FRICTION AND CREEP IN ROLLING CONTACT

NOVEMBER 1970

FEDERAL RAILROAD ADMINISTRATION
DEPARTMENT OF TRANSPORTATION

ERRATA
FOR "Friction and Creep in Rolling Contact"

1. Page 53, Equation (10) should read

$$\xi' = [1-(1-T')^{1/2}] \frac{4(1-\nu)}{\pi} .$$

2. Page 70, Line 4 should read

"..... becomes $p(0) \geq 2.34 \times 10^5$ psi."

3. Page 70, Line 14 should read

$$p(0) \geq 8.85 \times 10^4 \text{ psi.}$$

4. Page 80, Line 1 should read

"...elastic cylinder of radius r pressed..."

5. Page 80, Equation (46) should read

$$u = \frac{2(1-\nu^2)P}{\pi E} \left[\frac{2}{3} + \ell n \left(\frac{1.74rE}{P} \right) \right]$$

6. Page 80, Equation (48) should read

$$k_c = 1/ \frac{\partial u}{\partial P} = \frac{E/2(1-\nu^2)}{\frac{1}{\pi} \ell n \frac{1.74rE}{P} - \frac{1}{3\pi}}$$

The last two changes require a quantitative change in the discussion on p. 81 of contact resonance. The qualitative aspects of the discussion remain the same, however.

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FRA-RT-71-64		

4. Title and Subtitle	5. Report Date
Friction and Creep in Rolling Contact	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
P.R. Nayak S. Hariharan Raya Stern R. Abilock P.A. March P.K. Gupta	

9. Performing Organization Name and Address	10. Work Unit No.
Bolt, Beranek and Newman Inc. 50 Moulton Street Cambridge, Massachusetts 02138	11. Contract or Grant No.
	13. Type of Report and Period Covered

12. Sponsoring Agency Name and Address	14. Sponsoring Agency Code
Department of Transportation (FRA) Office of High Speed Ground Transportation 400 7th Street, S.W. Washington, D.C. 20591	

15. Supplementary Notes

16. Abstract

Experimental and analytical studies of friction and creep in rolling contact are reported. Factors examined for their influence on friction (adhesion) and creep are (a) surface roughness, (b) surface vibration, (c) surface contamination, (d) dynamic loading due to irregular track, and (e) rolling velocity. The following conclusions are reached. 1. Surface roughness does not influence the creep coefficients at operating loads. However, surface roughness influences the tractive capacity when the wheel and rail surfaces are either very clean or flooded with a contaminant. 2. Surface vibrations affect wheel-rail friction considerably. 3. Surface contamination decreases both friction and creep coefficients. The magnitude of the change in these coefficients depends on the oil viscosity temperature and pressure coefficients, the normal load on the wheel and the surface roughness. 4. Dynamic loads due to suspension resonances do not appear to influence the friction or creep coefficients significantly. 5. Observed decreses in the friction coefficient at increased rolling velocities are probably due to increased surface vibrations, decreased time for the formation of friction junctions, and elastohydrodynamic effects.

17. Key Words	18. Distribution Statement
Rolling contact; Friction; Creep; Vibration; surface Dynamic loading	Availability is unlimited. Reports may be purchased from the National Technical Information Service, Springfield, Virginia 22151

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price

Form DOT F 1700.7 (8-69)

REPORT SUMMARY

The two main themes of this work are friction and creep in rolling contact. The motivation for the research reported here is that a detailed knowledge of wheel-rail contact phenomena is essential for the design of the trains of tomorrow. A knowledge and understanding of the factors governing rolling friction under field conditions is essential for the development of high-capacity trunk lines, with their requirements of high acceleration and deceleration rates. A similar understanding of the factors governing creep in rolling contact is needed for the design of vehicles that will run stably at high speeds.

Factors examined for their influence on the rolling friction coefficient and on creep are (a) surface roughness, (b) surface contamination, (c) surface vibration, (d) rolling velocity and (e) dynamic loads due to track irregularities. The following conclusions are reached.

1. Surface roughness influences the rolling friction coefficient, rough surfaces generally having a higher coefficient than smooth surfaces. This is particularly true either when the surfaces are very clean (as with a plasma torch) or highly contaminated. Surface roughness has a smaller influence on boundary-lubricated surfaces.

2. The slightest amounts of surface contamination decrease friction coefficients by as much as 1/2. For contaminated surfaces, the effective tractive capacity of the contact (i.e., the tractive force that may be transmitted without excessive slip) depends on the rolling velocity. Such characteristics of the lubricant as its viscosity pressure and temperature coefficients, its compressibility and its tendency to react with the metal surface are thought to be important factors governing friction.

3. Surface vibrations may strongly influence friction in rolling contact. Large vibrations take place at suspension resonance frequencies and at the contact resonance frequency, and may markedly change the behavior of the friction junctions in such a way as to decrease friction. It is likely that the influence of surface vibrations on rough surfaces is smaller than their influence on smooth surfaces.

4. The rolling friction coefficient decreases as the rolling velocity increases. Two mechanisms are suggested: (a) the time spent by a friction junction in the contact region decreases, causing a decrease in the shear strength of the junction, and (b) increased surface vibrations cause a decrease in the shear strength of the junctions.

5. The effect of dynamic loads due to irregular track in decreasing the effective normal load on a wheel, and consequently its tractive capacity, is found to be small. If further investigations of rolling friction should show the effects to be not quite negligible, then low-frequency dynamic loads, at the car resonance frequency, are more important then high-frequency dynamic loads at the wheelset bounce frequency. This would suggest that greater emphasis be placed on the design of the secondary suspension from the viewpoint of tractive capacity.

6. Surface roughness has no effect on creep coefficients for dry contact at usual operating loads, if surface vibrations are not present.

7. Fluctuating normal and tractive forces do not alter creep coefficients if the period of fluctuation is large compared to the time to roll one contact length.

8. Contaminants such as water and oil decrease creep coefficients by amounts ranging from 1/4 to more than 2/3, depending on the normal load and the rolling velocity.

9. When suspension resonances and the contact resonance are excited, creep coefficients may alter drastically. The behavior would be expected to depend on the surface roughness.

10. Increasing rolling velocity does not alter the creep coefficients, unless at some rolling velocity some resonance is suddenly excited, an unlikely event with random track with a broad-band power spectral density.

In the foregoing, the "creep coefficient" is defined as the slope of the linear part of the creep-tractive force relationship, a quantity of prime interest in the linear theory of vehicle stability.

Studies are also made of the dynamic loading of random track by a high-speed vehicle, and consequent damage to the rail. It is found that regions of localized plastic flow, called "damaged zones", occur, and that their frequency of occurrence cannot safely be extrapolated to high speeds from data obtained at low speeds.

A refined version of the theory of the rolling contact* of rough surfaces is presented, along with numerical techniques for assessing the importance of surface roughness in any given application. The theory is well suited to the analysis of situations where surface roughness may be important, an example being the rolling contact of highly clean (and therefore highly adhesive) surfaces.

* See Ref. 21 of Chapter 2 on p. 85

Finally, a mathematical theory of the description and analysis of surface roughness is presented which removes many obstacles to the study of the importance of surface topography in a variety of interfacial phenomena.

FOREWORD

The theme that runs through the following chapters is that of friction in rolling contact, particularly as applied to railway wheels. And it is this theme that binds together seemingly disparate chapters. Some deal with rolling friction under laboratory conditions, and others investigate rolling contact under field conditions.

Chapter 1 is a survey of some of the most important literature. It concerns itself in the main with fundamental friction mechanisms, with a view to gaining insight into reasons for the discrepancies between data obtained in the laboratory and those obtained in the field.

Chapter 2 describes in detail an analytical and experimental investigation of creep in rolling contact under laboratory conditions. We are mainly concerned here with evaluating the importance of surface roughness as a parameter in rolling contact. Creep under dynamic conditions is also explored briefly, as is a phenomenon known as contact resonance, which promises to be of importance.

Fundamental friction mechanisms are the object of the investigations reported in Chapter 3 — a study of the deformation of macroscopic models of friction junctions under a variety of conditions.

Chapter 4 looks at rolling contact under field conditions, with the important element of the random nature of track geometry now included. Analytical techniques for obtaining the dynamic loads between wheel and rail are described first. Numerical results for a specific simple vehicle are then used in detailed investigations of both damage to the rail and of the tractive capacity of wheelsets.

Chapter 5 closes the circle by once again returning to tools necessary for fundamental investigations of interfacial phenomena, particularly friction. A detailed theory of the characterization of rough surfaces is presented, including techniques for obtaining characteristic indices for any surface.

Chapter 6 contains a review of experimental work on rolling contact performed at MIT. Finally, Chapter 7 presents recommendations for future work that we believe will further enhance knowledge of some of the central problems of wheel-rail interaction.

By way of acknowledgments, our debt is the greatest to Mr. Kenneth Lawson and Mr. Steven Ditmeyer of OHSGT, U.S. Department of Transportation, both for the encouragement and the free hand they gave us during the course of the entire investigation.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

1. A SURVEY OF ROLLING CONTACT LITERATURE

1.1 Introduction

In Sec. 1.1, we briefly describe the organization of this review. The review assumes a general knowledge of the theory of rolling contact on the part of the reader, and is intended only as a guide to the related literature; it is by no means comprehensive. In several places, we resort to "brainstorming" to explain experimental data; these are guesses and must be treated as such; their correctness can be proved only by further experiments. Their usefulness lies, obviously, in indicating what experiments are desirable.

Section 1.2 contains a brief review of the smooth-surface theory of rolling contact, together with its most important results. Section 1.3 contains a critique of this theory and an examination of the reasons why its predictions do not always match experimental results. Section 1.4 contains a survey of the most popular theories of friction, and an examination of such variables as surface roughness, contamination, and vibration on the friction coefficient. Section 1.5 reviews the mathematical theory of contact of rough surfaces. Section 1.6 reviews theory and experiment in the mechanics of deformation of friction junctions formed during contact of rough surfaces. Section 1.7 examines the problem of rolling contact of railway wheels and rails, with the attendant problems of surface vibration, contamination, and load variations. Section 1.8 describes some of the more significant experiments that have been made in the field of rolling contact.

1.2 Review of the Smooth-Surface Theory of Rolling Contact

When two bodies are pressed together, they form a contact region. When one of the bodies rolls on the other, the shape of the contact region and the normal pressure distribution within it remain unchanged in time if the following conditions are satisfied [*96*].

1. the normal load remains constant, and
2. the geometry of the bodies (i.e., their radii of curvature) near the contact region remains unchanged.

If a tangential load is applied to the rolling body, it will transmit this force to the other body through the contact region, by means of shear stresses in the contact region. These shear stresses are caused by the friction between the two bodies.

The critical assumptions of the smooth-surface theory are:

1. both bodies have topographically smooth surfaces, and
2. friction between the bodies is governed by Coulomb's law (see Secs. 1.3 and 1.4).

The theory suggests that the contact region may be separated into a locked region, in which there is no interfacial slip, and a slipped region, in which interfacial slip occurs. The essential problem is to predict the extent of the locked and slipped regions and the stresses in them, given the externally applied normal and tangential loads (see Fig. 1).

As a specific example, consider a wheel rolling on a track. The loads applied to it may be: (i) a normal load N, (ii) a load T, parallel to the track, (iii) a load T_1 parallel to the wheel axle, (iv) a moment M_a about the axle and a moment M_s about a line perpendicular to both the track and the axle.

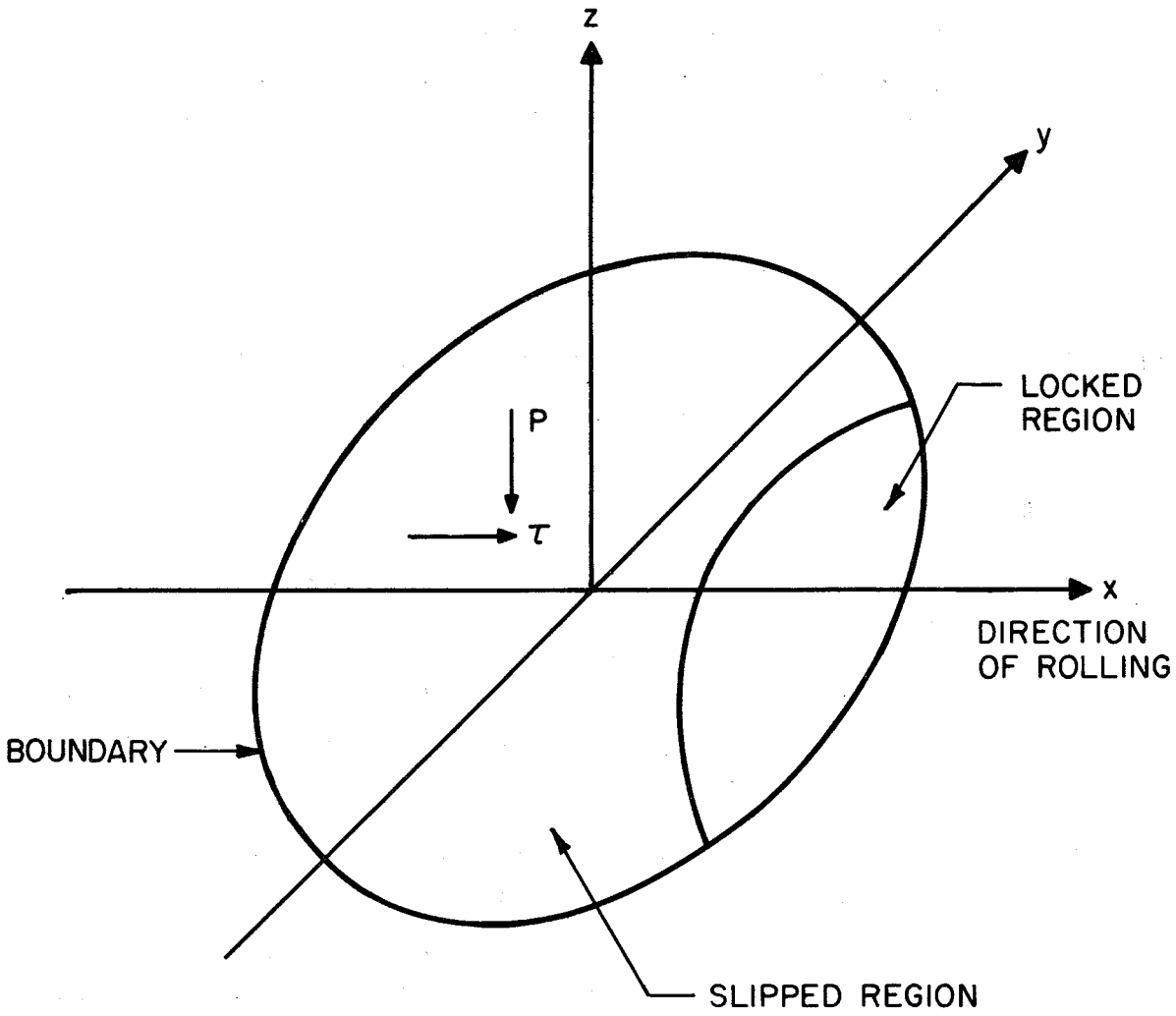


FIG. 1.1 LOCKED AND SLIPPED REGIONS WITH NORMAL PRESSURE P AND SHEAR STRESS τ TO BE DETERMINED AS A FUNCTION OF THE NORMAL LOAD N AND THE TANGENTIAL LOAD T

Let the wheel radius be R, its forward velocity V, and its angular velocity about its axle Ω. Also, let the wheel spin parallel to the $\overline{M}_s$ vector with "spin" angular velocity ω_s.

The forward creep is defined as:

$$\xi_x = \left|\frac{\Omega R - V}{V}\right| = \left|\frac{\Omega R}{V} - 1\right| .$$

If the wheel has a transverse velocity (parallel to the axle) V_y, the transverse creep is defined as:

$$\xi_y = \frac{V_y}{V} .$$

The relationships of interest to railway engineers are those between N, T_1, T_2, M_a, and ξ_x, ξ_y, and ω_s.

Some of these relationships as predicted by the smooth surface theory for steady rolling are shown in Figs. 1.2, 1.3, and 1.4 [*22*].

The earliest analytical work in rolling contact seems to have been that of Carter [*70*]. Poritsky's work [*88*] bore out Carter's; both dealt with two-dimensional problems, the rolling of cylinders on plane surfaces. The major steps forward in the theory of three-dimensional problems may be attributed to K.L. Johnson [*76-81*] and J.J. Kalker [*82-84*]. Johnson suggested what may be called "engineering solutions", whereas Kalker's solutions may be termed "exact". For other information in rolling contact theory, see Refs. 69-89.

