \left(\frac{LX}{A^2 + L^2} \right)^2 + (1-2\nu) N_x - 2(1-\nu) \frac{ZM_x}{L} + 2\nu \frac{ZM_z}{L} \quad (A1-2)$$

$$\frac{\sigma_y}{\sigma_0} = Q \left(\frac{LY}{1+L^2} \right)^2 + (1-2\nu) N_y - 2(1-\nu) \frac{ZM_y}{L} + 2\nu \frac{ZM_z}{L} \quad (A1-3)$$

$$\frac{\sigma_z}{\sigma_0} = Q \left(\frac{Z}{L} \right)^2 \quad (A1-4)$$

$$\frac{T_{xy}}{\sigma_0} = Q \left(-\frac{LX}{A^2 + L^2} \right) \left(\frac{LY}{1+L^2} \right) - (1-2\nu) N \quad (A1-5)$$

$$\frac{\tau_{xz}}{\sigma_0} = Q \frac{ZX}{A^2 + L^2} \quad (A1-6)$$

$$\frac{\tau_{yz}}{\sigma_0} = Q \frac{ZY}{1+L^2} \quad (A1-7)$$

in which,

$$A = \frac{a}{b}, \quad X = \frac{x}{b}, \quad Y = \frac{y}{b}, \quad Z = \frac{z}{b} \quad (A1-8a)$$

$$\sigma_0 = \frac{-3P}{2\pi ab} \quad (A1-8b)$$

and P is the total load

L is the largest positive root of the following equation

$$\frac{X^2}{A^2 + L^2} + \frac{Y^2}{1 + L^2} + \frac{Z^2}{L^2} = 1 \quad (A1-8c)$$

$$Q = \frac{Z}{L} \cdot \frac{A}{\sqrt{(A^2 + L^2)(1 + L^2)}} \cdot \frac{1}{\left(\frac{LX}{A^2 + L^2}\right)^2 + \left(\frac{LY}{1 + L^2}\right)^2 + \left(\frac{Z}{L}\right)^2} \quad (A1-8d)$$

$$\Phi = \frac{1}{\sqrt{A^2 - 1}} \arctan \left[\frac{Y \sqrt{A^2 - 1}}{1 + L^2 + \frac{Z}{L} \sqrt{(A^2 + L^2)(1 + L^2)}} \right] \quad (A1-8e)$$

$$\Psi = \frac{1}{\sqrt{A^2 - 1}} \operatorname{arctanh} \left[\frac{X \sqrt{A^2 - 1}}{A^2 + L^2 + \frac{Z}{L} \sqrt{(A^2 + L^2)(1 + L^2)}} \right] \quad (A1-8f)$$

$$N = \frac{A}{A^2 - 1} (X \Phi - Y \Psi) \quad (A1-8g)$$

$$N_x = \frac{A}{A^2 - 1} \left(1 - \frac{Z}{L} \sqrt{\frac{1 + L^2}{A^2 + L^2}} - Y\Phi - X\Psi \right) \quad (A1-8h)$$

$$N_y = \frac{A}{A^2 - 1} \left(\frac{Z}{L} \sqrt{\frac{A^2 + L^2}{1 + L^2}} - 1 + Y\Phi + X\Psi \right) \quad (A1-8i)$$

$$M_x = \frac{L(F - E)}{A^2 - 1} \quad (A1-8j)$$

$$M_y = \frac{L(A^2 E - F)}{A^2 - 1} - \frac{L^2 A}{\sqrt{(A^2 + L^2)(1 + L^2)}} \quad (A1-8k)$$

$$M_z = A \sqrt{\frac{1 + L^2}{A^2 + L^2}} - LE \quad (A1-8l)$$

where F and E are the ordinary elliptic integrals of the first and second kind, respectively, with modulus $k = \sqrt{1 - \frac{1}{A^2}}$ and argument $\theta = \arctan \frac{A}{L}$; ν is

Poisson's ratio, assumed to be 0.3 in all of the numerical calculations.

For points in the contact, i.e. $Z = 0$, the largest positive root L is zero. It follows from equation (A1-8c) that

$$\lim_{\substack{Z \rightarrow 0 \\ L \rightarrow 0}} \frac{Z}{L} = \sqrt{1 - \frac{X^2}{A^2} - Y^2} \quad (A1-9)$$

Using this relation, the stress formulas can easily be obtained from equations (A1-2) through (A1-7) as follows:

$$\frac{\sigma_x}{\sigma_0} = (1 - 2\nu) N_x + 2\nu \sqrt{1 - \frac{X^2}{A^2} - Y^2} \quad (A1-10)$$

$$\frac{\sigma_y}{\sigma_0} = (1 - 2\nu) N_y + 2\nu \sqrt{1 - \frac{X^2}{A^2} - Y^2} \quad (A1-11)$$

$$\frac{\sigma_z}{\sigma_0} = \sqrt{1 - \frac{X^2}{A^2} - Y^2} \quad (A1-12)$$

$$\frac{\tau_{xy}}{\sigma_0} = \dots (1-2\nu) N \quad (A1-13)$$

$$\tau_{xz} = 0 \quad (A1-14)$$

$$\tau_{yz} = 0 \quad (A1-15)$$

$$\text{where } N_x = \frac{A}{A^2 - 1} \left(1 - \frac{1}{A} \sqrt{1 - \frac{X^2}{A^2} - Y^2} - Y\phi - X\psi \right) \quad (A1-16a)$$

$$N_y = \frac{A}{A^2 - 1} \left(A \sqrt{1 - \frac{X^2}{A^2} - Y^2} - 1 + Y\phi + X\psi \right) \quad (A1-16b)$$

$$N = \frac{A}{A^2 - 1} (X\phi - Y\psi) \quad (A1-16c)$$

$$\phi = \frac{1}{\sqrt{A^2 - 1}} \arctan \left[\frac{Y \sqrt{A^2 - 1}}{1 + A \sqrt{1 - \frac{X^2}{A^2} - Y^2}} \right] \quad (A1-16d)$$

$$\psi = \frac{1}{\sqrt{A^2 - 1}} \operatorname{arctanh} \left[\frac{X \sqrt{A^2 - 1}}{A^2 + A \sqrt{1 - \frac{X^2}{A^2} - Y^2}} \right] \quad (A1-16e)$$

A, X and Y are as defined in equations (A1-8a).

APPENDIX II

FORMULAS FOR STRESSES CORRESPONDING TO AN INFINITELY NARROW CONTACT ELLIPSE IN A HERTZIAN STRESS FIELD

In the limiting case, $A \rightarrow \infty$ or $A^{-1} \rightarrow 0$, the formulas in Appendix I must be modified before they can be used in numerical calculations. The following approximations must be made in deriving the stress formulas.

It is proper that, before working on the stress formulas, attention be directed to the modifications of equations (A1-8). Designating

$$x' = \frac{x}{A} = \frac{x}{a}$$

Equation (A1-8c) becomes

$$x'^2 + \frac{y^2}{1+L^2} + \frac{z^2}{L^2} = 1 \quad (A2-1)$$

Thus L is the largest positive root satisfying equation (A2-1). Since

$$\lim_{A \rightarrow \infty} \frac{A}{\sqrt{A^2 + L^2}} = 1$$

it is evident that

$$Q = \frac{Z}{L} \cdot \frac{1}{\sqrt{1+L^2}} \cdot \frac{1}{\left(\frac{LY}{1+L^2}\right)^2 + \left(\frac{Z}{L}\right)^2}$$

Eliminating Y by means of equation (A2-1), this can further be reduced to

$$Q = \frac{Z}{L} \cdot \frac{\sqrt{1+L^2}}{L^2 (1-x'^2) + \left(\frac{Z}{L}\right)^2} \quad (A2-2a)$$

As $A \rightarrow \infty$, both ϕ and ψ approach zero. Thus

$$N \doteq \frac{x}{A} \phi = x' \phi = 0 \quad (A2-2b)$$

$$N_x = -\frac{K}{A} \psi = -K' \psi = 0 \quad (A2-2c)$$

$$N_y = \frac{Z}{L} \cdot \frac{1}{\sqrt{1+L^2}} + \frac{K}{A} \psi = \frac{Z}{L} \cdot \frac{1}{\sqrt{1+L^2}} \quad (A2-2d)$$

For the elliptic integrals the modulus k will approach unity while the argument β will be close to $\pi/2$. Therefore, it can be seen that (40)

$$\begin{aligned} \lim_{A \rightarrow \infty} \frac{F(\beta/k)}{A^2} &= \lim_{k' \rightarrow 0} k'^2 F(\beta/k) = \lim_{k' \rightarrow 0} k'^2 K \\ &= \lim_{k' \rightarrow 0} k'^2 \ln \frac{4}{k'} \\ &= -\lim_{k' \rightarrow 0} k'^2 \ln k'^2 = 0 \end{aligned}$$

$$\text{and } \lim_{A \rightarrow \infty} E = 1, \text{ where } k' = \sqrt{1-k^2}.$$

It then follows

$$M_x = 0 \quad (A2-2e)$$

$$M_y = L - \frac{L^2}{\sqrt{1+L^2}} \quad (A2-2f)$$

$$M_z = \sqrt{1+L^2} - L \quad (A2-2g)$$

Thus the stress formulas can be obtained directly from equations (A1-2) through (A1-7). These are

$$\frac{\sigma_x}{\sigma_0} = \frac{2\nu Z}{L} M_z \quad (A2-3)$$

$$\frac{\sigma_y}{\sigma_0} = Q \left(\frac{LY}{1+L^2} \right)^2 + (1-2\nu) N_y^{-2} (1-\nu) \frac{ZM_y}{L} + 2\nu \frac{ZM_z}{L} \quad (A2-4)$$

$$\frac{\sigma_z}{\sigma_0} = Q \left(\frac{Z}{L} \right)^2 \quad (A2-5)$$

$$T_{xy} = 0 \quad (A2-6)$$

$$T_{xz} = 0 \quad (A2-7)$$

$$\frac{T_{yz}}{\sigma_0} = Q \frac{ZY}{1+L^2} \quad (A2-8)$$

in which $\sigma_0 = -\frac{2p}{\pi b}$ and p is the load per unit length of the contact.

For points located in the contact zone, the largest positive root L is zero, and, from equation (A2-1) one has, in the limit,

$$\lim_{\substack{Z \rightarrow 0 \\ L \rightarrow 0}} \frac{Z}{L} = \sqrt{1 - X'^2 - Y'^2} \quad (A2-9)$$

The stress formulas follow directly from equations (A2-3) through (A2-8) by means of equation (A2-2). They are

$$\frac{\sigma_x}{\sigma_0} = 2\nu \sqrt{1 - X'^2 - Y'^2} \quad (A2-10)$$

$$\frac{\sigma_y}{\sigma_0} = \sqrt{1 - X'^2 - Y'^2} \quad (A2-11)$$

$$\frac{\sigma_z}{\sigma_0} = \sqrt{1 - X'^2 - Y'^2}$$

(A2-12)

$$\tau_{xy} = 0$$

(A2-13)

$$\tau_{xz} = 0$$

(A2-14)

$$\tau_{yz} = 0$$

(A2-15)

APPENDIX III

PLANE CONTACT OF ASPHERITIES

The governing equation relating the contact pressure $p(x)$ in the region $-a < x < a$ to a symmetrical surface profile designated by $f(x)$, is given by (36).

$$\frac{2}{E'} \int_{-a}^a \frac{p(x)}{x-t} dt = -f'(x) \quad (A3-1)$$

$$\text{where } E' = \left[(1-\nu_1^2) \nu_1 E_1 + (1-\nu_2^2) \nu_2 E_2 \right]^{-1}$$

$$f'(x) = df(x)/dx$$

Equation (A3-1) shows that the plane contact problem of two bodies having moduli E_1 and E_2 , respectively, is equivalent to that in which one of the solids is rigid ($E \rightarrow \infty$) and the other has a reduced Young's modulus, $E' (1-\nu^2) / \nu$, provided that $f(x)$ remains the same.