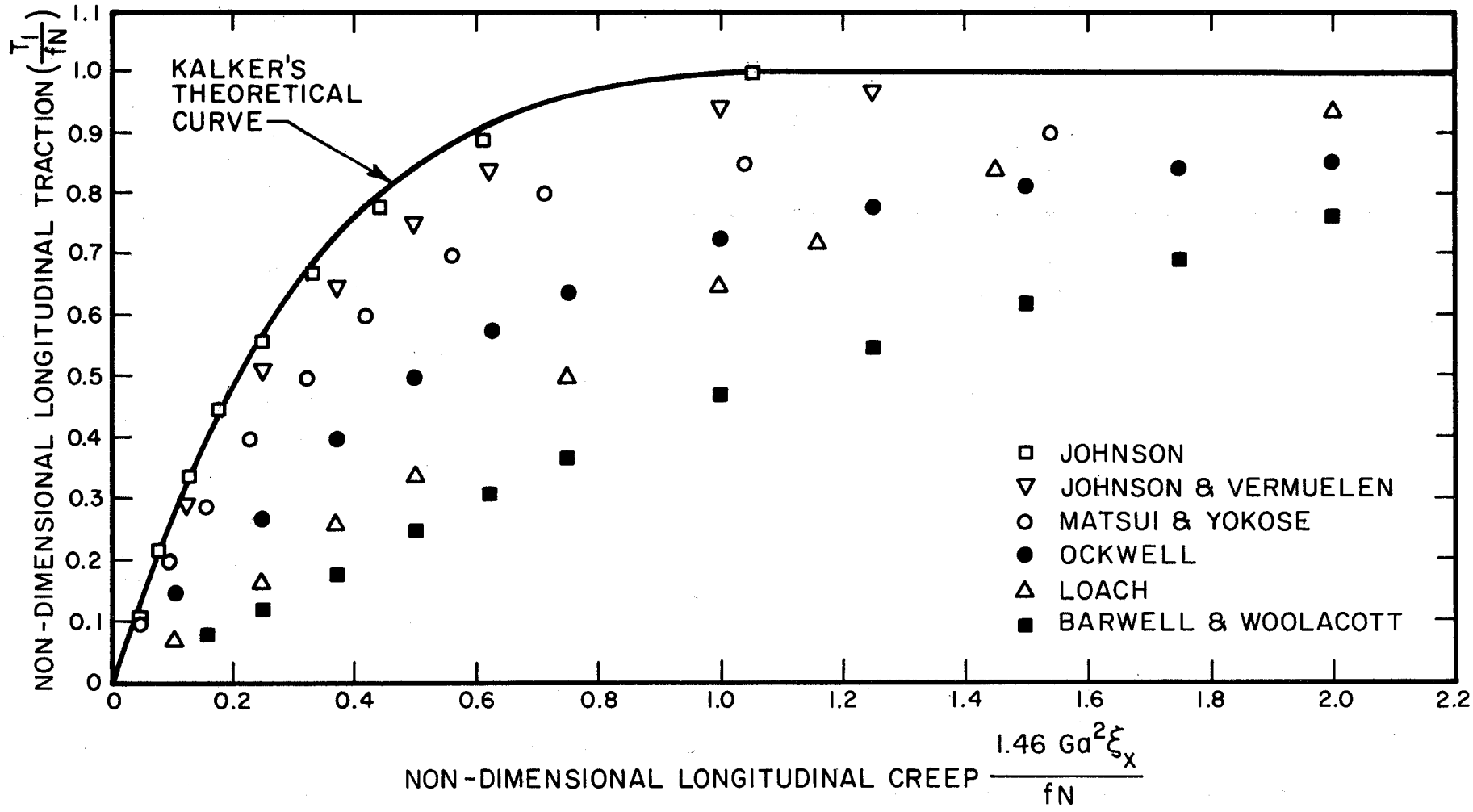


FIG. 1.2 COMPARISON OF LONGITUDINAL CREEP MEASUREMENTS WITH KALKER'S THEORETICAL CURVE

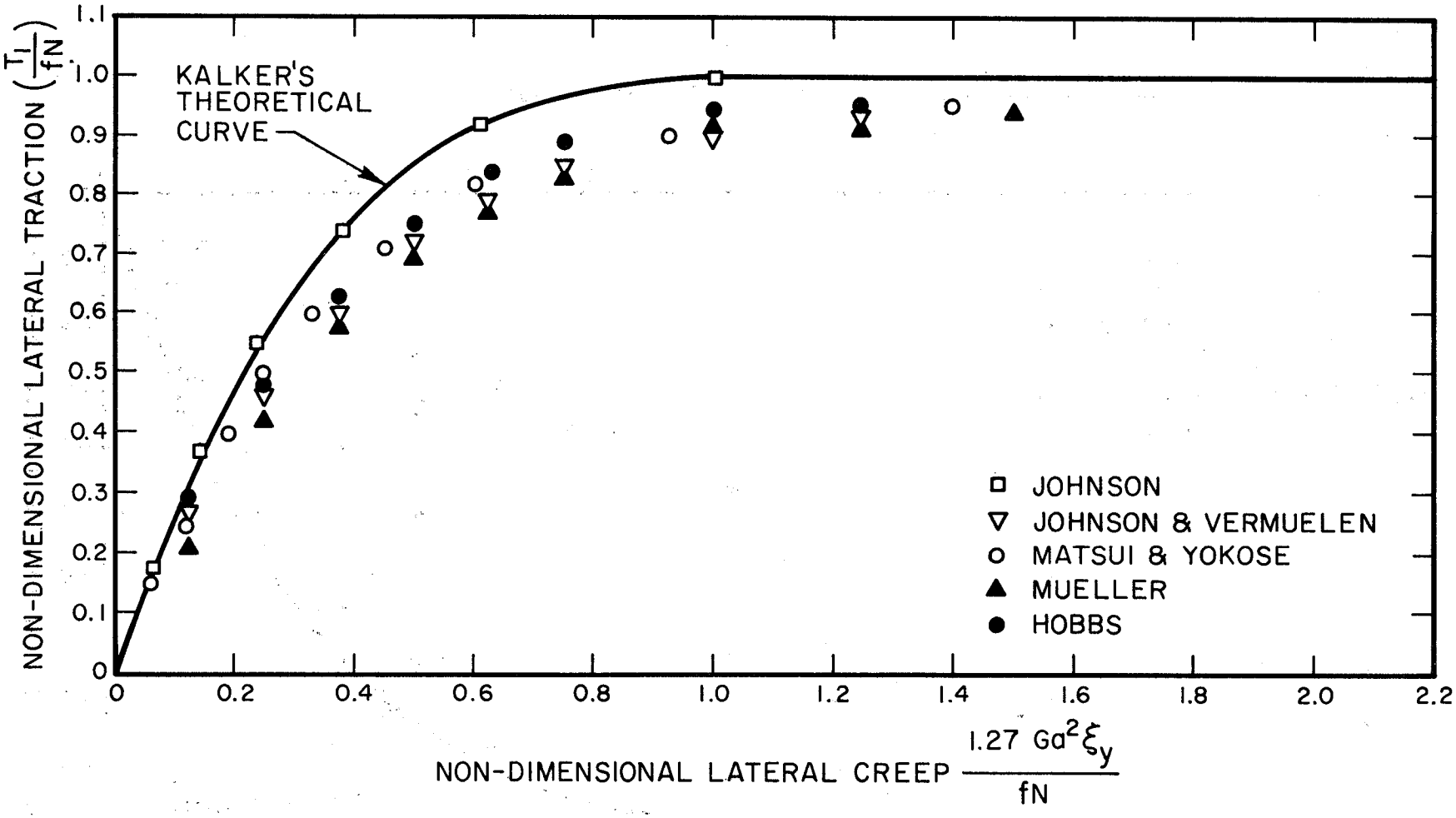


FIG. 1.3 COMPARISON OF LATERAL CREEP MEASUREMENTS WITH KALKER'S THEORETICAL CURVE

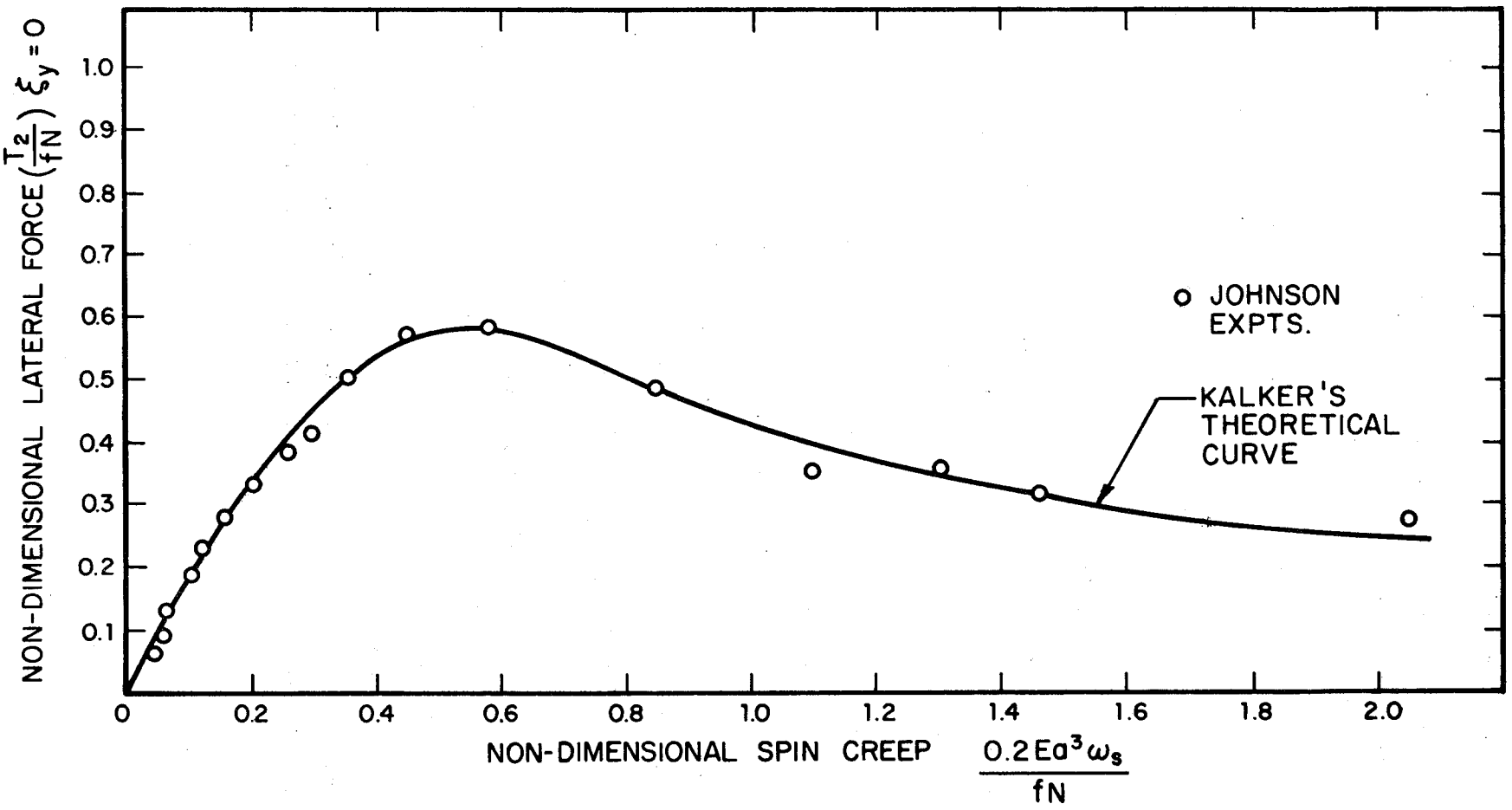


FIG. 1.4 COMPARISON OF SPIN CREEP VS LATERAL FORCE MEASUREMENTS BY JOHNSON WITH KALKER'S THEORETICAL CURVE

1.3 Critique of the Smooth-Surface Theory

A critical evaluation of the smooth-surface theory of rolling contact seems to indicate that it is valid under the following conditions:

1. reasonably heavy loads, where the elastic surface displacements are much greater than the rms roughness of the surface,
2. low rolling speed, and
3. slow fluctuations of loads in time.

These conditions are imposed as much by experimental evidence as by intuition. A factor that might influence rolling contact is surface roughness. This possibility is examined in some detail in Secs. 1.5 and 1.6.

Experimental investigations of rolling contact do not always produce results that agree with theory, as is shown in Figs. 1.2-1.4 [22]. Reasons suggested for this disagreement are:

1. unevenness of the rolling surfaces, producing fluctuating loads, and perhaps vibration of the rolling surfaces,
2. contamination of the surfaces, calling into doubt the validity of the assumption of Coulomb friction in the interface,
3. careless experimentation, in which parameters such as the spin ω_s are present but unobserved, all effects being ascribed to the creeps ξ_x and ξ_y.

To obtain an idea of the discrepancy between theory and experiment, we may observe from Fig. 1.2 that longitudinal creep has been measured that is more than four times the value predicted by theory.

The major criticism that may be levelled at the smooth-surface theory is its assumption that Coulomb's law of friction is valid in the contact region. Coulomb's law as used in rolling contact is:

$$\tau = fp \frac{\bar{V}_s}{|V_s|} \text{ if } |V_s| \neq 0$$

$$|\tau| \leq fp \text{ if } |V_s| = 0 \ ,$$

where τ is the shear stress, f is a friction coefficient, p is the normal pressure, and V_s is the slip velocity. Coulomb's law of friction is examined in some detail in Sec. 4, but here we may observe that it usually applies under the following conditions:

1. gross sliding,
2. reasonably heavy normal loads, and
3. slowly fluctuating loads.

Even under these conditions, Coulomb's law applies to the tangential and normal *loads*, not the stresses in the interface. The extension of the law to friction stresses in rolling contact is a pure hypothesis, which can only be verified by experimental testing of some deduced consequence.

A theory of rolling contact postulating surface roughness effects to be important is presented in Chap. 2. This theory drops both the assumption that the surfaces are smooth as well as the assumption that Coulomb's law of friction is valid in the contact region.

1.4 Friction

1.4.1 General

When a surface slides on another, energy is dissipated at the interface by shear stresses parallel to the interface, known as friction stresses. Detailed investigations of friction can be found in Refs. 90-97; here, we present only a brief compilation of relevant information.

The following mechanisms have been postulated for the generation of shear stresses in the interface:

a) Mechanical interlocking of the roughness peaks (asperities) on the two surfaces. This mechanism is now widely believed to be incorrect.

b) Digging into and ploughing up of a soft metal by the asperities on a harder one.

c) Electrostatic forces due to differences in potential which act across the interface. It is doubtful that such a mechanism could operate for any but the cleanest metal surfaces; even then, the electrostatic contribution to the total friction force is negligible.

d) Formation of a metallic bond and its subsequent shear.

e) Secondary bonds such as Van der Waal bonds, for materials without an affinity for each other.

f) Diffusion of atoms from one body into the other, chiefly for organic materials.

g) Viscous shear of fluid films between the surfaces. It is our belief that in the contact of metals, mechanisms b), d) and g) are the ones that generate friction forces.

Friction is also classified according to the nature of the contaminant layer between the surfaces (if any):

a) Dry friction, with no fluid present between the surfaces. Oxides of the bulk material may be present, however.

b) Boundary friction in which a monolayer of fluid separates the surfaces. Metal-to-metal contact may occur by breaking of the layer. Sometimes the layer is absorbed onto the surface.

c) Hydrodynamic friction where a multilayer (many molecular thicknesses) is present. Fluid dynamic effects are then dominant.

d) Solid-film friction, when a (possibly) multilayer thickness of a soft solid separates the two surfaces. (For example, molybdenum disulfide, MoS_2, between metals.)

1.4.2 Coulomb's law

For a brief, critical evaluation of Coulomb's law, see Ref. 97. Coulomb's law is a rough-and-ready rule for prediction of frictional behavior during sliding contact.

If N = normal load,

$\overline{T}$ = tangential load,

$\overline{V}$ = sliding velocity,

then Coulomb's law states that

$$\overline{T} = \frac{\overline{V}}{|\overline{V}|} f_k N , \qquad |\overline{V}| \neq 0 ,$$

$$|\overline{T}| \leq f_s N , \qquad |\overline{V}| = 0 ,$$

where f_k is the kinetic coefficient of friction and f_s is the static coefficient of friction. In the law as it is usually stated, f_s and f_k are assumed independent of surface roughness, velocity of sliding, normal load, and area of contact, for a given pair of materials with given surface contamination. The coefficients f_s and f_k are assumed to depend on

a) the nature of the materials in contact,

b) the temperature, and

c) the nature of the contamination.

It is also assumed — and this was until recently considered to be experimentally verified — that $f_s > f_k$.

The physical explanation usually given for the nondependence of the friction force on the normal load and apparent area of contact during sliding is as follows. When two rough surfaces are made to touch over an "apparent area of contact" A_a, real contact occurs at or near the tips of asperities, giving rise to a "real contact area" A_r. Assuming that the normal load N is sufficiently large, the tips will surely undergo plastic deformation (so the argument goes); the normal stress on each little patch of real contact must be h, the penetration hardness. We may then write:

$$h \cdot A_r = N, \quad \text{or} \quad A_r = N/h \ .$$

Thus, A_r is proportional to N. If one body is made to slide on the other, the patches of real contact must be sheared plastically, and the shear stress must be $k\tau_y$, where τ_y is the yield

strength in shear, and k is a factor accounting for contamination, oxidation of the surface, weakness of the metallic bond, etc. The tangential force is then

$$T = k\tau_y A_r = \frac{k\tau_y}{h} \cdot N \ .$$

For many metals, $\tau_y \approx 0.3\ \sigma_y \approx 0.1\ h$, where σ_y is the yield stress in compression.

$$T = 0.1\ kN = fN,$$

where

$$f \equiv 0.1\ k \ .$$

So runs the basic argument. The current argument is actually somewhat more complicated [*90*], taking into account the growth of the junctions during initial sliding, etc.

The main objection to the preceding explanation of Coulomb's law of friction is that it seems unlikely that the asperities of a surface will continue to deform plastically after a reasonable amount of running in, unless there is a considerable amount of wear taking place. We believe that when wear is insignificant, the asperity deformations will be predominantly elastic after the first few loadings. A little further on, we shall discuss an alternative explanation of Coulomb's law (Sec. 1.4.7 and 1.6).

It was stated above that it is generally assumed that $f_s > f_k$. Recent measurements [*30,97*] have shown however, that the friction coefficient is a continuous function of the sliding velocity. In many cases, f decreases with velocity, but in some cases it may

increase. Moreover, Tolstoi [*30*] has shown that the friction coefficient during sliding depends to a large extent on the stiffness and damping of the contacting bodies. Where damping is not very large, however, and for medium velocities [1 cm/sec 10^3 cm/sec], it seems safe to assume that f_k is a constant, and that $f_k < f_s$, where f_s is defined as

$$f_s = \lim_{V_s \to 0} f_k \ ,$$

and where V_s is the sliding velocity.

1.4.3 Coulomb's law in rolling contact

When Coulomb's law is applied to rolling contact, it talks about stresses in the contact region rather than gross loads (see Sec. 1.3). If at any point in the contact region there is a nonzero relative (or slip) velocity $\overline{V}_s$, the tangential stress is given by

$$\overline{\tau} = fp \frac{\overline{V}_s}{V_s} \ , \qquad |\overline{V}_s| \neq 0 \ .$$

If $|V_s| = 0$, the requirement on $\overline{\tau}$ is

$$|\tau| \leq fp \ , \qquad |\overline{V}_s| = 0 \ .$$

Note that f is taken to be the same regardless of whether there is or is not slip; in other words, the static and kinetic friction coefficients are taken to be the same. To our knowledge, there is very little discussion of the applicability of Coulomb's law to rolling friction analyses, much less a discussion of the differences between f_s and f_k.

We offer a brief justification for assuming $f_s=f_k$ in rolling contact; further observations on the nature of the friction law are offered in Secs. 1.4, 1.5 and 1.6.

In measurements of the static coefficient of friction, it is found [*97*] that the coefficient is strongly dependent on the time that the surfaces have been in contact before the measurement is made. This may be related to (a) creep effects, making the asperities of a surface flow into the valleys of the other surface, and (b) diffusion effects, determining the strength of the bond between various asperities. It is found that as the time of contact decreases, the measured friction coefficient decreases. In rolling contact, the time of contact in the regions of no slip is vanishingly small; moreover, vibrations of the surfaces are present, whereas they are absent in a static friction measurement. Both these factors tend to bring down the static friction coefficient (or the "no slip" friction coefficient) [*30,97*], possibly to the value of the kinetic friction coefficient.

1.4.4 Anamolous behavior of the sliding and rolling friction coefficients

Consider the following two ways of measuring a friction coefficient.

1. A cylinder of radius R is pressed onto a plane surface with a normal load N. The axis of the cylinder cannot move. A torque M is then applied so that the cylinder slips relative to the plate at a constant velocity. For equilibrium, we must have

$$T = M/R \ ,$$

where T is the friction force between the cylinder and the plane surface. The friction coefficient may then be calculated from

$$f_1 = T/N = M/RN \ .$$

2. The same cylinder can now roll along the plate and slip simultaneously. Let its (constant) linear and angular velocities by V and Ω. Then the slip velocity is

$$V_s = \Omega R - V \ .$$

The tangential force is again T=M/R. It is found that M reaches a limiting maximum value as V_s/V is made to increase. The friction coefficient may then be defined as

$$f_2 = T_{max}/N = M_{max}/RN \ .$$

It is found that $f_1 \neq f_2$, and that f_1 may be as much as 50% larger than f_2.

Thus there is some doubt as to what friction coefficient is to be used in comparing rolling contact experiment with theory. What is usually done is to use the friction coefficient f_2, so that the dimensionless tangential force T/f_2N has a maximum value of 1.

We can only speculate on the reason for the difference: in the first case, contact occurs at a fixed place on the plate; as a consequence, all contaminating layers are soon worn away,

and good adhesion is possible between the two metals. In the second case, new metallic surface is coming into contact, and oxides and contaminating layers are likely to reduce the friction.

1.4.5 Effects of cleanliness, contamination

In most laboratory experiments, surfaces are "cleaned" with some solvent such as Trichlorethylene or Carbon Tetrachloride. It has been shown, however, [*97*], that such surfaces are not really clean, the limit of cleanliness being imposed both by the oxidation of the surface and the purity of the solvent itself. It is questionable, moreover, whether surfaces cleaned in such a manner deserve investigation, as results from such an investigation may be of little use in practical problems. However, the study of thoroughly clean surfaces, such as those obtained with a plasma torch; is extremely important, as stressed in Chap. 2.

In most experiments on rolling contact, one can be reasonably certain that the surfaces have a film of oxide on them. For most metals, the oxide has a smaller tendency to form a strong adhesive bond with itself than does the metal; consequently, the presence of an oxide usually reduces the friction coefficient.

The effect of dust [*90*] is to increase the coefficient of friction. The effect of a thin layer of oil (a monolayer) is to reduce the coefficient of friction. The effect of a thick layer of oil at low speeds is to increase the coefficient of friction. The effect of a thick layer at high speeds is very complex, because of elastohydrodynamic effects [*93*].