The above singular integral equation can be integrated for $p(x)$. For the special case that $p(\pm a) = 0$, i. e. the contact pressure is zero at the contact edges (this is valid if the profile does not have a sharp corner or ridge at the edges), the contact pressure is given by (36)

$$p(x) = \frac{E'}{2\pi} \sqrt{a^2 - x^2} \int_{-a}^a \frac{f'(t) dt}{\sqrt{a^2 - t^2} (t-x)} \quad (A3-2)$$

For the particular profile depicted in Figure 12, one has

$$\begin{aligned} f'(t) &= t/R \quad \text{for } |t| < b; \\ &= b/R \quad \text{for } a > |t| > b \end{aligned} \quad (A3-3)$$

Substituting equation (A3-3) into equation (A3-2) and carrying out the integration, the pressure distribution is obtained as follows:

$$p(x) = \frac{E'}{2\pi R} \left[2\sqrt{a^2 - x^2} \beta_0 - (x+b) \log \left(\frac{\sin \left(\frac{\beta_0 + \alpha}{2} \right)}{\cos \left(\frac{\beta_0 - \alpha}{2} \right)} \right) - (x-b) \log \left(\frac{\cos \left(\frac{\beta_0 + \alpha}{2} \right)}{\sin \left(\frac{\beta_0 - \alpha}{2} \right)} \right) \right] \quad (A3-4)$$

where

$$\beta_0 = \sin^{-1} \frac{b}{a}, \quad \alpha = \sin^{-1} \frac{x}{a}.$$

The force acting between the single asperity and the half plane is given by

$$P = \frac{E'}{2\pi} \int_{-a}^a \frac{t f'(t) dt}{\sqrt{\frac{a^2}{2} - t^2}} = \frac{E' a^2}{2\pi R} \left(\beta_0 + \frac{1}{2} \sin 2\beta_0 \right) \quad (A3-5)$$

For small values of the approach between two bodies, i.e. $\frac{b}{a}$ very small,

$$P \cong E' \beta_0^2 a^2 / 2\pi R \quad (A3-6)$$

For a symmetrical profile $f(x)$ with respect to the z -axis (or the contact center), the subsurface stress distribution at $x = 0$, or the centerline of contact, is given by (36)

$$\sigma_z - \sigma_x = 4\pi \operatorname{Im} \Phi'(\zeta) \quad (A3-7)$$

$$T_{xz} = -2z \operatorname{Re} \Phi'(\zeta) \quad (A3-8)$$

where σ_x, σ_z = normal stresses acting in x and z direction, respectively

T_{xz} = orthogonal shear stress

$\zeta = x + iz$, a complex number

$\operatorname{Im}, \operatorname{Re}$ = imaginary and real parts of a complex number

$$\Phi(\zeta) = \frac{E'}{4\pi} \frac{\sqrt{\frac{a^2}{2} - \zeta^2}}{\sqrt{\frac{a^2}{2} - t^2}} \int_{-a}^a \frac{f'(t) dt}{(t-\zeta)} \quad (A3-9)$$

and a prime denotes differentiation with respect to ζ .

Substituting equations (A3-3) and (A3-9) into equation (A3-7) and performing the integration, one obtains

$$\begin{aligned} T_{460} \Big|_{x=0} &= \frac{1}{2} (\sigma_z - \sigma_x) \\ &= -\frac{E' \theta \zeta}{\pi^2} \cdot \frac{a}{b} \left[\frac{\pi}{2} - \tan^{-1} (\gamma \tan u) - \gamma \left(\frac{\pi}{2} - u \right) \right] \end{aligned} \quad (A3-10)$$

where

$$v = \cos^{-1} \left(\frac{b}{a} \right)$$

$$\gamma = (\xi) / \left[1 + (\xi)^2 \right]^{\frac{1}{2}}$$

$$\xi = z/a$$

$\theta = b/R$, the asperity slope angle

APPENDIX IV

COMPRESSION OF A HALF PLANE CONTAINING AN IDEALIZED SURFACE DEFECT

As depicted in Figure 22, the half plane is straight edged except at the depression. The line of symmetry of the profile is chosen as the y-axis whereas the edge of the half plane is chosen as the x-axis. The corner radii of the depression, assumed to be identical, are denoted by r . The approach of the two bodies is assumed to be normal to the surface of the half plane. Thus at infinity there is an uniform compression p_0 throughout the bodies. The contact region occupies the boundary of the half plane except at the center of the depression or $|x| < b$.

Using the same notations as in Appendix III, the governing integral equation, similar to equation (A3-1) in form except for the limit of the integral, is given below

$$\frac{2}{E'} \left[\int_{-\infty}^{-b} + \int_b^{\infty} \right] \frac{p(x)}{x-t} dt = -f'(x) \quad (A4-1)$$

The derivative of the profile expression $f(x)$ with respect to x for the present problem is given by:

$$\begin{aligned} f'(x) &= 0 & \Rightarrow |x| > c \\ f'(x) &= (c-x)/r & b < x < c \\ &= (-c-x)/r & -c < x < -b \end{aligned} \quad (A4-2)$$

According to (36) the complex potential in this problem, for the case $p(\pm b) = 0$, can be given as follows:

$$\Phi(z) = \frac{E'}{4\pi^2} \sqrt{c^2 - b^2} \left[\int_{-\infty}^{-b} + \int_b^{\infty} \right] \frac{f'(t) dt}{\sqrt{t^2 - b^2} (t-\zeta)} ; \quad \zeta = x + iz \quad (A4-3)$$

The contact pressure in the contact region is given by (36).

$$p(x) = \frac{E'}{2\pi^2} \sqrt{x^2 - b^2} \left[\int_{-\infty}^{-b} + \int_b^{\infty} \right] \frac{f'(t) dt}{\sqrt{t^2 - b^2} (t-x)} \quad (A4-4)$$

Substituting equation (A4-2) into equation (A4-4) and rearranging the terms, yields:

$$p(x) = \frac{E'}{2\pi^2} \sqrt{x^2 - b^2} \cdot \frac{x}{r} \int_b^c \frac{(c-t) dt}{\sqrt{t^2 - b^2} (t^2 - x^2)}$$

$$= \frac{E'}{2\pi^2 r} [I(x) - I(-x)] \quad (A4-5)$$

where $I(x) = (x-c) \log \left| \frac{(x-b)\eta - \sqrt{x^2 - b^2}}{(x-b)\eta + \sqrt{x^2 - b^2}} \right|$

$$\eta = \tan \left[0.5 \cos^{-1} (b/c) \right]$$

In the above solution, the quantity b is unknown and has to be determined from the magnitude of the undisturbed pressure p_0 , or the pressure at $x = \infty$. For $x = \infty$, equation (A4-5) can be reduced to the following form

$$p_0 = \frac{E'}{\pi} \frac{c}{r} \left[\log \left(\sqrt{\frac{c^2}{b^2} - 1} + \frac{c}{b} \right) - \sqrt{\frac{c^2}{b^2} - 1} \cdot \frac{b}{c} \right] \quad (A4-6)$$

Since p_0 , c , r and E' are known constants, it is possible to solve for b which is required in equation (A4-5). The determination of c/b can be made by using Figure 23 plotting c/b against the dimensional quantity $C_0 = \pi^2 p_0 r / c E'$.

APPENDIX V

AVERAGE SHEAR RANGE IN THE STRESSED AREA, S , ENCLOSED BY A CONTOUR OF EQUAL SHEAR RANGE, T_R , IN A HERTZIAN ELLIPTICAL CONTACT

The relationship between the area, S , enclosed by a contour of equal shear range T and T_0 is given in equation (7-5) of Section VII as follows:

$$S = c a z_0 \left[\frac{2T_0 - T}{T} \right]^{3/4} \quad (A5-1)$$

in which a = the major axis of the contact ellipse, z_0 = the depth corresponding to maximum reversing shear stress T_0 and c = constant.

$$\text{From (A4-1)} \quad T = 2 T_0 (c a z_0)^{4/3} \left[S^{4/3} + (c a z_0)^{4/3} \right]^{-1} \quad (A5-2)$$

For a specific stressed area S bounded by a contour of equal shear range the average shear stress becomes

$$\begin{aligned} (T_R)_{av.} &= \frac{1}{S} \int_0^S T \, dS = \frac{2 T_0}{S} \int_0^S \frac{(c a z_0)^{4/3}}{S^{4/3} + (c a z_0)^{4/3}} \, dS \\ &= \frac{2 T_0 c a z_0}{S} \int_0^{S/c a z_0} \frac{dw}{w^{4/3} + 1}, \end{aligned} \quad (A5-3)$$

The upper limit of the integral can be determined from equation (A5-1) by means of Figure 10, and therefore the integral can be evaluated using a numerical integration technique.

Using equation (A5-3), values of $\frac{(T_R)_{av.}}{T_0}$ have been calculated and plotted in (Figure 32) against $S/a z_0$. It can be shown that for $S/a z_0$ smaller than 4.7, the curve can be approximated by a straight line expressed as follows:

$$\frac{S}{a z_0} = -6.45 \frac{(T_R)_{av.}}{T_0} + 12.9 \quad (A5-4)$$

Solving for $(T_R)_{av.}$ in equation (A5-4), one has

$$(T_R)_{av.} / T_0 = 0.155 (12.9 - S/a z_0) \quad (A5-5)$$

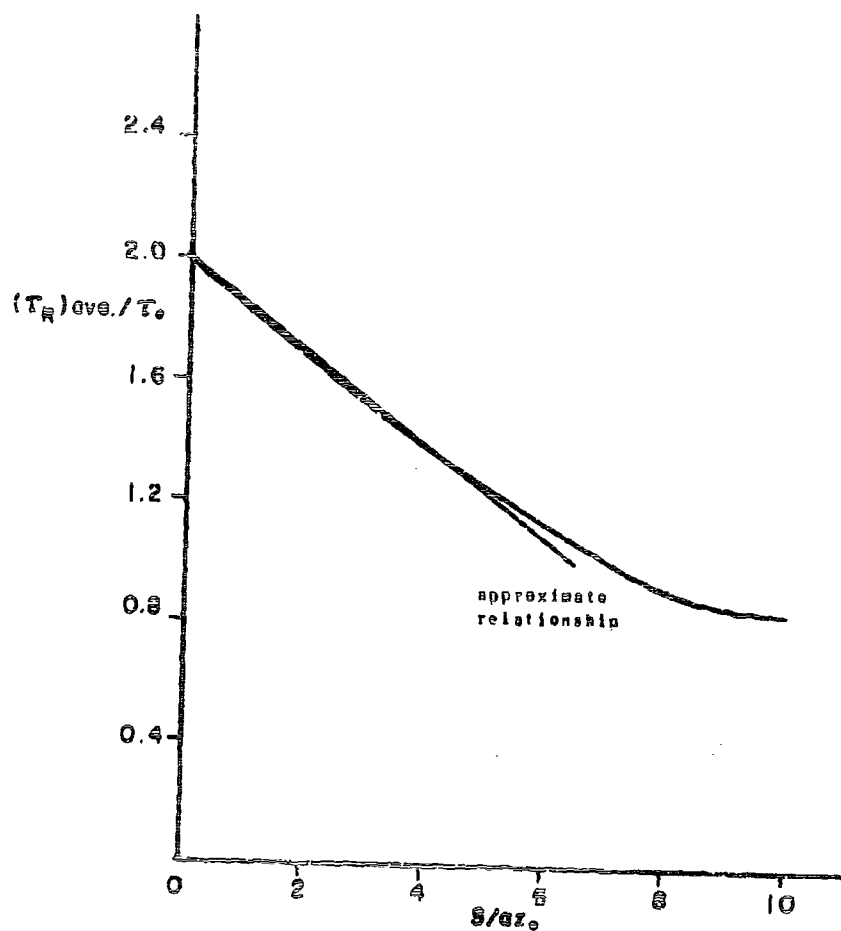


Figure 32. Variation of Average Shear Range in a Contour of Equal τ_R with Enclosed Area S