Rabinowicz [*97*] argues fairly convincingly that the friction coefficient must depend strongly on the quantity

$$G \equiv 1 - 2W_{ab} \cot\theta/vp \ ,$$

where

θ = average slope of rough surface

v = average junction radius

p = penetration hardness $\approx 3\sigma_y$

σ_y = yield stress in compression,

$W_{ab} = \gamma_a + \gamma_b - \gamma_{ab}$

γ_a = surface energy of body a (ergs/cm^2)

γ_b = surface energy of body b

γ_{ab} = interfacial surface energy for bodies a and b.

As a guide, Rabinowicz [*97*] observes that if

a and b are identical, $W_{ab} \approx 2\gamma_a$

a and b are "compatible", $W_{ab} \approx \frac{3}{4} (\gamma_a+\gamma_b)$

a and b are incompatible, $W_{ab} \approx \frac{1}{2} (\gamma_a+\gamma_b)$

If G is small compared to 1, the friction coefficient f is large; if G is near 1, f is small. γ/p for a wide range of materials is shown in Fig. 1.5 [*97*].

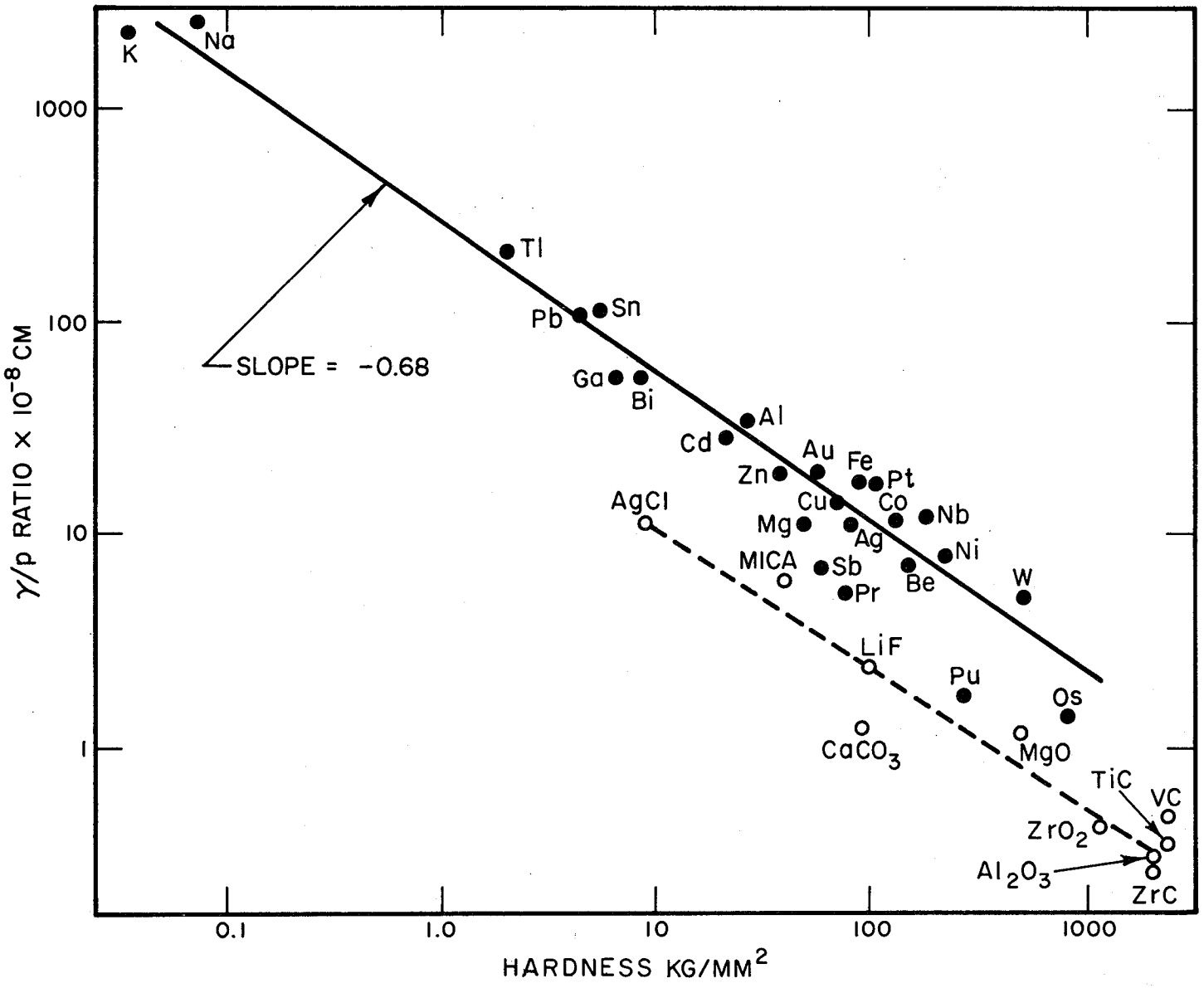


FIG. 1.5 γ/P RATIO AGAINST HARDNESS FOR VARIOUS MATERIALS

1.4.6 Effects of surface vibration

There have been a number of investigations of the influence of surface vibration on sliding friction. Of these we mention but two; Tolstoi investigated normal vibrations of the surface [*30*]; Mitskevich investigated tangential vibrations of the surface [*27*]. The vibration amplitudes in each case were quite small, of the order of the rms roughness of the surfaces.

Mitskevich shows that the effect of tangential surface vibration is to change the friction coefficient as follows:

$$f_{vib} = f \frac{2}{\pi} \frac{V}{\omega A}$$

where V = velocity of sliding

A = amplitude of vibration

ω = angular frequency of vibration

The analysis is rather simplistic, accounting only for the fact that sometimes the fixed (but vibrating surface) is moving with the sliding body and sometimes against it; however, his experimental results agree with his theoretical predictions.

Tolstoi's analysis and experiment were considerably more sophisticated. The following is a brief summary of his results:

a) Sliding motion is invariably associated with tangential and normal vibrations, caused by the sliding of the asperities of one surface past the asperities of the other.

b) If these vibrations are forcibly damped, the friction force increases.

c) If the contacting bodies are made to vibrate, the friction force is reduced, the maximum reduction being at a resonance of one of the bodies.

d) The variation of f with sliding velocity is continuous, no sudden drop being found from the static f to a kinetic f. The slope of the f - V_s curve, moreover, can be controlled by suitably controlling the stiffness and damping of the contacting elements.

We might add in conclusion that Tolstoi's method for damping normal vibrations is fairly complex, and would by no means be easy to use in most cases of "engineering" contact. However, the imposition of vibration to reduce friction seems a technique worth investigating in rolling contact.

Tolstoi accounts for the effects of normal vibrations by showing that the mean separation (averaged over time) between two surfaces increases as the amplitude of vibration increases. This means less real contact area between the two surfaces, and consequently less shear resistance.

Other effects that might enter are:

a) Dynamic elastic and plastic properties are relevant, rather than static properties.

b) Time effects, if the time to form a strong metallic junction is comparable to the period of vibration.

1.4.7 Effects of surface roughness

The effect of surface roughness on the sliding friction of copper on copper is shown in Fig. 1.6 [*97*]. In the regions of very small roughness, the real area of contact apparently

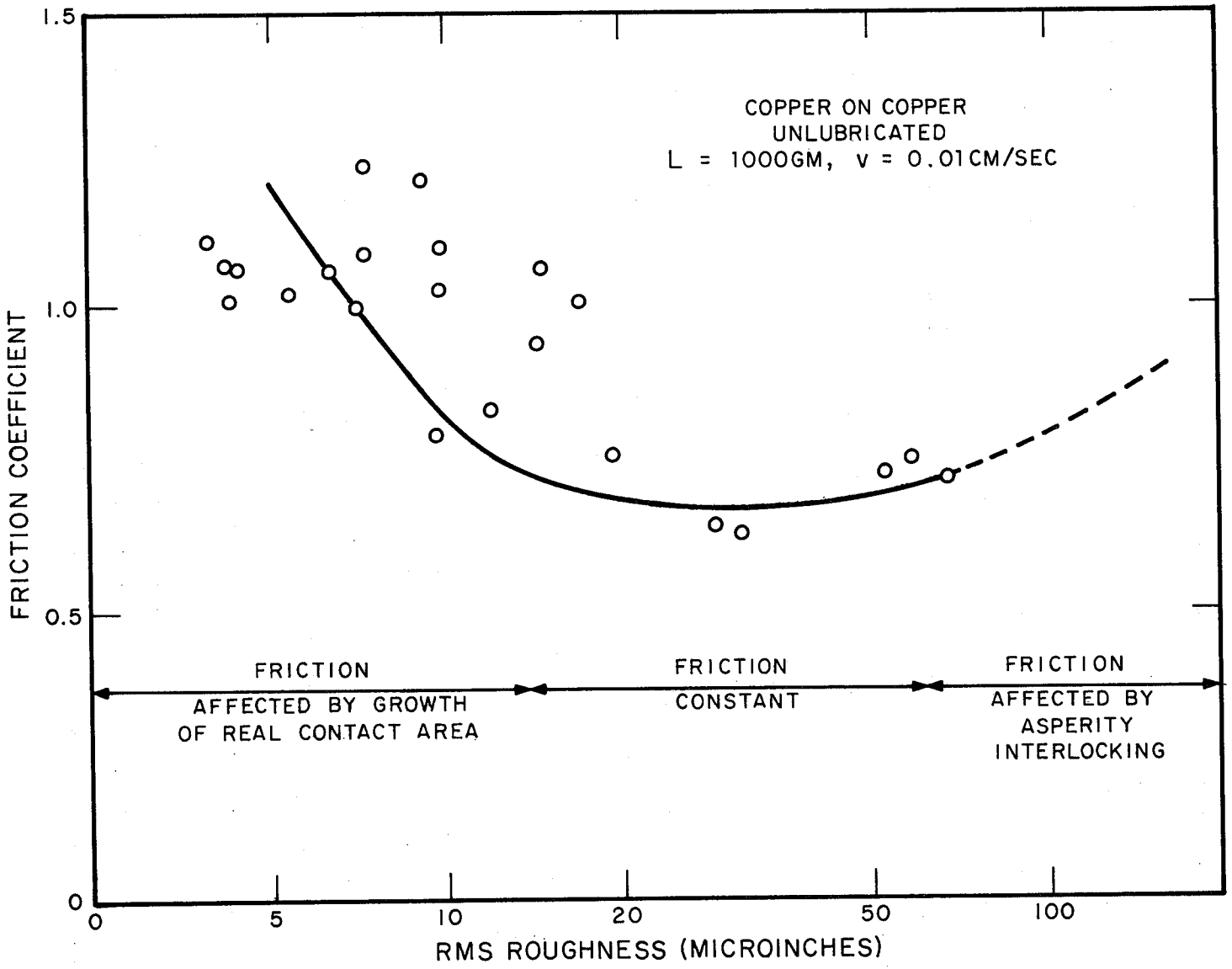


FIG. 1.6 FRICTION COEFFICIENT PLOTTED AGAINST RMS ROUGHNESS FOR UNLUBRICATED COPPER ON COPPER

becomes very large through junction growth [*90*]. In the regions of very great surface roughness, friction becomes a mechanical phenomenon, due to the interlocking of asperities. In the middle range of roughness, the analyses of Greenwood *et al* [*37,38*] indicate a porportionality between real area of contact and normal load, regardless of whether the asperity deformation is elastic or plastic. A simple extension of the analysis of Sec. 1.4.2 then provides a "proof" of Coulomb's law of friction, with the friction coefficient being independent of surface roughness. The essential step is the assumption that the shear strength of a junction is proportional to its area.

1.5 Static Contact of Rough Surfaces

1.5.1 Introduction

When contacting surfaces are modeled as rough, it is necessary to specify the following:

1. the nominal smooth surface geometry,
2. the waviness, or the departure from the nominal surface with wavelength greater than some minimum λ_0,
3. the roughness, or the departure from the nominal surface with wavelength less than some maximum λ_0.

Both the waviness and the roughness are specified as random functions. For example, we might specify the probability density function

$$p(h) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{h-\bar{h}}{\sigma}\right)^2} ,$$

which is the Gaussian density function. In this equation,

h = departure of any point from the nominal surface.

$\bar{h}$ = mean departure, usually zero.

σ = standard deviation of h.

Then $p(h_0)dh$ is the probability that an arbitrary point will lie at a height H away from the nominal surface such that $h_0+dh \geq H \geq h_0$.

Obviously, it is possible to "specify" any number of probability density functions, unless one has regard for their relevance to actual machined surfaces. No comprehensive study of this sort has been undertaken. What little information is available, however, supports the claim that polished, lapped or shot-peened surfaces are Gaussian.

When considering contact of two rough surfaces, the above description is inadequate. Contact usually takes place at or near the tips of asperities, and what is of greatest interest in the description of a rough surface is:

a) the density of asperities, no./in.^{2}

b) the height distribution of the asperities

c) the tip curvature of the asperities, and its distribution.

We shall consider some ways in which these three parameters are determined or specified.

1.5.2 Rough surface statistics

In a "profilometric" the analysis of the rough surfaces, a profilometer is used to obtain random profiles of the two

contacting surfaces. These profiles are then superposed, to simulate actual contact. Wherever the profiles overlap, contact is assumed to take place. It is also usually assumed that deformation is plastic. By this means, it is possible to obtain the following parameters as a function of the distance between the nominal surfaces:

a) The length of real contact per unit length of the profile,

b) the average length of a segment of real contact, and

c) the spacing between these segments of real contact.

If

L = length of profile

ℓ = length of real contact

d = average length of a segment of real contact, then it is possible to show from probability theory, that for the surfaces (as distinct from the profiles),

$$\frac{\text{real area of contact}}{\text{unit area of apparent contact}} = \frac{\ell}{L} \quad ,$$

and

$$\text{average area of a junction} \approx d^2 \quad .$$

Furthermore, if it is assumed that plastic deformation is taking place, we may say that over the area of real contact, the normal pressure is equal to the indentation hardness $h \approx 3\sigma_y$, where σ_y = compressive yield stress. With this additional observation,

the normal load per unit apparent area can be related to the separation between the nominal surfaces.

A more involved version of the profilometric approach has been used by Cooper *et al* and by Mikic [*36,46*]. They replace the process of placing the profiles against each other (on transparent paper) by a conceptual mating of the surfaces. Using probability theory, they obtain all the data in the profilometric theory, given (1) the root-mean-square (rms) roughness of the surfaces and (2) the average slopes. Given these parameters, the density of peaks and their height distribution may be predicted, with the help of some results from Statistical Communication Theory. It must be pointed out however, that the specific results used by Cooper *et al* [*36*] are valid only for what is known as a Gaussian narrow-band surface, which looks more-or-less like a series of sinusoidal ridges.

A variation of the approach used by Cooper has been used by Greenwood *et al* [*37-40*] (Greenwood's work predates Cooper's). The approach used by Greenwood assumes that the asperity height distribution, asperity density, and asperity tip curvature are known from profilometric measurements. Greenwood then takes more-or-less the same analytical route as Cooper, but does need to assume either solely elastic or solely plastic deformation.

A fourth approach, used by Rightmire [*53*], is somewhat more sophisticated then the approaches of Cooper, Mikic and Greenwood. Cooper *et al* assume in their analysis that the asperities of a surface touch a flat portion of the mating surface, if they touch it at all. This is based on the argument that the average slopes of most surfaces are quite small, and that an

asperity will therefore have a negligible probability of seeing anything but a gentle slope, with negligible curvature. Rightmire allows for the possibility of interaction between tips of asperities.

Most analytical approaches need to make one of the following assumptions:

a) the contacting surfaces are narrow-band Gaussian surfaces or

b) information on the surface peak height and tip curvature distributions can be directly obtained from a profilogram.

We may say, quite generally, that very few surfaces are both — Gaussian and narrow-band. Assumption (a) must therefore be made very cautiously and with a full awareness of the possibility that such statistics as peak height and surface slope distributions may not be predicted well analytically. Assumption (b) has never, as far as we know, been seriously questioned. However, it seems obvious that the peak height distribution on a surface may at best be related in a fairly complex way with the peak height distribution on a line sample of the surface (i.e., a profilogram); the two distributions are quite unlikely to be identical. Similar remarks quite obviously apply to the tip curvature statistics. This reasoning led to the development of a new model of rough surfaces, which is described in detail in Chap. 5.

1.6 Junction Deformations

In Sec. 1.4, we indicated a certain dissatisfaction with Coulomb's law of friction, especially as applied to rolling contact. What alternative is there? Bowden and Tabor [*90*], in

their investigations of the nature of friction, postulated that friction must be due to the resistance to shear deformation of the metallic junctions which we discussed in Sec. 1.5. As we pointed out in Sec. 1.4.2, the junction-deformation theory of friction is often assumed to lead to Coulomb's law of friction. This may be true in sliding situations, where shear deformations are invariably large, but may not be true in rolling contact [*96*].

There are two schools of thought in the junction-deformation theory of friction. One claims that the junctions must necessarily be plastic due to the normal load. The other claims that this may be true for the first few contacts, but that work-hardening of the asperities must eventually set in (unless there is significant wear); during subsequent contacts, the asperities must be elastic.

Analyses of plastic junctions have been made by Bowden and Tabor [*90*] and by Green [*6,7*]. One of the important characteristics of a plastic junction is that when subject to a shear load, there is a significant growth in the size of the junction. In all these analyses, it is usually assumed that the bond across the interface is perfect, and that the only important parameters are the yield strength of the asperity materials and the geometry of the contact.

Elastic junctions do not seem to have been analyzed in any detail. Some work by Tomlinson on the contact of spherical asperities is, however, relevant, and is reported, for example, in Kragelskii [*10*]. Tomlinson considers the nature of the frictional bond, and tries to relate the friction force during sliding to the area of contact.

The only experimental work seems to have been that of Greenwood and Tabor [8], who built large-scale models of friction junctions, and studied their force-deformation relations. Their results, on the whole, support the theoretical work of Green. Greenwood and Tabor propose the following classification of junctions: (a) Strong, welded junctions, which are formed when the contacting surfaces are very clean and have an affinity for each other. This is the type of junction analyzed by Green. (b) Weak, dry junctions, which are formed when the surfaces are contaminated by oxides, monolayers, etc. (c) Lubricated junctions, which are formed when there is a film of fluid (many molecular layers thick) between the asperities. Figure 1.7 shows the relation between the normal and tangential forces on the junction and the tangential displacement of one half of the function relative to the other, for the three types of function.

Questions that to our knowledge have not been answered, but which are of obvious interest are:

1. What are the rate mechanisms involved in the formation of junctions (e.g., formation of atomic bonds, diffusion of atoms from one asperity into the other)? What are the time constants of these mechanisms?

2. What is the relation between the strength of the adhesive bond and the resistance to shear deformation?

3. What is the effect of the relative velocity of deformation on the force-deformation curve of a junction?

4. What is the effect of surface vibration on the force-deformation curve of a junction?

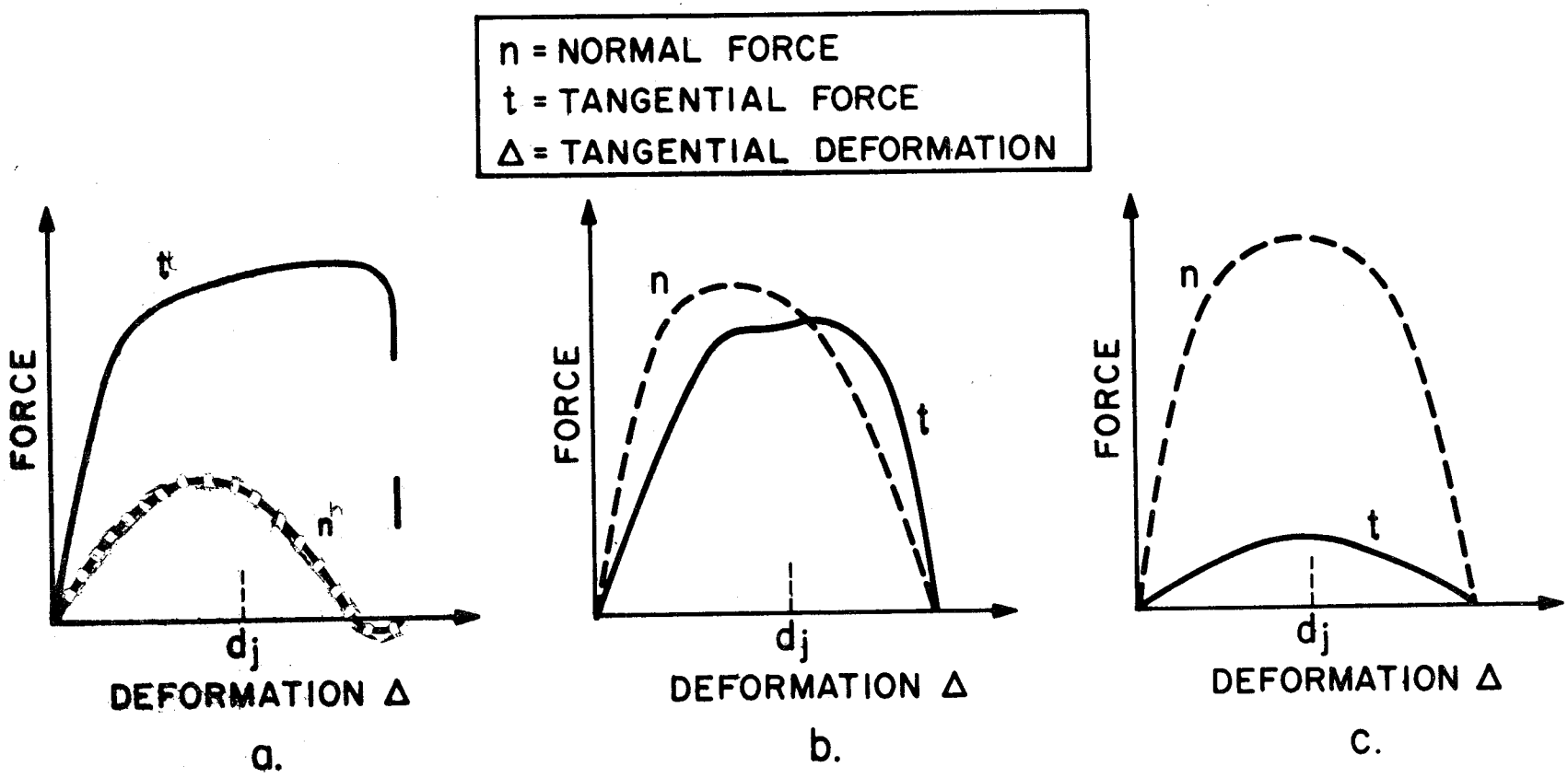


FIG. 1.7 FORCE-DEFORMATION CURVES FOR (a) STRONG JUNCTION; (b) WEAK, DRY JUNCTION; (c) LUBRICATED JUNCTION

5. What are the relative contributions to the tangential force on a junction of (1) mechanical interference, leading to gross elastic or plastic deformation of the asperities, for which the properties of the interface are unimportant, and (2) surface bonding effects?

Figure 1.7 shows the behavior of junctions for monotonic deformation. What if the direction of deformation is reversed part way through the experiment? There are no experimental data available on the behavior of the junction, but it is plausible that the tangential force-deformation curve would be something like that shown in Fig. 1.8.

Consider now what happens when two rough surfaces are pressed against each other and then made to slide, the sliding starting at time t=0. At t=0, there are a large number of junctions of different sizes. As sliding begins, these exert a resistance to shear. For a small time Δt — the time it takes to slide a distance of the order of the average junction diameter — the shear force for each junction increases, as does the total shear force. After some time, however, most of the initial junctions have been broken off, and there is an ensemble of junctions in various stages of their life process — being formed, being sheared, being destroyed — and the statistics of this ensemble do not change with time. What this means is that the total shear force is approximately constant in time after a certain amount of sliding.

It is obvious from this discussion that the relevant independent variable is not really the time t, but the distance moved, x. Thus, we could plot the friction coefficient f (shear force/normal force) as a function of x. If after some distance x_0 the sliding is reversed, the force on each junction falls off steeply,

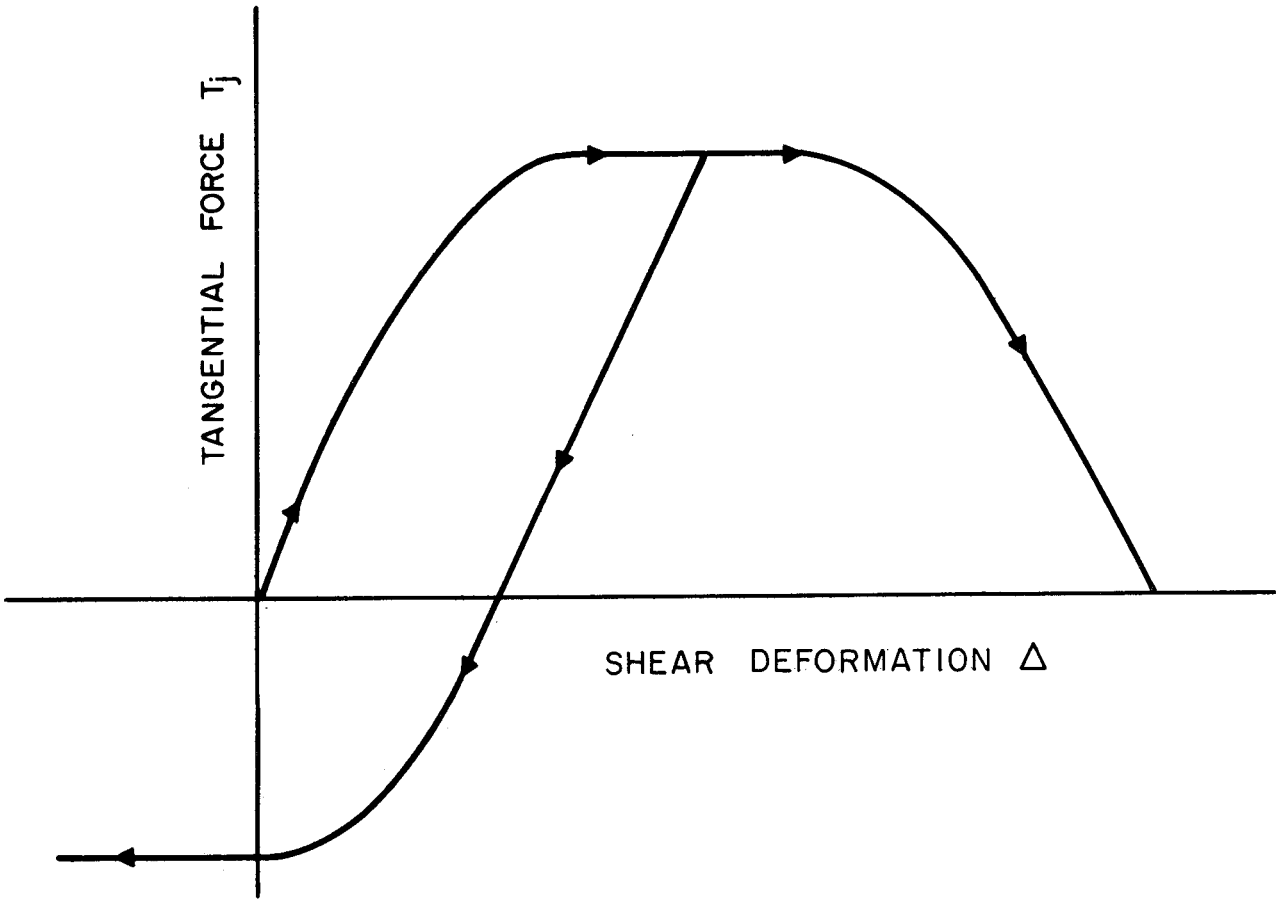


FIG. 1.8 PROPOSED JUNCTION DEFORMATION WITH VELOCITY REVERSAL

as in Fig. 1.7, and so does the total tangential force. The behavior of the friction coefficient with the distance slid, x, would be qualitatively that shown in Fig. 1.9.

What is the relation between the friction law of Fig. 1.9 and Coulomb's law of friction, which was discussed in Sec. 1.4.3? From Fig. 1.9, we see that except for small distances near points where the velocity is reversed, the friction coefficient is a constant, independent of velocity. In the region of velocity reversals, however, f as a function of time may be determined from the velocity history V(t), instead of being indeterminate, as in Coulomb's law.

In sliding situations, the difference between these two laws of friction is negligible. In rolling contact or oscillating contact, however, the effect may be significant [*96*].

Some experiments on junction deformation performed during the course of our contract are described in Chap. 3.

1.7 Rolling of Railway Wheels

A number of factors influence the rolling contact of railway wheels and rails, which have not been studied in detail in the laboratory. Some of these factors are:

a) Due to various reasons, there is a great deal of vibration of the contacting surfaces.

b) The contacting surfaces are usually contaminated, often by a large variety of contaminants.

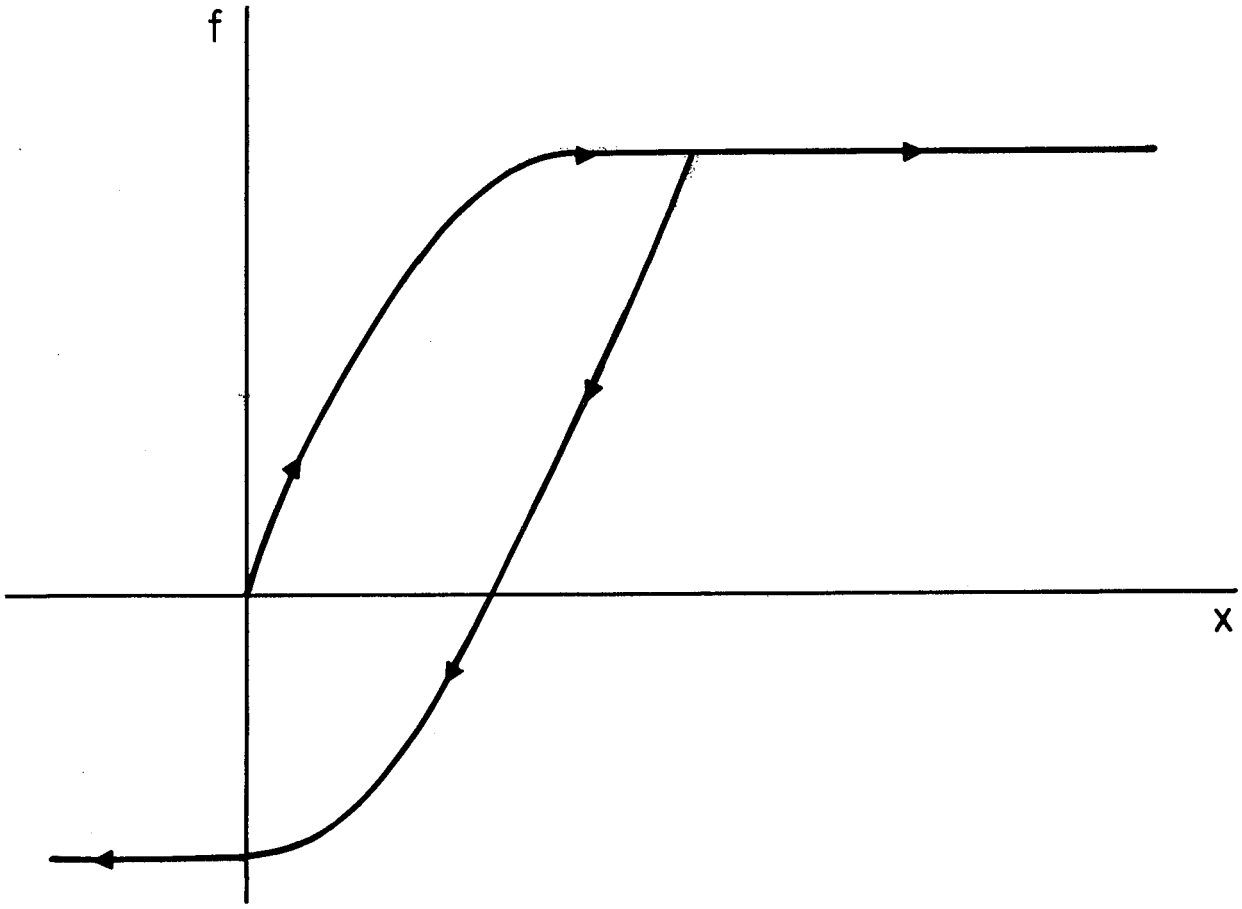


FIG. 1.9 FRICTION COEFFICIENT AS A FUNCTION OF DISTANCE SLID (WITH REVERSAL)

c) There are significant variations in time of the normal and tangential loads on the wheel. The normal load variations may arise from oscillations of the truck and carriage, or from the waviness of the track. The tangential load variations may arise from flange contact, from oscillations of the truck and carriage, from windgusts, from track curvature, etc.

The effects of speed of rolling or of surface vibration on the mechanics of the contact may be the following:

a) As the speed increases, there is less time for the junctions to form, and their strength may decrease. Some rate processes involved in the formation of junctions are the formation of atomic bonds (very fast), the diffusion of oxygen atoms away from the interface (slow), and the diffusion of surface metal atoms to provide more compatible surfaces (very slow).

b) Just as in sliding contact, the effect of surface vibration may be to decrease the friction coefficient f. The effect of this, with wheel loads held constant, would be to increase the slip.

c) Normal load variations may invalidate the results of the steady-state theory of rolling contact. The net effect is not likely to be great, however. The problem of unsteady contact presents a very prickly mathematical aspect, as evidenced by the work of Kalker [*23*]. Kalker assumes Coulomb's law of friction to hold in the contact region, and investigates the effects of tangential load variations on the friction stresses and on the slip.

d) Even though the smooth-surface, steady-state theory of rolling contact may be valid, dynamic normal load variations may cause a decrease in the effective friction coefficient between the rail and wheel. This possibility is important not merely for creep effects in the rolling of the wheels, but in predictions of the tractive capacity of the train, which governs the acceleration and braking of the train. This question is explored in detail in Chap. 4.

Detailed investigations of wheel-rail friction may be found in Refs. 14, 18-20, 24, 67 and 98-100.

1.8 Some Experiments

Although experimental investigations of rolling contact were begun as early as 1928 [*56*] (and possibly earlier), the first comprehensive studies were made by Johnson [*63,64*] in the 1950's and by Haines and Ollerton [*61*] and Haines [*62*] in the 1960's.

Johnson investigated the rolling contact of a steel ball against a steel surface, at fairly low rolling speeds and fairly high normal loads (leading to average pressures not less than 70,000 psi. The deformations are nevertheless elastic, because the normal stresses in the material near the contact region are mainly hydrostatic. By rolling a loaded ball down an inclined plane with a restraining force, Johnson was able to investigate both longitudinal and lateral creep, and found very good agreement with an approximate theory he had developed.

Johnson also investigated the effect of spin on a rolling ball, by observing balls in thrust bearings. He observed that a

sideward force was generated due to the spin, and that unless a restraining force was applied, the ball would creep sideways as well as spinning. Again, the rolling velocities were low, the average pressure in the contact region was high, and good agreement was found between experiment and approximate theory.

Halling [*59*] also investigated lateral and transverse creep of steel balls on steel plates with an apparatus somewhat more sophisticated than Johnson's, and also found good agreement between experiment and theory. As in Johnson's experiments, the velocities were low and the loads high (average pressures greater than 115,000 psi).

Halling and Brothers [*60*] investigated the effects of surface roughness on the creep of rolling balls, and found that surface roughness had no apparent effect. It must again be said, however, that the normal loads used were quite high (average pressures greater than 110,000 psi), and the only reasonable conclusion is that surface roughness is unimportant if the loads are sufficiently high.

Haines and Ollerton [*61*], Haines [*62*] and Poon [*68*] all investigated creep or spin by means of frozen photoelastic techniques. All found excellent agreement between stresses calculated from experiment and those predicted by theory. We must note, however, that these photoelastic techniques implicitly assume that the surfaces of the contacting bodies are smooth.

In conclusion, it is worth noting that creep measurements made in the field sometimes disagree substantially with theory and with measurements made in the laboratory [*2.2*]. Some reasons for this were discussed in Sec. 1.7.

REFERENCES

A. *Junction Deformation*

1. S.B. Ainbinder and A.S. Prancs, "On the Mechanism of the Formation and Destruction of Adhesion Junctions Between Bodies in Frictional Contact," *Wear* **9**, 209-227 (1966).

2. F.P. Bowden and D. Tabor, *The Friction and Lubrication of Solids, Vol. II*, Oxford University Press, England (1964), pp. 52-86.

3. M. Cocks, "Shearing of Junctions Between Metal Surfaces," *Wear* **9**, 320-328 (1966).

4. J.S. Courtney-Pratt and E. Eisner, "The Effect of a Tangential Force on the Contact of Metallic Bodies," *Proc. Roy. Soc.* (London) **238**, 529-550 (1957).

5. E. Finkin, "Friction, Adhesive Wear and Abrasive Wear, Principally in Metals: A Review," Paper No. 63-WA-89 presented at the ASME Winter Annual Meeting, Philadelphia, Pa. (1963).

6. A.P. Green, "The Plastic Yielding of Metal Junctions Due to Combined Shear and Pressure," *J. Mech. Phys. Solids* **2**, 197-211 (1954).

7. A.P. Green, "Friction Between Unlubricated Metals: A Theoretical Analysis of the Junction Model," *Proc. Roy. Soc.* (London) **228**, 191-204 (1955).

8. J.A. Greenwood and D. Tabor, "Deformation Properties of Junctions," *Proc. Phys. Soc.* (London) **58**, 609-619 (1955).

9. I.V. Kragelskii and V.P. Sabelnikov, "Experimental Check of Elementary Law of Friction: Dry Friction," *Proc. Inst. Mech. Engrs., Conference on Lubrication and Wear*, 246-251 (1957).

10. I.V. Kragelskii, *Friction and Wear*, Butterworth and Co. Ltd., England (1965), pp. 138-190.

11. I.V. Kragelskii, "Calculation of Dry Friction Forces," *Proc. Inst. Mech. Engrs., Conference on Lubrication and Wear*, 302-307 (1957).

12. C.H. Popelar, "On the Basic Equation of Junction Growth," *ASME Trans. J. Appl. Mech.* 91, 132-133 (1969).

13. A. Seireg and E.J. Weiter, "Viscoelastic Behavior of Frictional Hertzian Contacts Under Ramp-Type Loads," *Proc. Inst. Mech. Engrs.* 181, Pt. 30, 200-206 (1966-1967).

B. Vibration Effects and Related Topics

14. H.T. Albachten, "Locomotive Wheel-to-Rail Traction," *ASME Paper* No. 66-WA/RR-7 (1966).

15. F.P. Bowden, "Recent Experimental Studies of Solid Friction," *Friction and Wear,* R. Davies, Ed. (Elsevier Publishing Co., New York (1959).

16. J.D. Coran, "A Study of Non-Constant Creep Rate in Rolling Contact," BS Thesis, Dept. of Mech. Eng., MIT (May 1968).

17. R. Courtel, "Normal Vibrations in Contact Friction," *Wear* 11, 77 (1968).

18. M.D. Hersey, "Rolling Friction, II - Cast-Iron Car Wheels," *ASME Paper* No. 68-Lub-C (1968).

19. M.D. Hersey, "Rolling Friction, III - Review of Later Investigations," *ASME Paper* No. 68-Lub-D (1968).

20. M.D. Hersey, "Rolling Friction, IV, Additional Car Wheel Experiments," *J. Lub. Tech., ASME Trans.* (1969).

21. L.O. Hewko, F.G. Rounds, and R.L. Scott, "Tractive Capacity and Efficiency of Rolling Contacts," *Rolling Contact Phenomena,* J.B. Bidwell, Ed., Elsevier Publishing Co., New York (1962).