APPENDIX VI

PLAUSIBLE DEFECT SEVERITY DISTRIBUTIONS

In what follows we will explore the behavior of the function F necessary for the satisfaction of Equation (6-9) for two plausible choices of $F(d)$.

It may be reasoned that the distribution of defect severity is such that most of the defects have very small severity and a diminishing proportion have greater severities. That is, the proportion of defects with severity in the interval $d_a \pm \Delta d$ is larger than the proportion in the interval $d_b \pm \Delta d$ as long as $d_a < d_b$.

The density of the two parameter exponential distribution with unit location parameter to be given below has mode at unity and decreases with increasing variate values and thus satisfies the condition postulated above.

With this law the probability that a randomly selected defect has severity less than a value d is given by

$$\begin{aligned} \text{Prob} [d_1 < d] &= F(d) = 1 - \exp - \left(\frac{d-1}{d_0} \right); \quad d > 1 \\ &= 0; \quad d < 1 \end{aligned} \tag{A6-1}$$

The quantity d_0 is a constant parameter of the distribution related to the average severity $\bar{d}$ by $d_0 = \bar{d} - 1$. $\bar{d}$ may vary with different materials.

Another distribution that satisfies the above postulated condition is

$$\begin{aligned} F(d) &= 1 - d^{-c}; \quad d \geq 1, \quad c > 0 \\ &= 0; \quad d < 1 \end{aligned} \tag{A6-2}$$

where c is a constant parameter.

Equation (A6-2) is a form of Pareto's distribution (47), and has been found useful in economic studies. The parameter c is expressible in terms of the average defect severity $\bar{d}$ as

$$c = \frac{\bar{d}}{\bar{d} - 1} \tag{A6-3}$$

The distributions of Equations (A6-1) and (A6-2) are shown on Figure 30, plotted for $\bar{d} = 2.0$. The corresponding density functions are shown on Figure 31.

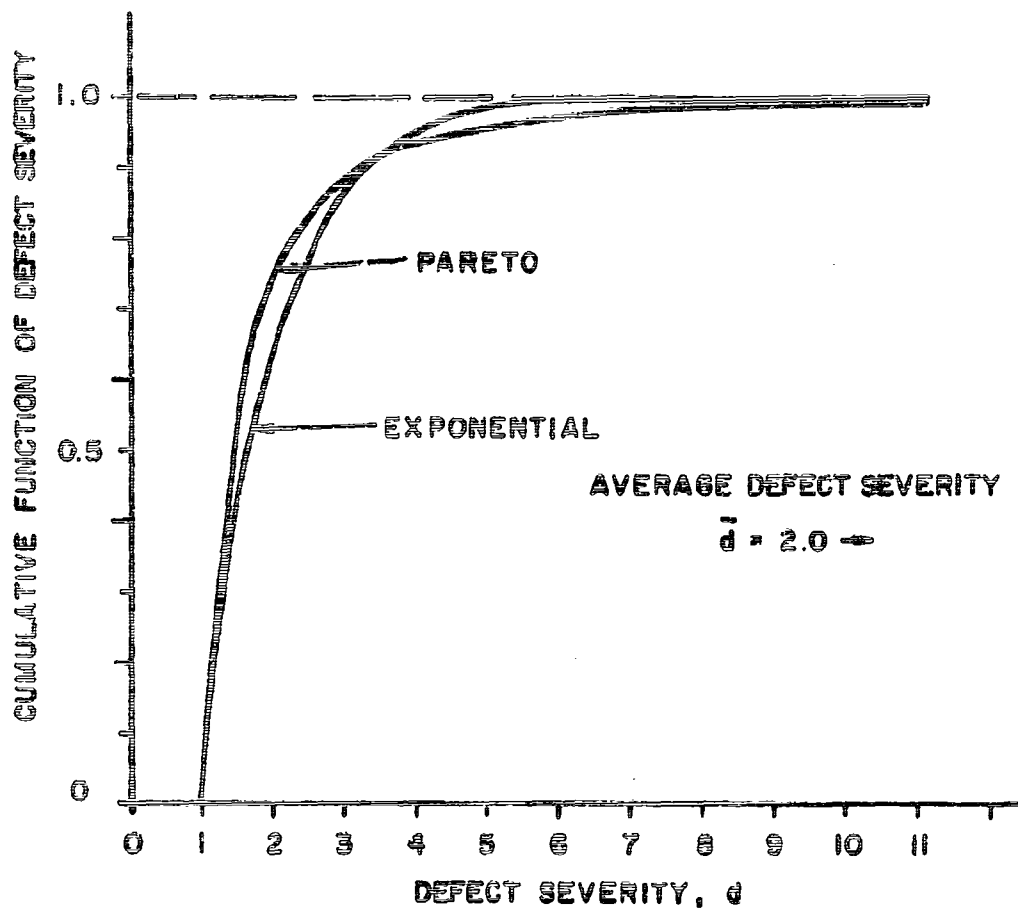


Figure 30. Cumulative Distribution Function of Defect Severity with $\bar{d} = 2.0$