22. A.E.W. Hobbs, "A Survey of Creep," British Railways Research Dept., DYN52, pp. 6-9 (1967).

23. J.J. Kalker, "Transient Phenomena in Two Elastically Similar Rolling Cylinders in the Presence of Dry Friction," Laboratorium Voor Technische Mechanica, Technological University, Delft, Netherlands, Report No. 384 (Jan. 1968).

24. W.B. Kirk, "High-Speed Train Operation," *ASME Paper* No. 65-WA/RR-5 (1965).

25. R.V. Klint, "Oscillating Tangential Forces on Cylindrical Specimens in Contact at Displacements within the Region of No Slip," *Am. Soc. Lub. Eng.* 3, 255-264 (1960).

26. R.D. Mindlin, W.P. Mason, J.F. Osmer, and H. Deresiewicz, "Effects of an Oscillating Tangential Force on the Contact Surfaces of Elastic Spheres," *Proc. First U.S. National Congress of Appl. Mech.*, 203-208 (1951).

27. A.M. Mitskevich, "Motion of a Body Over a Tangentially Vibrating Surface, Taking Account of Friction," *Soviet Phys.-Acoust.* 13, 348-351 (1968).

28. Y. Yokoyama, "Effects of Vibration on Static and Kinetic Friction," *J. Jap. Soc. Lub. Eng.* 13, 62-70 (1968).

29. A. Seireg and E.J. Weiter, "Frictional Interface Behavior Under Dynamic Excitation," *Wear* 6, 66-77 (1963).

30. D.M. Tolstoi, "Significance of the Normal Degree of Freedom and Natural Normal Vibrations in Contact Friction," *Wear* 10, 199-213 (1967).

C. Rough Surface Analyses and Experiments

31. J.F. Archard, "Contact and Rubbing of Flat Surfaces," *J. Appl. Phys.* 24, 981-988 (1953).

32. J.F. Archard, "Elastic Deformation and the Contact of Surfaces," *Nature*, 918-919 (1953).

33. J.F. Archard, "Elastic Deformation and the Laws of Friction," *Research Correspondence* 9, S26-S27 (July 1956).

34. F.P. Bowden and D. Tabor, *The Friction and Lubrication of Solids, Vol. I*, Oxford University Press, England, pp. 10-30 (1950).

35. M. Cocks, "Interaction of Sliding Surfaces," *J. Appl. Phys.* 33, 2152-2161 (1962).

36. M.G. Cooper, B.B. Mikic, and M.M. Yovanovich, "Thermal Contact Conductance," MIT (1968).

37. J.A. Greenwood and J.B.P. Williamson, "Contact of Nominally Flat Surfaces," *Proc. Roy. Soc.* (London) **295**, 300-319 (1966).

38. J.A. Greenwood, "The Area of Contact of Rough Surfaces and Flats," *J. Lub. Tech., ASME Trans.* **89**, 81-91 (1967).

39. J.A. Greenwood and J.H. Tripp, "The Elastic Contact of Rough Spheres," *J. Appl. Mech., ASME Trans.* **89**, 153-159 (1967).

40. J.A. Greenwood and H. Tripp, "The Contact of Two Nominally Flat Surfaces," (to be published).

41. J. Halling and B.G. Brothers, "The Effect of Surface Finish on the Creep and Wear of a Rolling Ball Subject to Normal and Tangential Surface Tractions," *Wear* **9**, 199-208 (1966).

42. I.V. Kragelskii, *Friction and Wear,* Butterworth and Co. Ltd., England (1965).

43. J.D. Huffington, "A Theory of Friction Involving a Micro-topographical Deformation Index," *Wear* **3**, 473-476 (1960).

44. F.F. Ling, "Some Factors Influencing the Area-Load Characteristics for Semismooth Contiguous Surfaces Under 'Static' Loading," *J. Appl. Mech. ASME Trans.* **80**, 1113-1120 (1958).

45. W.P. Mason, "Adhesion Between Metals and Its Effects on Fixed and Sliding Contacts," *Am. Soc. Lub. Eng.* **2**, 39-49 (1959-1960).

46. B.B. Mikic, "Effect of Previous Loading on Thermal Contact Conductance," MIT (1968).

47. D.F. Moore, "A Note on the Mathematical Representation of Surface Roughness," *Wear* **9**, 477-479 (1966).

48. N.O. Myers, "Characterization of Surface Roughness," *Wear* **5**, 182-189 (1962).

49. J.J. O'Connor and K.L. Johnson, "The Role of Surface Asperities in Transmitting Tangential Forces Between Metals," *Wear* **6**, 118-139 (1963).

50. E. Rabinowicz, "Autocorrelation Analysis of the Sliding Process," *J. Appl. Phys.* **27**, 131-135 (1956).

51. E. Rabinowicz, "Investigation of Size Effects in Sliding by Means of Statistical Techniques," *Proc. Inst. Mech. Engrs., Conference on Lubrication and Wear,* Paper No. 36 (1957).

52. B.G. Rightmire, "Probable Behavior of Contacts in the Sliding Process," *Proc. Inst. Mech. Engrs., Conference on Lubrication and Wear,* 281-285 (1957).

53. B.G. Rightmire, "Solid-Solid Interfaces," Unpublished MIT Notes (1965).

54. T.E. Tallian, Y.P. Chiu, D.F. Huttenlocher, J.A. Kamenshine, L.B. Sibley, and N.E. Sindlinger, "Lubricant Films in Rolling Contact of Rough Surfaces," *Am. Soc. Lub. Eng. Trans.* 7, 109-126 (1964).

55. T. Tsukizoe and T. Hisakado, "On the Mechanism of Contact Between Metal Surfaces — The Penetrating Depth and the Average Clearance," *J. Basic Eng., ASME Trans.* 87, 666-674 (1965).

D. Rolling Contact Experiments

56. A. Palmgren, "Investigations with Regard to Rolling Under Tangential Pressure," *SKF Ball Bearing J.* 3, 67-76 (1928).

57. F.T. Barwell and R.G. Woolacott, "The N.E.L. Contribution to Adhesion Studies," *Inst. Mech. Engrs., Proc. Conf. on Adhesion,* 145-160 (1963).

58. R.D. Davies, "Some Experiments on the Lateral Oscillations of Railway Vehicles," *J. Inst. Civil Eng.* 11, 224-261 (1939).

59. J. Halling, "The Rolling of a Ball Subjected to Normal and Tangential Loads," *Wear* 7, 516-534 (1964).

60. J. Halling and B.G. Brothers, "The Effect of Surface Finish on the Creep and Wear of a Rolling Ball Subjected to Normal and Tangential Surface Tractions," *Wear* 9, 199-208 (1966).

61. D.J. Haines and E. Ollerton, "Contact Stress Distribution on Elliptical Contact Surfaces Subject to Radial Forces," *Proc. Inst. Mech. Engrs.* 177, 95-114 (1963).

62. D.J. Haines, "Contact Stress in Flat Elliptical Contact Surfaces which Support Radial and Shearing Forces During Rolling," *Proc. Inst. Mech. Engrs.* 179, 154-165 (1964-1965).

63. K.L. Johnson, "The Effect of Spin Upon the Rolling Motion of an Elastic Sphere on a Plane," *J. Appl. Mech., ASME Trans.* 80, 332-338 (1958).

64. K.L. Johnson, "The Effect of a Tangential Force Upon the Rolling Motion of an Elastic Sphere on a Plane," *J. Appl. Mech., ASME Trans.* 80, 339-346 (1958).

65. K.L. Johnson, "The Influence of Elastic Deformation Upon the Motion of a Ball Rolling Between Two Surfaces," *Proc. Inst. Mech. Engrs.* 173 (1959).

66. W.B. Kirk, "High Speed Train Operation," *ASME Paper* No. 65-WA/RR-5 (1965).

67. N. Matsui and K. Yokose, "On the Slip Phenomenon Between Wheel and Rail," Railway Technical Research Institute, *Jap. Natl. Railways* 3 (1966).

68. S.Y. Poon, "An Experimental Study of the Shear Traction Distribution in Rolling with Spin," *Wear* 10, 61-69 (1967).

E. Theory

69. R.H. Bentall and K.L. Johnson, "Slip in the Rolling Contact of Two Dissimilar Elastic Rollers," *Int. J. Mech. Sci.* 9, 289-404 (1967).

70. F.W. Carter, "On the Action of a Locomotive Driving Wheel," *Proc. Roy. Soc.* (London) 112, 151 (1926).

71. A.D. DePater, "On the Reciprocal Pressure Between Two Elastic Bodies," in *Rolling Contact Phenomena*, J.B. Bidwell, Ed. Elsevier Publishing Co., New York (1962).

72. *Fatigue in Rolling Contact*, Proceedings of a Symposium arranged by The Institution of Mechanical Engineers, London (1964).

73. L.A. Galin, "Contact Problems in the Theory of Elasticity," Dept. of Math., School of Phys. Sciences and Appl. Math., North Carolina State College (1961).

74. D.J. Haines and E. Ollerton, see Ref. 61.

75. D.J. Haines, see Ref. 62.

76. K.L. Johnson, see Ref. 63.

77. K.L. Johnson, see Ref. 64.

78. K.L. Johnson, see Ref. 65.

79. K.L. Johnson, "Tangential Tractions and Microslip in Rolling Contact," in *Rolling Contact Phenomena*, J.B. Bidwell, Ed. Elsevier Publishing Co., New York (1962).

80. K.L. Johnson, "A Review of the Theory of Rolling Contact Stresses," *Wear* **9**, 4 (1966).

81. K.L. Johnson and D. Tabor, "Rolling Friction," *Proc. Inst. Mech. Engrs.* **182**, 168 (1967-1968).

82. J.J. Kalker, "Rolling with Slip in the Presence of Dry Friction," *Wear* **9**, 20 (1966).

83. J.J. Kalker, "A Strip Theory for Rolling with Slip and Spin," Parts I-III, *Proc. Koninkl. Nederl. Akademie van Wetenschappen-Amsterdam* **70**, 10-62 (1967).

84. J.J. Kalker, "On the Rolling Contact of Two Elastic Bodies in the Presence of Dry Friction," Doctoral Dissertation, Technische Hogeschool, Delft, Netherlands (1967).

85. J.J. Kalker, see Ref. 23.

86. J.E. Merwin and K.L. Johnson, "An Analysis of Plastic Deformation in Rolling Contact," *Proc. Inst. Mech. Engrs.* **177**, 676 (1963).

87. V.I. Mossakovskii and I.I. Mishchishin, "Rolling of Elastic Bodies," *Prikl. Math. i Mekh* **31**, 870 (1967).

88. H. Poritsky, "Stresses and Deflections of Cylindrical Bodies in Contact," *J. Appl. Mech., ASME Trans.* **17**, 191 (1950).

89. P.J. Vermeulen and K.L. Johnson, "Contact of Non-Spherical Elastic Bodies Transmitting Tangential Forces," *J. Appl. Mech., ASME Trans.* **86**, 338 (1964).

F. General References

90. F.P. Bowden and D. Tabor, *The Friction and Lubrication of Solids, Vols. I and II,* Oxford University Press, England (1950, 1964).

91. F.P. Bowden and D. Tabor, *Friction and Lubrication*, Methuen and Co., Ltd., London (1967).

92. R. Davies, Ed., *Friction and Wear,* Elsevier Publishing Co., New York (1959).

93. D. Dowson and G.R. Higginson, *Elasto-Hydrodynamic Lubrication,* Pergamon Press, New York (1966).

94. R. Houwink and G. Salomon, *Adhesion and Adhesives, Vols. I and II,* Elsevier Publishing Co., New York (1967).

95. I.V. Kragelskii, *Friction and Wear,* Butterworths, Washington (1965).

96. P.R. Nayak, "A New Theory of Rolling Contact," PhD Thesis, Dept. of Mech. Eng., MIT (April 1968).

97. E. Rabinowicz, *Friction and Wear of Materials*, John Wiley and Sons, Inc., New York (1965).

G. Additional References

98. K.R. Kilburn, "An Introduction to Rail Wear and Rail Lubrication Problems," *Wear* 7, 255 (1964).

99. H.A. Marta and K.D. Mels, "Wheel-Rail Adhesion," *J. Engg. Ind.* 91, Series B, No. 3, p. 839 (1969).

100. R.T. Spurr, "Some Measurements of Rail-Tyre Adhesion," *Wear* 3, 463 (1960).

2. ROLLING CONTACT OF ROUGH SURFACES

Summary

A refined version of the rough-surface theory of rolling contact is presented. The main conclusions of the work reported in this chapter are as follows:

1. It is now possible, by a profilometric examination of the wheel rim and rail surfaces, to determine whether or not surface roughness effects will be important in any given case.
2. For fairly smooth wheel and rail surfaces, with rms roughness ≈ 10-20μ in. and at present operating loads, surface roughness is not important.
3. The above conclusion has to be qualified, however, by the statement that surface roughness effects may become important in any of the following cases.
   a) Normal loads are decreased.
   b) Surfaces become rougher.
   c) Asperity vibrations are determined to be of importance.
   d) Adhesion in the interface is increased by treatment of the contacting surfaces.
4. When the surfaces are cleaned with a plasma torch, the behavior of the junctions in the interface becomes markedly different from the behavior of the boundary-lubricated junctions assumed in the present theory. However, the theoretical framework is well suited for adaptation to a study of the rolling contact of clean surfaces. An experimental and theoretical investigation of the rolling contact of plasma-cleaned surfaces seems highly desirable.

5. The concept of contact resonance is studied in some detail, and it is shown that it may be of importance in the frictional behavior of rolling elements. It appears, however, that even when the contact resonance frequency is excited, it is necessary to examine the effect of asperity vibrations on the behavior of junctions in order to explain decreases in the friction coefficient. Thus, other resonances, such as those of the suspension supporting the rolling elements, are likely to have the same effect on the friction as the contact resonance, the most important thing being the magnitude of the contact vibrations; these are large at all resonances. Some experimental observations of creep in the presence of vibrations confirm the importance of the contact resonance. Further studies of the effects of resonances on friction and creep in the presence of broadband random excitation are needed.

2.1 Introduction

In all analyses of the rolling contact of elastic bodies, starting with the work of Carter [1], several restrictive assumptions have been made, two of the most basic of which are: (1) Coulomb's law of friction [2] applies in the contact region between the rolling body and the supporting surface; and (2) the contacting surfaces are ideally smooth, in the sense that there are no surface irregularities.

We present a theory of rolling contact which is new in that it relaxes both these assumptions. We show that the results in rolling contact theory depend on a surface roughness parameter and indicate the nature of this dependence. Assuming that friction in the contact region arises from the deformation of friction junctions [2] between the surfaces, we propose a law of friction based on experimental observations of junction deformation. The theoretical results show that surface roughness has an influence on creep measurements in rolling contact. This influence may be large in some cases and negligible in others. Partial experimental validation of this conclusion is presented.

All the results reported here are for two-dimensional, steady-state contact of bodies of like materials. However, extensions to the three-dimensional problem are indicated.

2.2 The Smooth-Surface Theory

A summary of the current theory of rolling contact will serve to provide background for our new theory. This older theory we term "the smooth-surface theory".

Consider the problem depicted in Fig. 2.1(a). A smooth infinite, elastic cylinder of radius R is pressed against a smooth, elastic, semi-infinite solid. Both are made of the same material — with Young's modulus E, Poisson's ratio ν, and shear modulus G. The load normal to the surface of the semi-infinite solid is N lb/in.; the tangential load is T lb/in. A moment M = TR lb-in./in. is applied to restrain cylinder acceleration. Finally, Coulomb's law of friction is assumed to apply in the contact region. What is desired is the extent of the contact region, the normal pressure p, and the friction stress τ in the contact region. In addition, a quantity ξ, called the "creep" or "slip" is desired. If, in m revolutions of the cylinder, the center of the cylinder moves through a distance L, the creep is defined as

$$\xi = \frac{2\pi mR - L}{L} = \frac{\text{distance slipped}}{\text{distance rolled}} . \tag{1}$$

The solution to this problem was first presented by Carter [*1*], and later, in greater detail, by Poritsky [*3*]. We shall present some of their results.

We shall assume a familiarity with the Eulerian approach, in which the contact region is considered to be fixed, with the material of the two contacting bodies flowing through it (Fig. 2.2). Let the velocities of the upper and lower surfaces at points distant from the contact region be V_{10} and V_{20}. If u(x) is the elastic displacement parallel to the surface of surface points of the cylinder, the velocities in the contact region are:

$$V_1 = V_{10}\left(1 + \frac{\partial u}{\partial x}\right) , \quad V_2 = V_{20}\left(1 - \frac{\partial u}{\partial x}\right) . \tag{2}$$

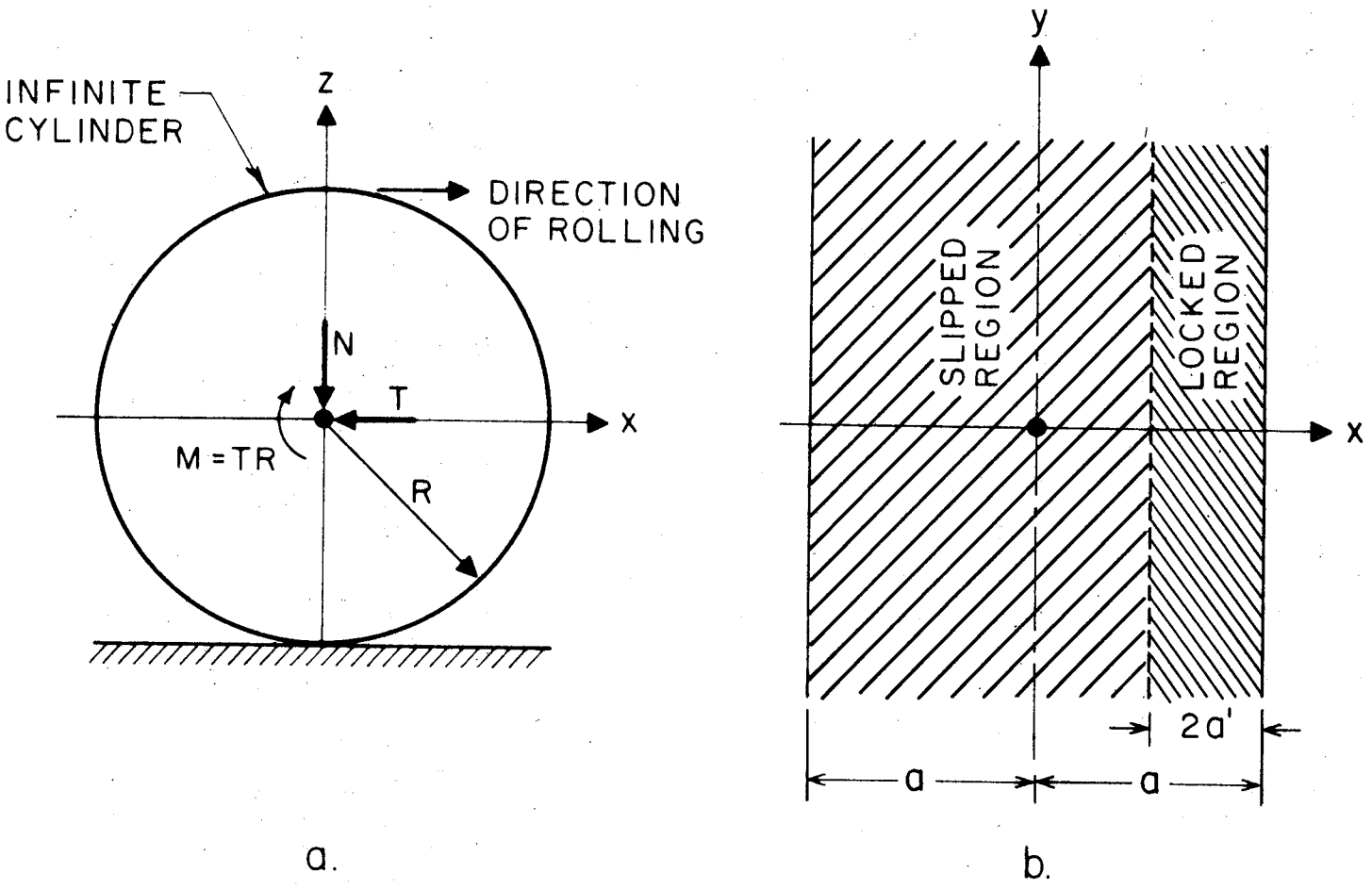


FIG.2-1 CYLINDER ROLLING ON A PLANE. (a) FORCES (b) CONTACT REGION

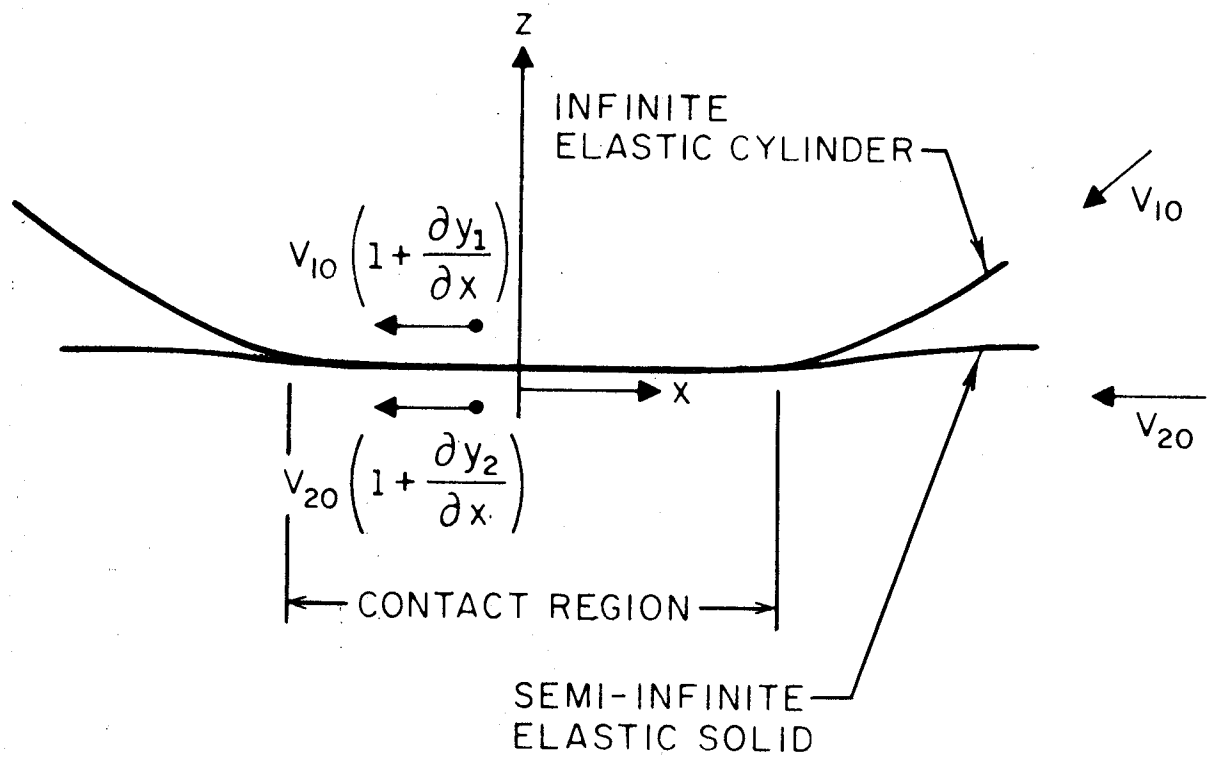


FIG.2-2 PARTICLE VELOCITIES IN THE CONTACT REGION

The slip velocity at which the lower surface slips past the upper along the positive x-axis is:

$$V_s = V_1 - V_2 = V\left(\xi + 2\frac{\partial u}{\partial x}\right) , \tag{3}$$

where $V \equiv 1/2(V_{10}+V_{20})$ and $\xi \equiv (V_{10}-V_{20})/V$. Coulomb's law of friction, as applied to rolling contact, states that the friction stress τ on the cylinder in the contact region is given by:

$$|\tau| \leq fp, \quad \text{if } V_s = 0 ;$$

$$\tau = fp \ \text{sign}(V_s) \ \text{otherwise} , \tag{4}$$

where f is a friction coefficient. Note that we define positive shear stresses on the cylinder as acting along the positive x axis.

The semiwidth of the contact region, "a", and the normal pressure in the contact region, p, are independent of the tangential load T:

$$a = \left\{\frac{8(1-\nu^2)NR}{\pi E}\right\}^{1/2} , \tag{5}$$

and

$$p = \frac{2N}{\pi a}\left(1 - \frac{x^2}{a^2}\right)^{1/2} , \tag{6}$$

The solution is completed by locating a "locked" region adjacent to the leading edge, in which the slip velocity is zero. This

locked region extends over the range $a - 2a' \leq x \leq a$, where a' is given by

$$a' = a\left(1 - \frac{T}{fN}\right)^{1/2} . \tag{7}$$

The friction stress in the locked region is given by

$$\tau = fp\left\{1 - \left[\frac{(x-a+a')^2 - a'^2}{a^2}\right]^{1/2}\right\} . \tag{8}$$

The rest of the contact region is a "slipped" region, in which the slip velocity is nonzero. It extends over the range $-a \leq x \leq a - 2a'$. The stress in the slipped region is given by

$$\tau = fp . \tag{9}$$

With the friction stresses given by Eqs. 10 and 11, we find that Coulomb's law of friction is satisfied everywhere. The following expression then holds for the creep:

$$\xi' = \left[1-(1-T'^2)^{1/2}\right] \frac{4(1-\nu)}{\pi} , \tag{10}$$

where $\xi' \equiv \xi aG/fN$ and $T' \equiv T/fN$ are a dimensionless creep and tangential load, respectively. It is worth remarking that a unique, one-to-one relation is found between T' and ξ'.

2.3 The Rough-Surface Theory

2.3.1 The static contact of rough surfaces

The surface of a rough circular cylinder may be described by the nominal radius R and the probability distributions of various statistics of the surface roughness. These statistics are, for

example, the height of the surface, the number of peaks (or asperities) per unit area on the surface, the probability distribution of the heights of these asperities, the radius of curvature of the tips of the asperities, their side slopes, etc.

The contact of rough surfaces has been analyzed in some detail for various geometries by Greenwood and Williamson [*4*], Greenwood [*5*], Greenwood and Tripp [*6*,*7*], Lo [*8*], and by Rightmire [*9*]. The following discussion briefly summarizes the results of these analyses, for the contact of a rough cylinder and a half-space.

1. If the elastic displacement normal to the surface, $v(x)$, is much larger than the root-mean-square roughness of the surfaces, then the extent of the contact region and the pressure p within it are given by Eqs. 5 and 6. This condition is satisfied when

$$\frac{2N(1-\nu^2)}{\pi E\sigma} >> 1 , \qquad (11)$$

where σ = root-mean-square surface roughness of the rougher surface.

2. For a given normal load N, the expected value of the diameter of the junctions is the same everywhere in the contact region.

3. The expected value of the density of junctions η (that is, the number of junctions per unit length perpendicular to the cylinder axis) is approximately proportional to the local normal pressure, p.

4. The average junction diameter increases with the normal load N, but only very slowly.

To derive these results, we must assume that the deformation of the asperities is elastic — a valid assumption when the surfaces have been repeatedly brought into contact. We will not discuss this assumption further, preferring to refer the interested reader to the literature. Suffice it to say that we could work equally well with a more complex model, taking into account plastic or elasto-plastic deformation of the asperities.

2.3.2 The rolling contact of rough surfaces

Suppose a rough cylinder is rolling on a plane, as in Fig. 2.1. Taking the Eulerian viewpoint, we assume that the contact region is fixed, and that the material of the two bodies flows through the region, as in Fig. 2.2.

To obtain the friction stress $\tau(x)$, we no longer assume that Coulomb's law of friction holds. Instead, we adopt a more realistic viewpoint, and consider the friction stresses to arise from the shear deformation of junctions. In order to expound this point of view, we must describe briefly the behavior of friction junctions.

Greenwood and Tabor [*10*] describe three types of junction: (1) Strong, welded junctions, formed when the surfaces are very clean and have an affinity for each other; (2) weak, dry junctions, formed when the surfaces are contaminated by oxides, monolayers, etc.; and (3) lubricated junctions, formed when there is a film of fluid between the contacting surfaces. A relation can be obtained between the tangential load t lbf or the normal load n lbf, and the shear displacement Δ of one half of the junction relative to the other. Greenwood and Tabor [*10*] obtained experimental data on the behavior of the three types of junction, and their results are shown qualitatively in Fig. 2.3.

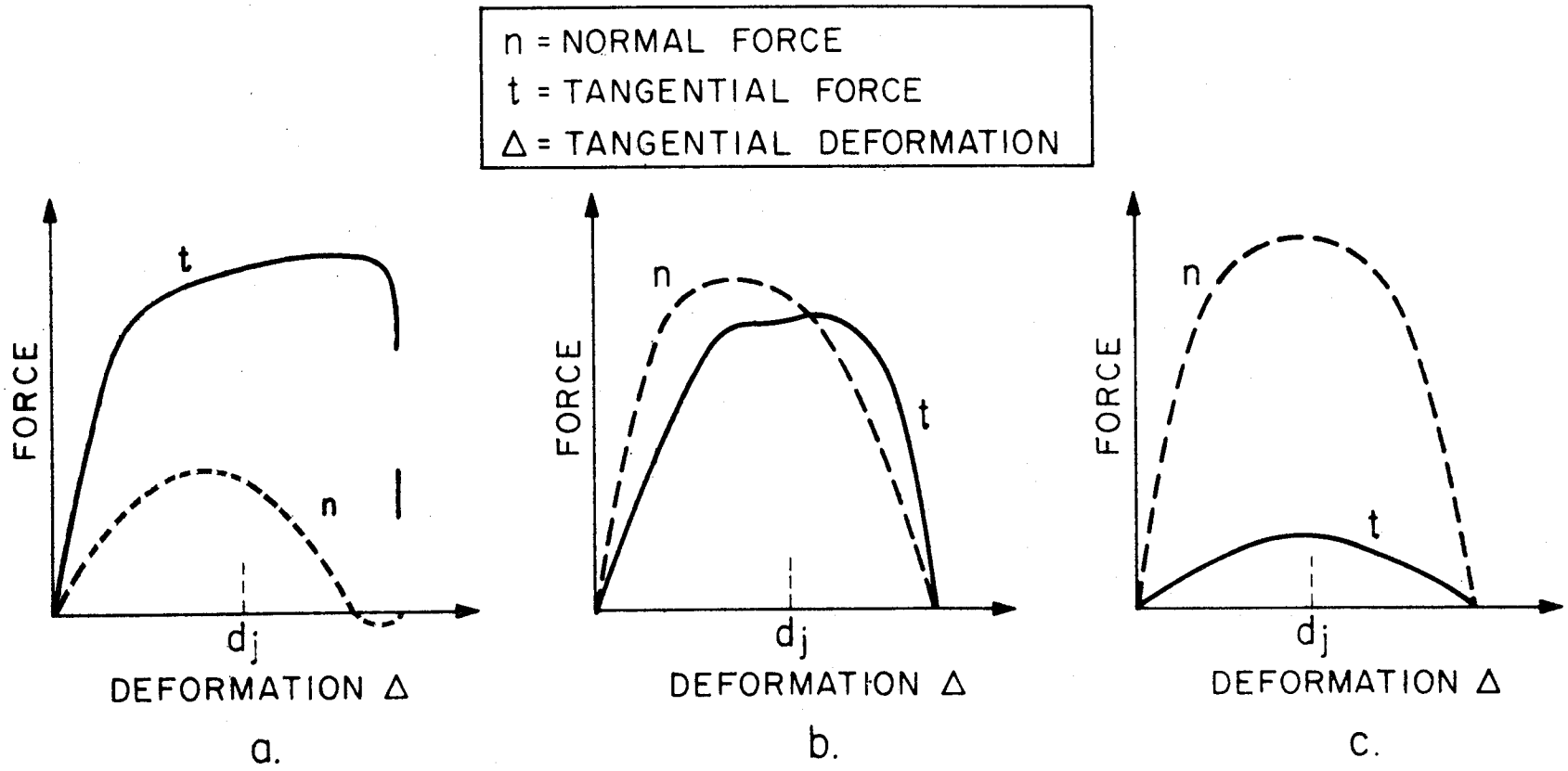


FIG.2-3 FORCE-DEFORMATION RELATIONS FOR (a) STRONG, (b) WEAK, DRY AND (c) LUBRICATED JUNCTIONS

We shall now examine the behavior of these junctions in the contact region of Fig. 2.2, and shall make a few assumptions that serve to simplify the mathematics, with the justification that we are interested more in the qualitative aspects of the theory, than in the exact numerical results obtained.

Assume that all junctions formed have the same diameter d_j. An appropriate value for d_j is the expected value of the junction diameter, which is the same everywhere in the contact region. Assume further that as the junction travels through the contact region, it does not change in size. Next, assume that all the junctions are weak, dry junctions, with no tendency to adhere to each other, and with a force-deformation relation such as shown in Fig. 2.3(b).

For deformations Δ that are small compared to d_j, the tangential force never decreases with Δ, and the force-deformation curve may be modeled by the equation

$$t = t_{max} \{1-\exp(-\Delta/cd_j)\} , \tag{12}$$

where c is a numerical constant. The tangential force is related to the tangential stress in the junction interface,

$$\tau = \frac{t}{\frac{\pi d_j^2}{4}} = \frac{t_{max}}{\frac{\pi d_j^2}{4}} \{1-\exp(-\Delta/cd_j)\} \equiv \tau_{max}\{1-\exp(-\Delta/cd_j)\} . \tag{13}$$

Consider a junction formed in the region (x_0, x_0+dx_0) and travelling with the material of the two bodies in the direction of the negative x-axis (Fig. 2.2). Whenever the slip velocity V_s, given by Eq. 8, is nonzero, the upper half of the junction tends to slip past the lower half. In the neighborhood of some point x_1, the slip velocity is given by Eq. 3. The junction

traverses a distance $(-dx_1)$ near x_1 in the time $dt = (-dx_1)/V$. In this time, the amount of shear deformation is

$$d(\Delta) = V_s(x_1)\cdot dt = -\left(\xi+2\frac{\partial u}{\partial x}\right)_{x_1} dx_1 \ . \qquad (14)$$

In going from x_0 to a point x, the total shear deformation of the junction and the shear stress on it are

$$\Delta(x,x_0) = -\int_{x_0}^{x}\left(\xi+2\frac{\partial u}{\partial x}\right)_{x_1} dx_1 = \xi(x_0-x)+2[u(x_0)-u(x)] \ , \qquad (15)$$

$$\tau = \tau_{max}\left\{1-\exp[-\Delta(x,x_0)/cd_j]\right\} \ . \qquad (16)$$

Now, of all the junctions passing through x, only a few will have originated in the neighborhood of x_0; others will have originated at all points between x and the leading edge, at x = a. Assuming that the expected value of the density of junctions at any point x is the same as in the static case, we may write

$$\eta(x) = kp(x) \ , \qquad (17)$$

where η = density of junctions (no./in.^{2}), p = pressure, and k = constant. The number of junctions crossing a point x_0 per unit time is $V\eta(x_0)$. The number crossing $x_0 + dx_0$ is $V\eta(x_0 + dx_0)$; thus the rate at which junctions are produced in dx_0 is $Vd\eta = V\left(\frac{\partial \eta}{\partial x}\right)_{x_0} dx_0$.

In our statistical averaging of the stress at x, we must weight the stress due to junctions starting near x_0 by the factor $(d\eta)_{x_0}/\eta(x)$, (thus accounting for the probability of a random junction passing through x having originated near x_0) and by the factor $\pi d_j^2\eta(x)/4$ (accounting for the fact that stress at x is

nonzero for only this fraction of the total time). The stress at x may be written, with the help of Eqs. 16 and 17, as:

$$\tau = \frac{\pi}{4} \int_a^x \tau_{max} \left\{1-\exp\left[-\frac{\Delta(x,x_0)}{cd_j}\right]\right\} kd_j^2 \left(\frac{\partial p}{\partial x}\right)_{x_0} dx_0 \quad . \qquad (18)$$

The quantity $\frac{\pi}{4} kd_j^2 \tau_{max}$ is dimensionless. It can be identified with the macroscopic friction coefficient f by noting that when gross slip occurs, $\Delta(x,x_0) \to \infty$, and $\tau/p \equiv f \to \frac{\pi}{4} kd_j^2 \tau_{max}$. Thus,

$$f \equiv \frac{\pi}{4} kd_j^2 \tau_{max} \quad .$$

Equation 18 may then be written:

$$\tau = f \int_a^x \left(\frac{\partial p}{\partial x}\right)_{x_0} \left\{1-\exp\left[-\frac{\Delta(x,x_0)}{cd_j}\right]\right\} dx_0 \quad . \qquad (19)$$

This equation is good so long as $(\partial p/\partial x)_x$ is negative, i.e., for $0 \leq x \leq a$.

In the region $-a \leq x \leq 0$, we postulate that some of the junctions that cross the point $x = 0$ are destroyed by the decrease in pressure along the path of the junctions; that is, the equilibrium density of junctions at $x < 0$ is less than at $x = 0$, since $p(x<0) < p(0)$. Specifically, we assume that at a point $x < 0$, all junctions formed in the region $(0,|x|)$ are broken off. Thus, the stress in the region $-a \leq x \leq 0$ may be written as:

$$\tau = f \int_a^{|x|} \left(\frac{\partial p}{\partial x}\right)_{x_0} \left\{1-\exp\left[-\frac{\Delta(x,x_0)}{cd_j}\right]\right\} dx_0 \quad . \qquad (20)$$

This equation for the stress includes Eq. 18 as a special case, and is thus valid for all x.

The expression for the tangential displacement that enters into the expression for $\Delta(x,x_0)$ in Eq. 15 is the usual one [*11*] for the displacements at the surface of a semi-infinite solid due to a surface stress distribution:

$$u(x) = \int_{-a}^{a} \frac{2(1-\nu^2)}{\pi E} \tau(x_0) \ell n|x-x_0| dx_0 \ . \tag{21}$$

Before attempting to solve these equations, we cast them into dimensionless form:

$$\Delta^*(x^*,x_0^*) = \xi^*(x_0^*-x^*) + \frac{2}{D}[u^*(x_0^*)-u^*(x^*)] \ , \tag{22}$$

$$u^*(x^*) = \int_{-1}^{1} \tau^*(x_0^*) \ell n|x_0^*-x^*| dx_0^* \ , \tag{23}$$

$$\tau^*(x^*) = \int_1^{|x^*|} \left(\frac{\partial p^*}{\partial x^*}\right)_{x_0^*} \{1-\exp[-\Delta^*(x^*,x_0^*)]\} dx_0^* \ , \tag{24}$$

where

$$x^* \equiv x/a \ , \qquad V_s^* \equiv V_s/V \ , \qquad p^* \equiv p/p(0) \ ,$$

$$\tau^* \equiv \tau/fp(0) \ , \quad u^* \equiv Du/d_j c \ , \quad \Delta^* \equiv \Delta/cd_j \ ,$$

$$\xi^* \equiv \xi a/cd_j \ , \quad \text{and} \quad D \equiv \pi E c d_j/[2fp(0)a(1-\nu^2)] \ . \tag{25}$$

D is a parameter that indicates the effects of surface roughness; the smaller the value of D, the more nearly do the contacting surfaces behave as though they were ideally smooth.

The procedure for solving Eqs. 22 through 24 is as follows. Suppose values of ξ, f, p(0), a and D have been specified. (Methods for estimating the value of d_j are discussed in Refs. 8 and 20.)

1. Assume some $\tau^* = \tau_1^*(x)$.
2. Calculate $u^* = u_1^*(x)$ from Eq. 23.
3. Calculate $\Delta^* = \Delta_1^*$ from Eq. 22.
4. Calculate $\tau^* = \tau_2^*(x)$ from Eq. 24.
5. Continue the iteration until there is sufficient convergence.