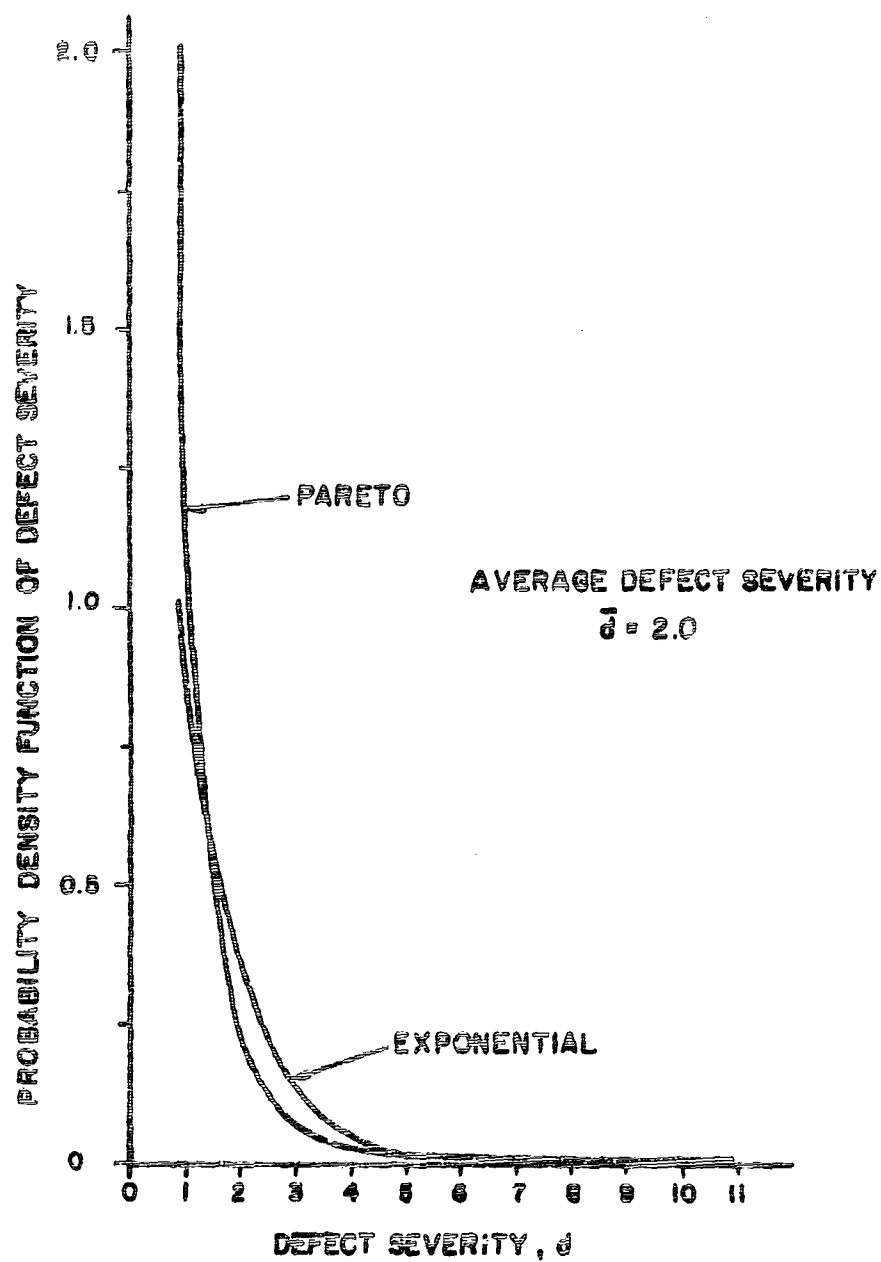


Figure 31. Probability Density Function of Defect Severity with $\bar{d} = 2.0$

1. BEHAVIOR OF Γ AS $N \rightarrow N_0$

Using the exponential distribution of Equation (A6-1) as the defect life distribution in Equation (8-1) gives

$$G(N_I) = \exp - \left\{ \frac{1}{d_0} \Gamma^{-1} \left(\frac{B}{N_I} \right) \right\} \quad (\text{A6-4})$$

Invoking the requirement of Equation (8-9) gives

$$\lim_{N \rightarrow N_0} \exp - \left\{ \frac{1}{d_0} \Gamma^{-1} \left(\frac{B}{N_I} \right) \right\} = \beta N_I^k \quad (\text{A6-5})$$

Taking logarithms,

$$\lim_{N \rightarrow N_0} \frac{1}{d_0} \Gamma^{-1} \left(\frac{B}{N_I} \right) = \log \frac{1}{\beta N_I^k} = \log \frac{1}{\beta'} \left(\frac{B}{N_I} \right)^k \quad (\text{A6-6})$$

where

$$\beta' = \beta B^k \quad (\text{A6-7})$$

Thus Γ^{-1} must behave as $\log \frac{1}{\beta'} \left(\frac{B}{N_I} \right)^k$ as $N \rightarrow N_0 \approx 0$.

Taking the inverse

$$\frac{B}{N_I} = \Gamma(d) = (\beta')^{1/kd_0} e^{d/kd_0} \quad (\text{A6-8})$$

Equation (A6-8) states that as d becomes large, $\Gamma(d)$ behaves as an exponential function. The result is required only asymptotically and cannot be true over the whole range of d since according to Equation (A6-8) $\Gamma(d)$ does not approach zero as $d \rightarrow 1$ as required by Equation (5-12) ($d = 1$ corresponds to no severity).

Turning to the Pareto distribution of Equation (A6-2) one finds, using Equation (6-8)

$$\left\{ \Gamma^{-1} \left(\frac{B}{N_I} \right) \right\}^{-c} = \beta N_I^k \quad (\text{A6-9})$$

which is satisfied if $\Gamma^{-1} \left(\frac{B}{N_I} \right)$ is given by

$$\Gamma^{-1} \left(\frac{B}{N_I} \right) = \frac{1}{\beta'} \left(\frac{B}{N_I} \right)^{+k/c}$$

where $\beta' = (B^k \beta)^{1/c}$

The function $\Gamma(d)$ in this case behaves for $d \rightarrow \infty$ as

$$\Gamma(d) = (\beta')^{c/k} d^{c/k} \quad (\text{A6-10})$$

that is, as a power function in d .