The total tangential load is obtained by integrating the shear stress over the contact region:

$$T = \int_{-a}^{a} \tau(x)dx \ . \qquad (26)$$

The iterative procedure described above was carried out, for 4 values of the parameter D, on an IBM 360/65 computer. Convergence of the shear stresses to four significant figures was usually obtained in four iterations. The computation time for this iterative process was about 20 sec. The contact region was divided into 40 segments for evaluation of the integrals. Some curves obtained for ξ as a function of the loads are shown in Fig. 2.4. Also included in Fig. 2.4 is the curve obtained from the smooth-surface theory (Eq. 10). It may be seen that, as the value of D decreases, the rough-surface theoretical curve approaches the smooth-surface theoretical curve.

The significance of the parameter D may be made clearer by considering the case when the tangential load T approaches the value fN. If we assume the smooth-surface theory to be valid, the relative displacement at some point x in the locked region of surface points that came together at x = a is, from Eq. 15,

$$\Delta(x,a) = 0 = \xi(a-x) + 2[u(a)-u(x)]_{smooth} \ . \qquad (27)$$

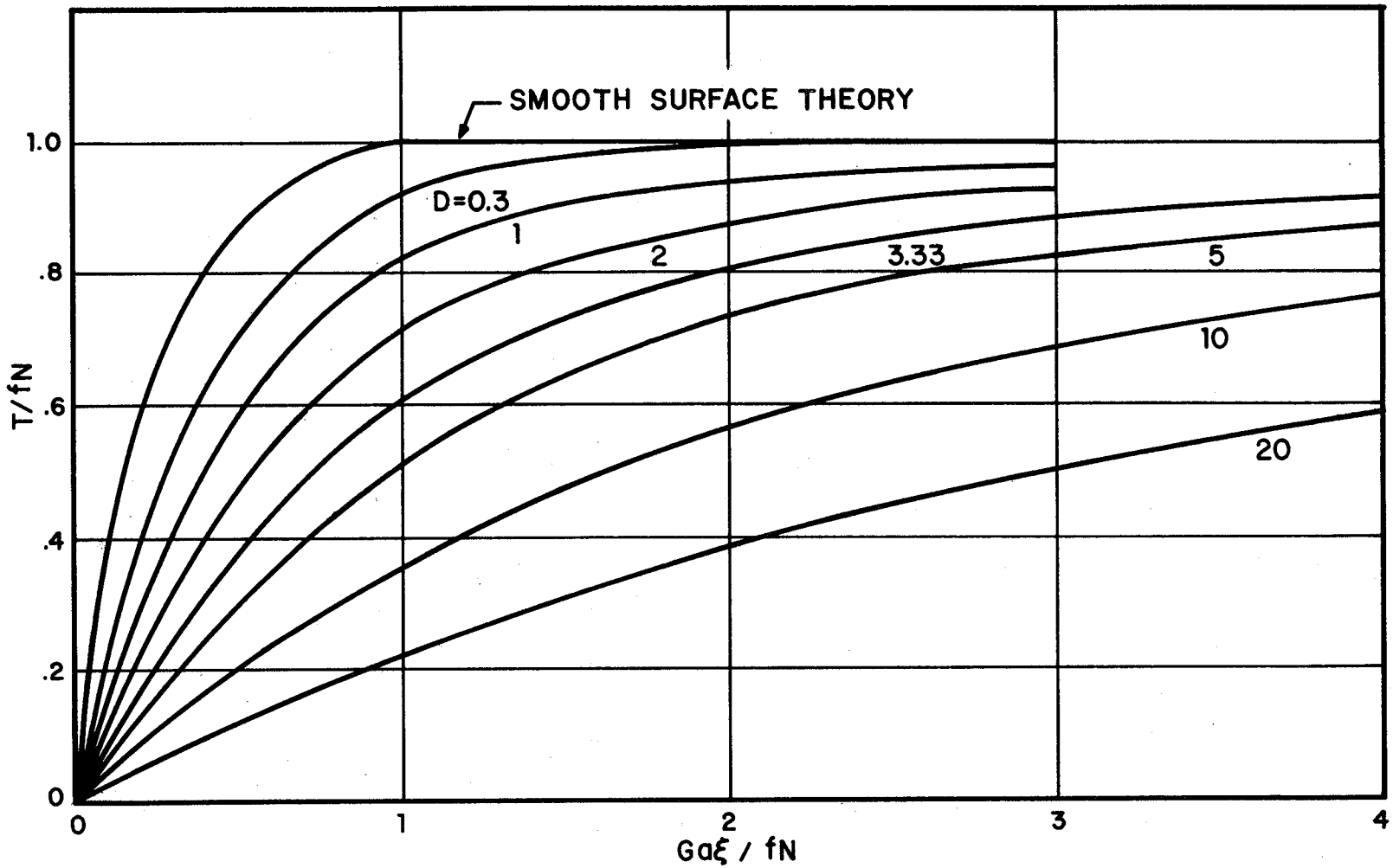


FIG. 2.4 THEORETICAL PREDICTION OF CREEP ACCORDING TO THE ROUGH-SURFACE THEORY. D IS A ROUGHNESS PARAMETER.

The stress at x is given by Eq. 8. If we take the alternate view, and say that this stress must be caused by junction deformation, we can see from Eq. 13 that the junction must have undergone a shear deformation of the order of cd_j:

$$\Delta(x,a) \approx cd_j = \xi(a-x) + 2[u(a)-u(x)]_{rough} . \tag{28}$$

Examining Eqs. 27 and 28, we observe that the rough-surface theory requires an increase in the value of u(a)-u(x), when compared with the smooth-surface theory:

$$[u(a)-u(x)]_{rough} = [u(a)-u(x)]_{smooth} + \frac{1}{2} cd_j . \tag{29}$$

Thus, we may roughly say that the displacements given by the smooth-surface theory are good or bad accordingly, as

$$F \equiv |cd_j/2[u(a)-u(x)]_{smooth}| \tag{30}$$

is small or large. If we look at the point x = 0, we find [3]

$$F = \frac{\pi E cd_j}{4fN(1-\nu^2)} . \tag{31}$$

Obtaining N as a function of p(0) from Eq. 6, we may write

$$F = \frac{Ecd_j}{2fp(0)a(1-\nu^2)} , \tag{32}$$

which differs from D by only a numerical factor of π. Thus, we may say, approximately, that the roughness of the surfaces is important when the size of the friction junctions is comparable to or greater than the magnitude of the tangential elastic displacements in the contact region, and is not important otherwise. In

other words, surface roughness is important either when the surfaces are very rough or when the normal load is light, for either of which $F = D/\pi$ is about 1 or more.

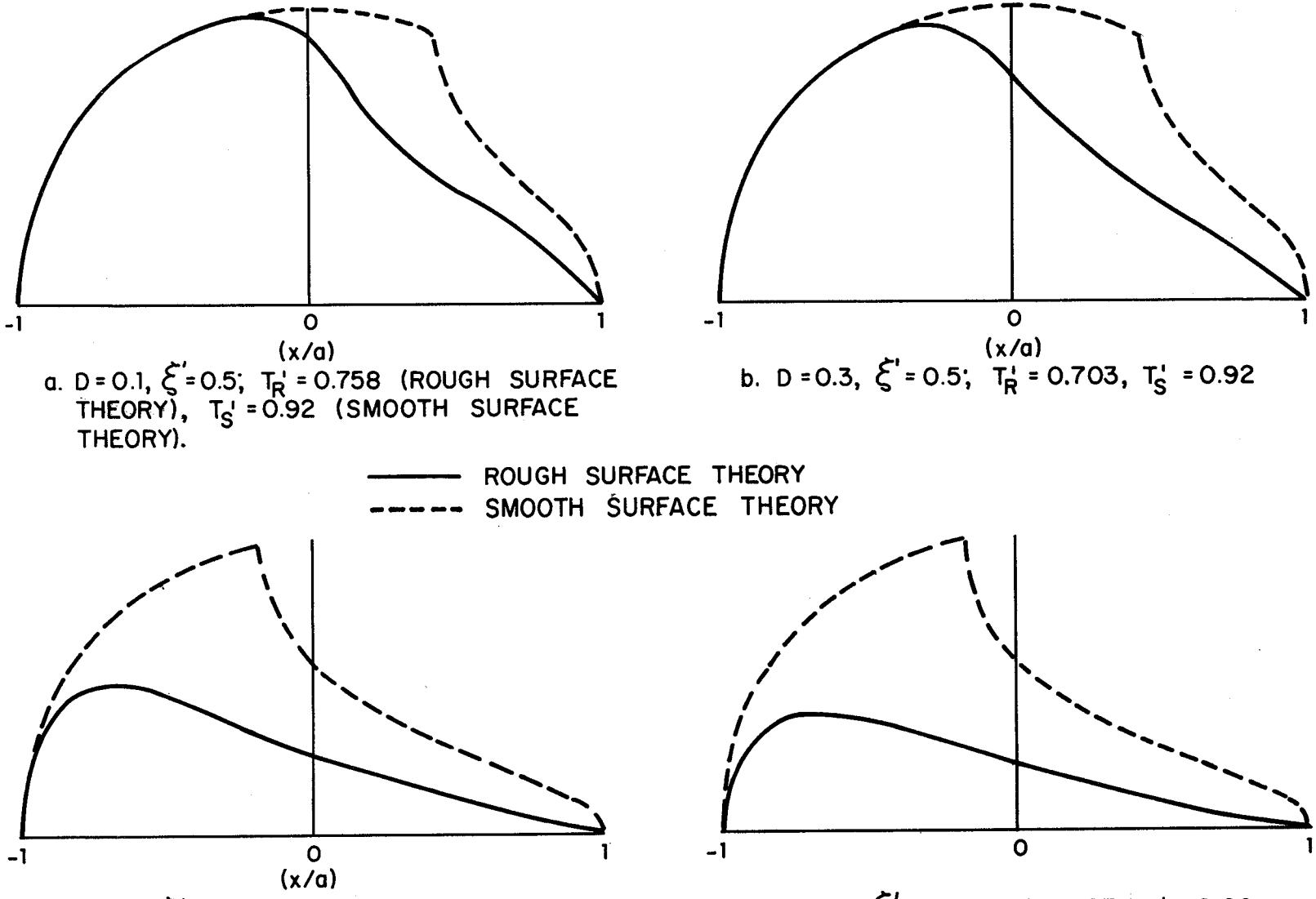
In regard to the stresses and the slip velocities, there are two distinct differences from the smooth-surface theory: (1) the stress distributions given by the rough-surface theory have continuous slopes along the x-axis, whereas the stress distributions given by the smooth-surface theory have a discontinuity in the slope, at the boundary between the locked and slipped regions. (2) More significantly, the rough-surface theory does not separate the contact region into a locked and a slipped region. The interfacial slip velocity is nonzero everywhere, and increases monotonically from the leading edge to the trailing edge. However, as the value of the parameter D grows small, there tends to be a strip adjacent to the leading edge, in which the slip velocity is small. In the limiting case of D tending to zero, this strip becomes the locked region of the smooth-surface theory. Figure 2.5 shows some shear stress distributions obtained from the rough-surface theory. Stress distributions from the smooth-surface theory are shown for comparison.

2.4 Application

In order to assess the significance of rough surface effects in actual physical applications, we shall obtain the value of the roughness parameter D for the rolling contact of a rough cylinder and a smooth half-space.

We assume:

1. That the cylinder and the half-space have the same elastic constants.



a. D = 0.1, $\xi' = 0.5$; $T_R' = 0.758$ (ROUGH SURFACE THEORY), $T_S' = 0.92$ (SMOOTH SURFACE THEORY).

b. D = 0.3, $\xi' = 0.5$; $T_R' = 0.703$, $T_S' = 0.92$

c. D = 1, $\xi' = 0.25$; $T_R' = 0.34$, $T_S' = 0.66$

d. D = 2, $\xi' = 0.25$; $T_R' = 0.27$, $T_S' = 0.66$

FIG. 2.5 SHEAR STRESS DISTRIBUTIONS IN THE INTERFACE ACCORDING TO THE ROUGH-SURFACE THEORY.

2. The contact is elastic.

3. The normal load is sufficiently large so that Hertzian analysis can be used for the normal problem.

4. The rough cylinder has asperities with Gaussian peak height distribution, with zero mean and standard deviation σ.

5. The density of asperities on the cylinder is η_A, and all asperities have the same tip radius β.

Let

η = number of junctions/unit length of cylinder

A = total area of junctions/unit length of cylinder

$E' = E/2(1-\nu^2)$

R = cylinder radius.

Then η is found to be as in Fig. 2.6 [*8*]. The parameter δ is defined as

$$\delta = \frac{8}{3} \frac{\sqrt{R\beta}\ \eta_A \sigma}{\pi^{3/2}} . \tag{33}$$

A representative diameter for the junctions is

$$d_j = \left(\frac{4A}{\pi\eta}\right)^{1/2} . \tag{34}$$

The value of the coefficient c in Eq. 25 may be estimated by considering each junction of diameter d_j to be a Hertzian contact subject to normal and tangential loading. Then the tangential displacement that the junction undergoes before gross slip occurs is [*12*]

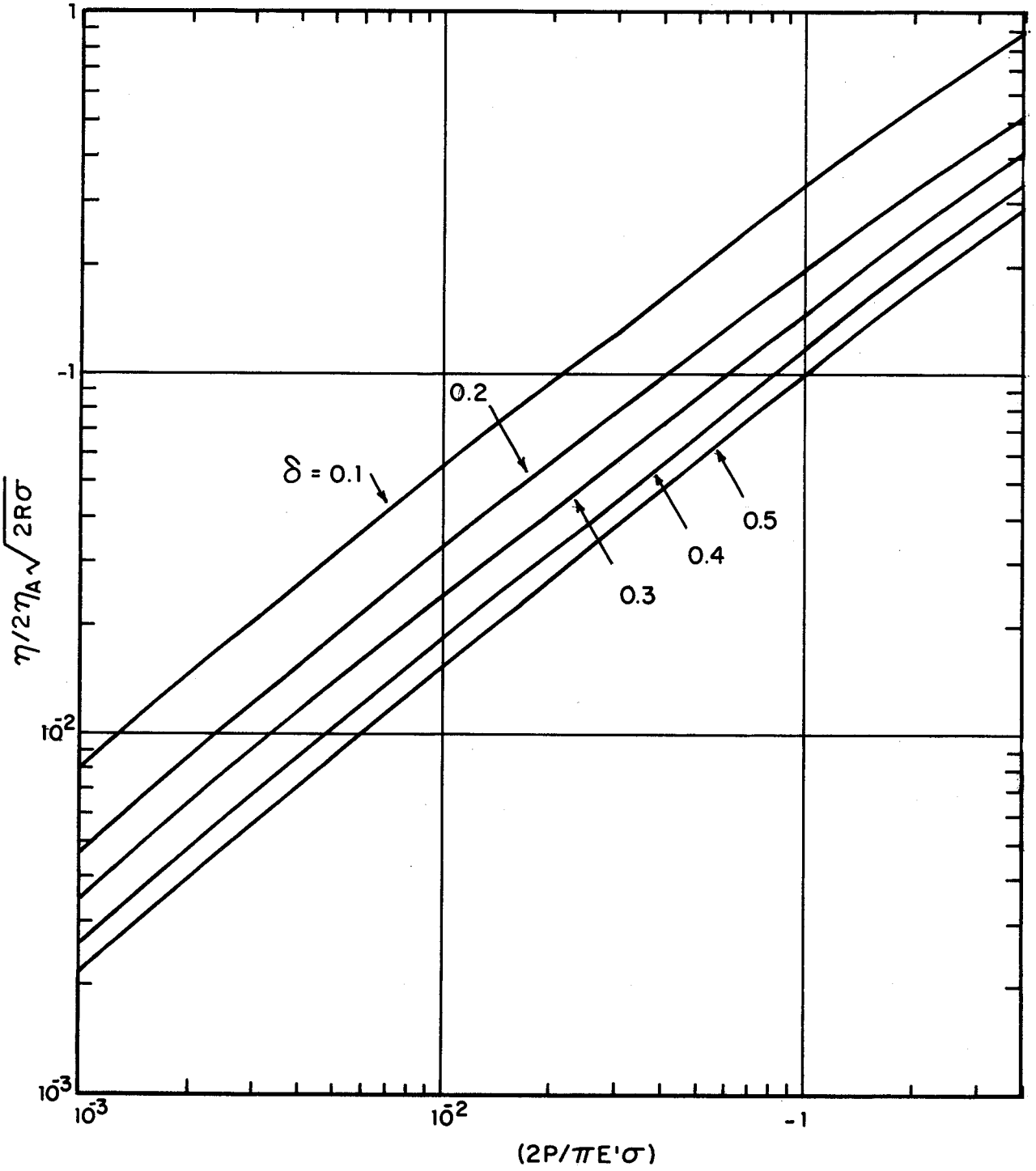


FIG. 2.6 DENSITY OF JUNCTIONS IN THE CONTACT REGION OF A ROUGH CYLINDER AND A PLANE.

$$\Delta = \frac{(2-\nu)\ f' d_j}{16\ \beta\ (1-\nu)}\ d_j\ , \tag{35}$$

where f' is the friction coefficient for each junction. The relationship between f' and the observed overall friction coefficient f is

$$f' = 2af/A\ , \tag{36}$$

where a is the semi-width of the contact region.

Combining Eqs. 35 and 36, the coefficient c (see Eq. 12) may be approximated by

$$c = \frac{af\ (2-\nu)\ d_j}{24\ \beta\ (1-\nu)\ A}\ . \tag{37}$$

We have assumed in Eq. 12 that $\Delta = 3cd_j$ represents gross slip. Finally, Eqs. 5, 6, 25, and 37 may be combined to yield an estimate of the roughness parameter D:

$$D = \frac{(2-\nu)}{3\ (1-\nu)\ \beta\eta} \left\{\frac{\pi RE'}{4N}\right\}^{1/2}\ . \tag{38}$$

The density of junctions η is obtained from Fig. 2.6. Thus the surface roughness parameters necessary for estimating D are β, η_A and σ, all of which may be obtained through profilometry [*13*].

2.5 Numerical Example

For the purposes of numerical calculation, the curves of Fig. 2.6 may be described by a single equation:

$$\eta = .65\ \eta_A\ \sqrt{R}\ (N/\delta E')^{0.75}\ (\sigma)^{-0.25}\ . \tag{39}$$

After introducing the definition of δ, Eq. 33, and combining Eqs. 38 and 39, the following expression is obtained for the parameter D:

$$D \approx \frac{0.26(2-\nu)}{(1-\nu)} \cdot \left(\frac{\sigma}{R}\right) \left(\frac{R}{\beta}\right)^{5/8} \left(\frac{E'R}{N}\right)^{5/4} \Big/ (\eta_A R^2)^{1/4} \quad . \qquad (40)$$

This equation illustrates the effect of various parameters on D: an increase in σ, β or R or a decrease in η_A or N (while all the others are held constant) causes an increase in D. The strongest influence on D is that of σ and N. The variation of D with R mentioned, however, is misleading. An increase in R causes a decrease in the maximum pressure p(0). To determine the effect of R on D while p(0) is held constant, Eqs. 5 and 6 may be combined with Eq. 40 to give

$$D \approx \frac{.062(2-\nu)}{(1-\nu)} \left(\frac{\sigma}{R}\right) \left(\frac{R}{\beta}\right)^{5/8} [E'/p(0)]^{5/2} \Big/ (\eta_A R^2)^{1/4} \quad . \qquad (41)$$

It is now clear that with p(0) constant, an increase in R causes a decrease in D.

Equation 41 may be used to determine the value of p(0) above which the smooth-surface theory is valid. If we make the approximate judgment from Fig. 2.5 that this is so if $D \leq 0.3$, Eq. 41 yields

$$p(0) \geq \left[\frac{0.2(2-\nu)}{(1-\nu)} \frac{\sigma}{R}\right]^{2/5} \left(\frac{R}{\beta}\right)^{1/4} E'/(\eta_A R^2)^{1/10} \quad . \qquad (42)$$

In other words, if inequality (41) is satisfied, the smooth-surface theory should yield good results.