Here again, $\Gamma(d)$ does not approach zero as $d \rightarrow 1$ so that Equation (A6-10) cannot be valid over the entire range of d .

2. FATIGUE LIFE DISTRIBUTIONS WITH EXPONENTIAL AND PARETO DEFECT SEVERITY DISTRIBUTIONS

With the exponential severity distribution, the Weibull characteristic life N^* of Equation (6-8) becomes, using the expression for β given in Equation (A6-7),

$$N^* = \frac{B}{(\beta' \eta V)^{1/k}}$$

$$\text{or since } B = \frac{f_I(A_p)}{\gamma_I},$$

$$N^* = \frac{f_I(A_p)}{\gamma_I (\beta' \eta V)^{1/k}}$$

(A6-11)

With the Pareto defect severity distribution, the characteristic life becomes

$$N^* = \frac{B}{\left[(\beta')^{-\alpha} n V \right]^{1/k}} = \frac{\frac{\bar{\epsilon}_I(A_P)}{\gamma_I}}{\left[(\beta')^{-\alpha} n V \right]^{1/k}} \quad (\text{A6-12})$$

Equations (A6-11) and (A6-12) are identical relationships in form; for either of the two assumed defect severity distributions, the characteristic life varies inversely with the stressed volume and the matrix parameter γ_I .

REFERENCES

1. Lundberg, G. and Palmgren, A., Dynamic Capacity of Rolling Bearings, Acta Polytechnica No. 196 (1947)
2. Lundberg, G. and Palmgren, A., Dynamic Capacity of Roller Bearings, Acta Polytechnica No. 210 (1952)
3. Method of Evaluating Load Ratings for Ball Bearings, AFBMA Standards Section No. 9.
4. Load Ratings for Ball and Roller Bearings, ASA Standard No. B3, 11-1959.
5. ISO Recommendation 76.
6. Tallian, T.E., "Rolling Contact Failure Control Through Lubrication", Conference on Lubrication and Wear, Inst. of Mech. Engrs., London, England, (Paper No. 14), Sept. (1967).
7. Hasson, S.S., Hirschberg, M.H., "Fatigue Behavior in Strain Cycling in the Low - and Intermediate - Cycle Range" in Fatigue, an Interdisciplinary Approach, Syracuse University Press, 1964.
8. Coffin, L.F., "A Study of the Effects of Cyclic Thermal Stresses on a Ductile Metal", Trans. ASME 76, pp 931-936, Aug. 1954.
9. Morrow, J. D., "Cyclic Plastic Strain Energy and Fatigue of Metals", ASTM STP 370, Am. Soc. Testing Mats., 1965
10. Frost, H.E., and Dugdale, D.S., "The Propagation of Fatigue Cracks in Sheet Specimens", Journal of Mechanisms of Physical Solids, Vol. 6, No. 2, 1958.
11. Paris, P. C., "The Fracture Mechanisms Approach to Fatigue", Fatigue, an Interdisciplinary Approach, Syracuse University Press, 1964
12. Laird, C., "The Influence of Metallurgical Structure on the Mechanisms of Fatigue Crack Propagation", in Fatigue Crack Propagation, ASTM Spec. Tech. Publ. No. 415, p.131 (1967).
13. Wood, H.A., "Recent Observations on Fatigue in Metals", Symposium on Basic Mechanisms of Fatigue, ASTM STP 237, 1958

14. Grosskreutz, J.C., "A Critical Review of Micromechanism in Fatigue" in Fatigue, an Interdisciplinary Approach, Syracuse University Press, 1964.
15. Manson, S.S. and Mircenberg, M.H., "Crack Initiation and Propagation in Notched Fatigue Specimens", Proceedings of the First International Conference on Fracture (1965), Vol. 1, pp 479-498.
16. Peterson, R.E., "Design Approaches for Low-Cycle Fatigue Problems in Power Apparatus" in Fatigue, an Interdisciplinary Approach, Syracuse University Press, 1964.
17. Martin, J. A. and Eberhardt, A.D., "Identification of Potential Failure Nuclei in Rolling Contact Fatigue", ASME Paper No. 67-WA/CF-1.
18. Littman, W., and Widner, R., "Propagation of Contact Fatigue from Surface and Sub-surface Origins", Trans. ASME, Engr. Sect., 1966, pp. 624-636.
19. Esin, A., "The Microplastic Strain Energy Criterion Applied to Fatigue", ASME Paper No. 67-WA/Met-3.
20. Cross and Greenert, Basic Information on the Bearing Properties of Various Materials in Liquid Metals, U. S. Navy Engrg. Experiment Station Report 9C(4)966051, Nov. 1952.
21. Manson, S.S., Thermal Stress and Low-Cycle Fatigue, McGraw Hill Book Co. (1966).
22. Tetelman, A.S., McEvily, A.J., Jr., Fracture of Structural Materials, John Wiley & Sons, Inc. (1967).
23. Bridgman, P.W., Studies in Large Plastic Flow and Fracture, McGraw Hill Book Co., N. Y. (1952).
24. Forsyth, P.J.E., "A Two Stage Process of Fatigue Growth", Proc. Crack Propagation Symp. Cranfield, the College of Aeronautics, (1962).
25. Sinclair, G.M. and Dolan, T.J., "Effect of Stress Amplitude on Statistical Variability in Fatigue Life of 75S-T6 Aluminum Alloy", Trans. ASME, 75, 867 (1953).
26. Manson, S.S., "Fatigue, A Complex Subject - Some Simple Approximations", Experimental Mechanics, July, 1963.

27. Bergese, S., "An Electron Fractographic Study of Spalls Formed in Rolling Contact", ASME Paper No. 67-WA/CF-3.
28. Greenert, W.J., "The Toroid Contact Roller Test as Applied to the Study of Bearing Steels", J. of Basic Eng. ASME Trans. 1962, 84, p. 101.
29. Johnson, K.L., "Plastic Flow and Residual Stresses in Rolling and Sliding Contact", Fatigue in Rolling Contact, Inst. Mech. Engrs., Paper No. 5, 1963.
30. Poritsky, H., "Stress and Deflection of Cylindrical Bodies in Contact with Application to Contact of Gears and of Locomotive Wheels", J. App. Mech., Trans ASME 92, 192-201 (1950).
31. Smith, J.D. and Liu, C.K., "Stresses Due to Tangential and Normal Load in an Elastic Solid with Application to Some Contact Stress Problems", J. Appl. Mech. 75, 157-185 (1953).
32. Hertz, H., "On the Contact of Rigid Elastic Solids and on Hardness", Miscellaneous Papers, Macmillan and Co., London (1896) pp. 163-183.
33. Little, A.D., Inc., Final Report on A Study of Contact Fatigue, (Sponsored by ASME) Jan. 1, 1967
34. Lundberg, G. and Sjoval, H., Stress and Deformation in Elastic Contacts, Publication No. 4, Chalmers University of Technology, Gothenburg, Sweden, 1958.
35. Archard, J.F., "Elastic Deformation and the Laws of Friction", Proc. Roy. Soc. (London) A243 (1957).
36. Muskhelishvili, N.I., Some Basic Problems of the Mathematical Theory of Elasticity, P.N. Noordhoff, Ltd., p. 479 (1953).
37. Tallian, T.E., Chiu, Y.P., Huttenlocher, D.F., Kamenshine, J.A., Sibley, L.B., and Sindlinger, N.E., "Lubricant Films in Rolling Contact of Rough Surfaces", ASLE Trans. 7, 2, p. 109 (1964).
38. Eldredge, K.R. and Tabor, D., "The Mechanism of Rolling Friction, I. The Plastic Range", Proc. Roy Soc. A229 181 (1955).

39. Tallian, T.E., "Discussion to Paper entitled 'Topographs of Solid Surfaces', by J.B.P. Williamson", to appear in NASA Symposium, Interdisciplinary Approach to Friction and Wear.
40. Dwight, H., Tables of Integrals and Other Mathematical Data, Macmillan Co., New York, 3rd Ed., 1957.
41. Rice, J.R. and Brown, E.J., Discussion to paper by M. Freudenthal, entitled "Random Fatigue Failure of a Multiple-Load-Path Redundant Structure" in Fatigue, an Interdisciplinary Approach, Syracuse University Press, 1964.
42. Tallian, T.E. "Weibull Distribution of Rolling Contact Fatigue Life and Deviations Therefrom", ASLE Trans. 5, No. 1, 1962, pp. 183-196.
43. Meyer, G.J., A Mechanics Analysis of Rolling Element Failures, Ph.D. Thesis submitted to Univ. of Illinois, T & AM Report No. 132 (1960).
44. Valeri, S.E., Tallian, T.E. and Sibley, L.B., "Elastohydrodynamic Film Effects on the Load-Life Behavior of Rolling Contacts", ASME Paper 65-Lub-11.
45. Wilks, S., Mathematical Statistics, John Wiley and Sons, 1962, p. 236.
46. Epstein, S., "Elements of the Theory of Extreme Values", Technometrics 2, No. 1, 1960.
47. Gumbel, E.J., Statistics of Extremes, Columbia Univ. Press, 1958.
48. Dowson, D., and Whitaker, A.V., "A Numerical Procedure for the Elastohydrodynamic Problems of Rolling and Sliding Contact Lubricated by a Newtonian Fluid", EHD Lubrication Symposium at Leeds. Paper No. 4, Inst. of Mech. Engrs., London, (1965).
49. Morrow, J., "Fatigue Properties of Metals", Unpublished paper presented to the membership of SAE, ISTC Div. 4, November 4, 1965, University of Illinois.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)		2a. REPORT SECURITY CLASSIFICATION	
SKF Industries, Inc. Research Laboratory King of Prussia, Pa. 19406		Unclassified	
3. REPORT TITLE		2b. GROUP	
Development of a Mathematical Model for Predicting Life of Rolling Bearings			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)			
Final Report 28 Dec 66 22 Dec 67			
5. AUTHOR(S) (First name, middle initial, last name)			
Y.P. Chiu J.A. Martin J.I. McCool			
6. REPORT DATE		7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 1968		160	49
8a. CONTRACT OR GRANT NO.		9a. ORIGINATOR'S REPORT NUMBER(S)	
F30602-67-C-0147		AL68P003	
b. PROJECT NO.		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
5519		RADC-TR-68-54	
c. Task 551902			
d.			
10. DISTRIBUTION STATEMENT			
This document is subject to special export controls and each transmittal to foreign governments, foreign nationals or representatives thereto may be made only with prior approval of RADC (EMEAM), GAFB, N.Y. 13440.			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY	
		Rome Air Development Center (EMEAM) Griffiss AFB NY 13440	
13. ABSTRACT			
A description of rolling contact failure modes is given and the variables affecting the life of a rolling contact are identified. A mathematical model of subsurface and surface crack propagation is presented. The life to failure of volume elements in the vicinity of a defect is formulated. A term "severity" of a microdefect has been defined. The model is characterized by the inclusion of bulk material parameters, defect characteristics and parameters of geometry, stress, lubrication and surface topography. A statistical expression for the life of an entire rolling body is based on the defect life formula. The new model includes current standard bearing life prediction formulas as a special case. To assist in interpretation of the model, the stressed volume in a Hertzian elliptical stress field has been determined from the computed contours of equal reversing shear stress. A stress analysis has been conducted on the stresses near interacting asperities and around a surface defect (furrow).			

DD FORM 1473
1 NOV 65

Unclassified

Security Classification

Security Classification

Unclassified

Security Classification