As an example, consider a cylinder with the parameters

$$R = 1 \text{ in.}, \quad E' = 1.65 \times 10^7 \text{ psi}, \quad \eta_A = 3 \times 10^5/\text{in.}^2,$$

$$\beta = 1.6 \times 10^{-3} \text{ in.}, \quad \sigma = 2 \times 10^{-5} \text{ in. and } \nu = 0.3.$$

Then inequality (41) becomes $p(0) \geq 1.56 \times 10^5$ psi. Thus, a fairly high pressure is necessary for the smooth-surface theory to be valid. It may be noted, however, that the surface considered is not particularly smooth. The quantity σ is the standard-deviation of the height distribution of surface asperities. The standard-deviation of the surface height will, in general, be greater than this, see Chap. 5.

As another example, consider a cylinder with R = 16 in., all other parameters being the same as above. In this case, inequality (41) becomes

$$p(0) \geq 5.9 \times 10^4 \text{ psi} .$$

Thus, for large cylinders (such as railway wheels), operating at usual pressures [$p(0) \approx 2 \times 10^5$ psi], rough surface effects will be unimportant, unless the surface roughness is very high.

2.6 Experimental Verification

An experimental verification of a theory such as the one outlined above must involve creep measurements. We know of no technique for measuring stresses or slip velocities in the interface, and the photoelastic techniques of Haines and Ollerton [*13,14*] and of Poon [*15*] implicitly assume that the surfaces are smooth.

In our measurements of creep, the surface roughness was held more-or-less constant. The surfaces were artificially roughened

with No. 100 emery cloth, but no profilometer measurements were made. The surfaces were run in before creep measurements were made. Measurements of creep were made at varying values of the tangential load T, for a constant normal load N. The value of the parameter D was then changed by changing the normal load.

The experimental apparatus is illustrated in Fig. 2.7. The rolling element is an aluminum cylinder, the contacting surfaces being two 1/8 in. wide strips, one at each end of the cylinder. The rolling radius of these strips is 1 in. Two stub axles project from the flat ends of the cylinder. A restraining yoke held in the spindle of the milling machine presses against ball bearings on the axles, thus preventing the cylinder from moving along the aluminum plate surfaces on which it rests. The aluminum plate – 5/8 in. thick – is clamped to the table of the milling machine. The table is made to move by hand feed, the distance traversed being measured on the milling machine's micrometer. The tangential load is applied by means of a nylon string glued to the center of the cylindrical surface. The string passes over a pulley and is then loaded with weights. The normal load is applied by the restraining yoke, which is made in two pieces that can slide relative to each other. Two springs between the pieces provide the compressive force, which is transmitted to the axles of the cylinder. The normal load can be varied by changing the compression of the springs; this is accomplished merely by raising the milling machine table.

The creep was measured as follows. A little flat plate was glued on to one of the flat ends of the cylinder, near the rim. A reading was taken on the plate with a dial gauge standing on a flat reference surface. After several traverses of the milling machine table, a reading was again taken on the plate. Then, for

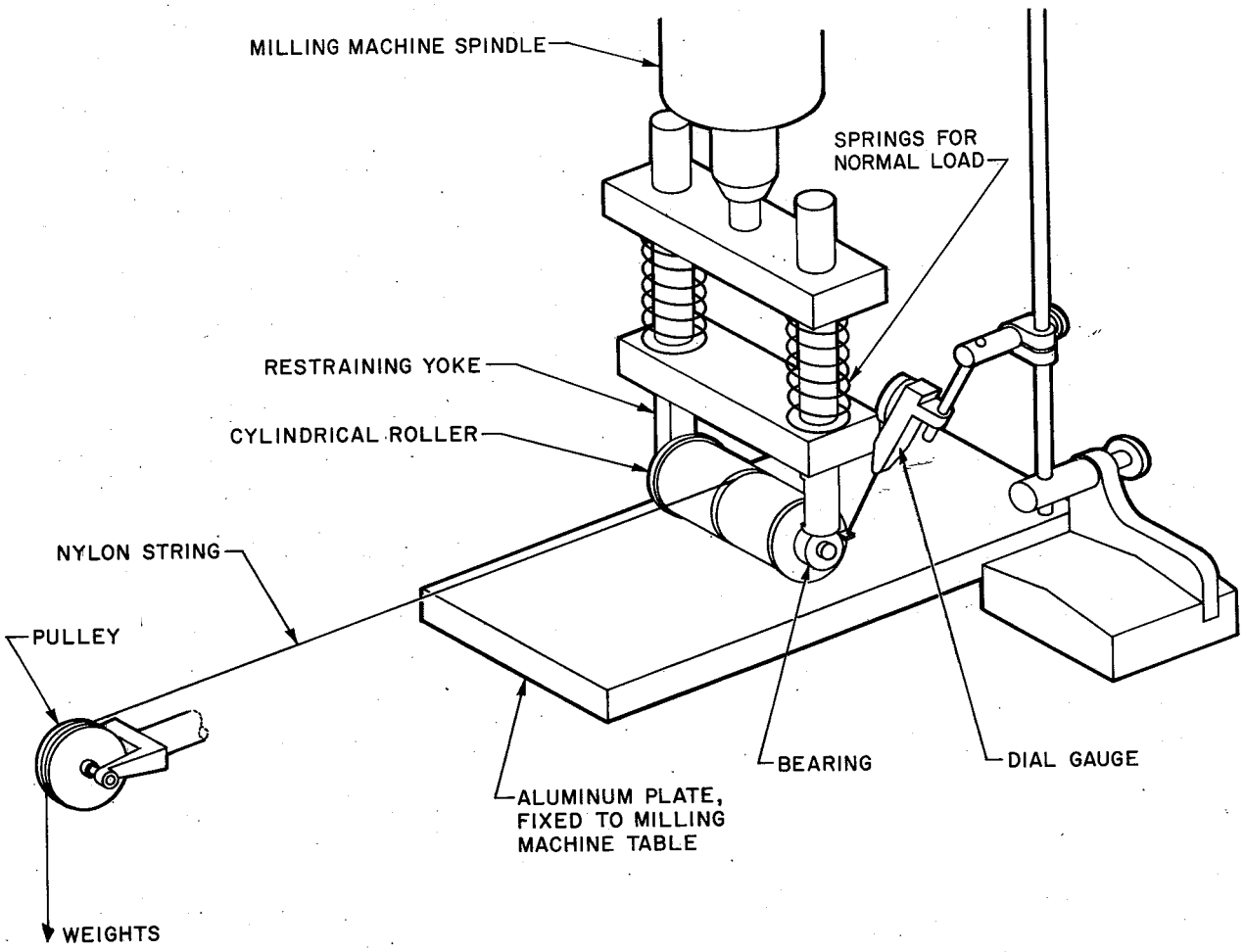


FIG. 2.7 EXPERIMENTAL APPARATUS FOR MEASUREMENT OF CREEP.

small amounts of slip, the difference between the final and the initial readings on the dial gauge gives the distance slipped directly.

Some properties of the aluminum used are:

$E = 3.8 \times 10^6$ psi, $\nu = 0.33$, σ_y = compressive yield stress $\approx$ 25,000 psi.

The results of the experiment are shown in Fig. 2.8. Some useful data concerning the three curves of Fig. 2.8 are presented in Table 2.1

TABLE 2.1

	Curve 1	Curve 2	Curve 3 of Fig. 2.8
Normal load N lb/in.	3.7	76	400
Friction coefficient f	.172	.189	.179
Contact width 2a, in.	3.04×10^{-3}	13.8×10^{-3}	31.6×10^{-3}
Maximum pressure, psi	1500	7000	16,200

At high values of the tangential load, a considerable amount of scatter was obtained in the creep measurements. This we attribute to time-variations in the friction coefficient, leading to instantaneous values of T/fN greater than 1, and to gross slipping. The choice of the value of f used to normalize the data is not easy to justify. We defined f as being given by

$$f = T_0/N \ , \tag{43}$$

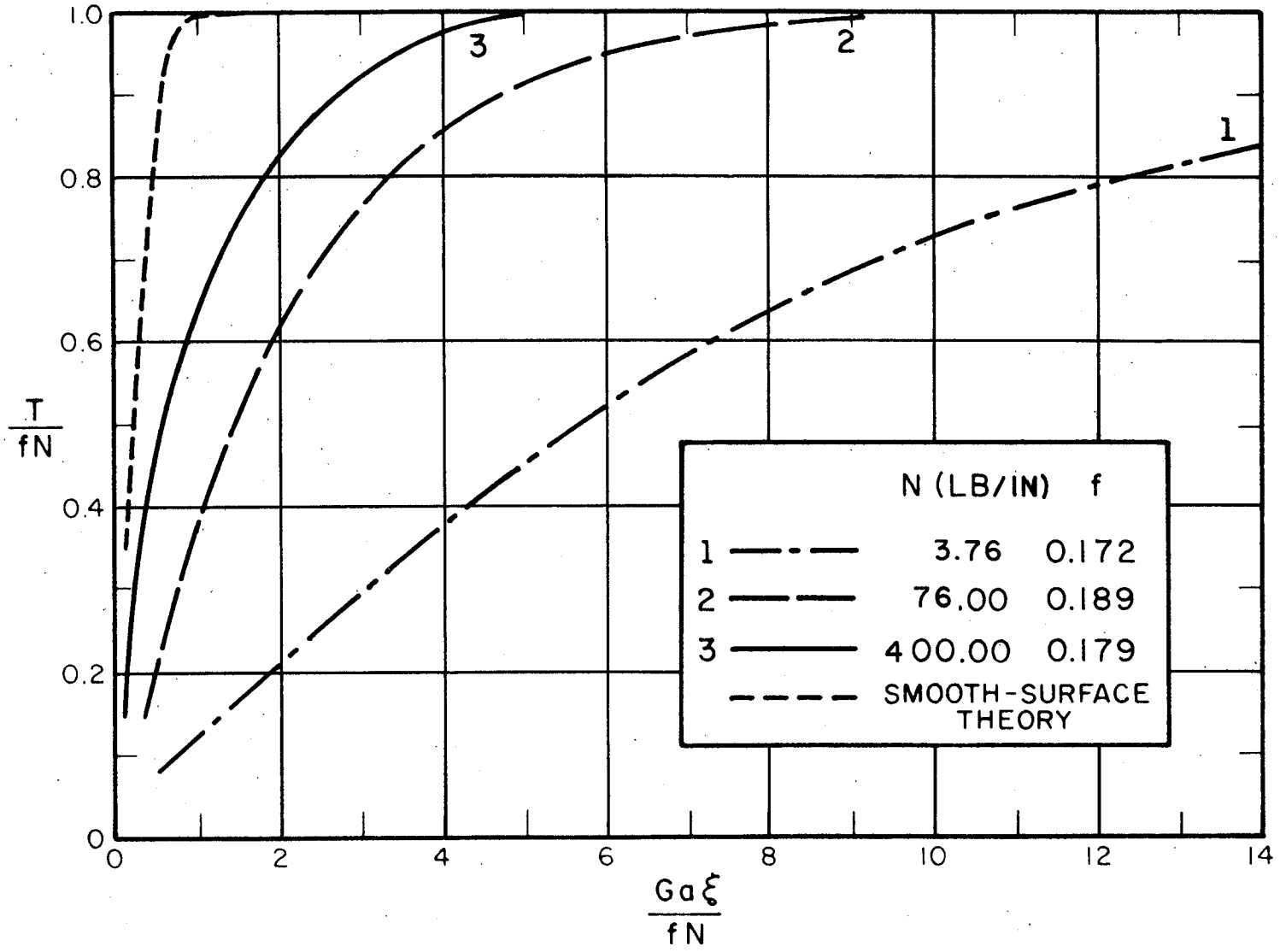


FIG.2-8 EXPERIMENTALLY MEASURED VALUES OF CREEP

where T_0 is the tangential load at which large scatter is first obtained.

The contacting surfaces were regularly cleaned with trichlorethylene. Measurements were made with all obvious nearby sources of vibration turned off, to make sure that there were no anomalous effects due to vibration, such as are observed in sliding contact [*16,17*].

We made certain that contact was being made over the whole width of the two strips on the cylinder by inserting a thin sheet of carbon paper between the cylinder and the plate, and noting that the tracks left by the cylinder were both 1/8 in. wide and of approximately uniform darkness across their width.

The curves of Fig. 2.8 seem to provide qualitative support of the rough-surface theory. Increasing normal loads should — and do — bring experimental measurements and the smooth-surface theory in closer agreement, due to a decrease in the parameter D.

In addition to these low-velocity, vibration-free creep studies, a study was made of the effects of surface vibration on creep. The apparatus used was that shown in Fig. 2.7, with the plate mounted in such a way as to permit significant bending vibrations. The plate was made to vibrate by attaching to it an electromagnetic shaker. The input to the shaker was a sinusoidal signal, amplified by a power amplifier. The amplitude of the acceleration due to the vibration was measured with an accelerometer at an arbitrary point on the plate. We did not try to differentiate between vertical and horizontal vibration, as the apparatus was not sophisticated enough to permit this. Both the frequency and the amplitude of the applied vibration were changed. The results of the experiment are shown in Fig. 2.9. The loads and some other relevant parameters had the following values:

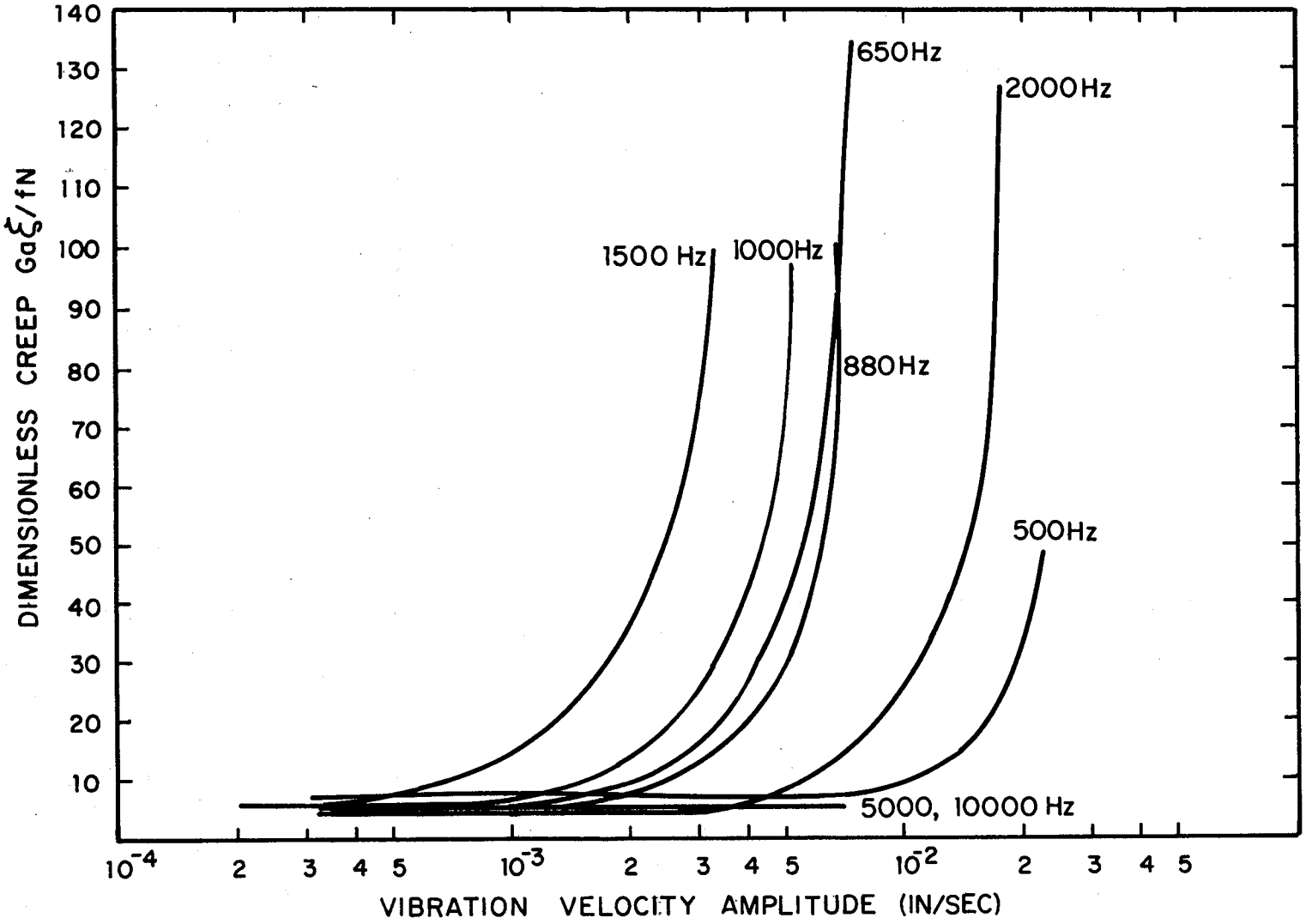


FIG. 2.9 EFFECT OF SURFACE VIBRATION AND AMPLITUDE ON CREEP.

Normal Load N = 3.54 lb/in.

Tangential Load T = 0.275 lb/in.

Half-Width of Contact Region a = 1.5×10^{-3} in.

Young's Modulus E = 3.8×10^{6} psi

Material: Aluminum

Cylinder Radius r = 0.996 in.

On the vertical axis in Fig. 2.9, we have plotted the dimensionless creep, defined as

$$\xi^* = \frac{Ga\xi}{fN} ,$$

where

$$\xi = \text{longitudinal creep} = \frac{\text{distance slipped}}{\text{distance rolled}}$$

$$f = \text{friction coefficient.}$$

The friction coefficient was assumed to have a constant value of f = 0.137. The normal and tangential loads were held constant during the entire experiment. On the horizontal axis, we have plotted the amplitude of the plate *velocity*, calculated from the measured plate acceleration.

We checked to make sure there was no loss of contact between the cylinder and the plate by completing an electrical circuit between them, and making sure that both the plate and roller were always at the same voltage.

At large values of the velocity amplitude, where ξ^* increases rapidly, there is considerable scatter in the data. We have omitted this for purposes of clarity and because we feel that there is, nevertheless, a significant effect.

One possible mechanism which might cause the effect we have observed is that the effect of surface vibrations is to cause a decrease in the effective coefficient of friction f. Thus, in the presence of vibration, the friction coefficient to be used in calculating the dimensionless tangential load $T^* = T/fN$ may be much smaller than f = 0.137. This would increase T^*, and also the theoretically predicted value for ξ^*. For example, with T = 0.275 and N = 3.54 lb/in. and f = 0.137, we get $T^* = 0.56$. However, if the effect of vibration is to reduce f to .07, we would get $T^* = 1.1 > 1.0$. This would mean that there would be gross slipping, which would show up in our measurements as large creep, with a lot of scatter. A decrease in friction due to surface vibrations has been observed frequently, as mentioned in Chap. 1.

Another mechanism that is sometimes suggested involves the concept of contact resonance. The Hertzian contact itself acts as a spring. There is thus a "contact resonance frequency", dependent on the mass of the contacting elements. If this contact resonance frequency is excited in some way, it is likely that large surface deflections will take place. These deflections lead to vibrations of the surface asperities, which may result in a decrease in the friction coefficient. If this mechanism were the correct one, then a large decrease in the friction coefficient — or in our experiment, a large increase in the creep — should result. Figure 2.10 shows the measured dimensionless creep as a function of the excitation frequency, at constant vibration velocity amplitude. It must be noted that the velocity is measured at a point distant from the contact region; it may be significantly different from the amplitude at the contact. It may be seen from Fig. 2.10 that there is a large increase in the creep at 1500 Hz. We now proceed to a quantitative examination of the concept of contact resonance, in order to see whether the measured resonance frequency is near the contact resonance frequency.

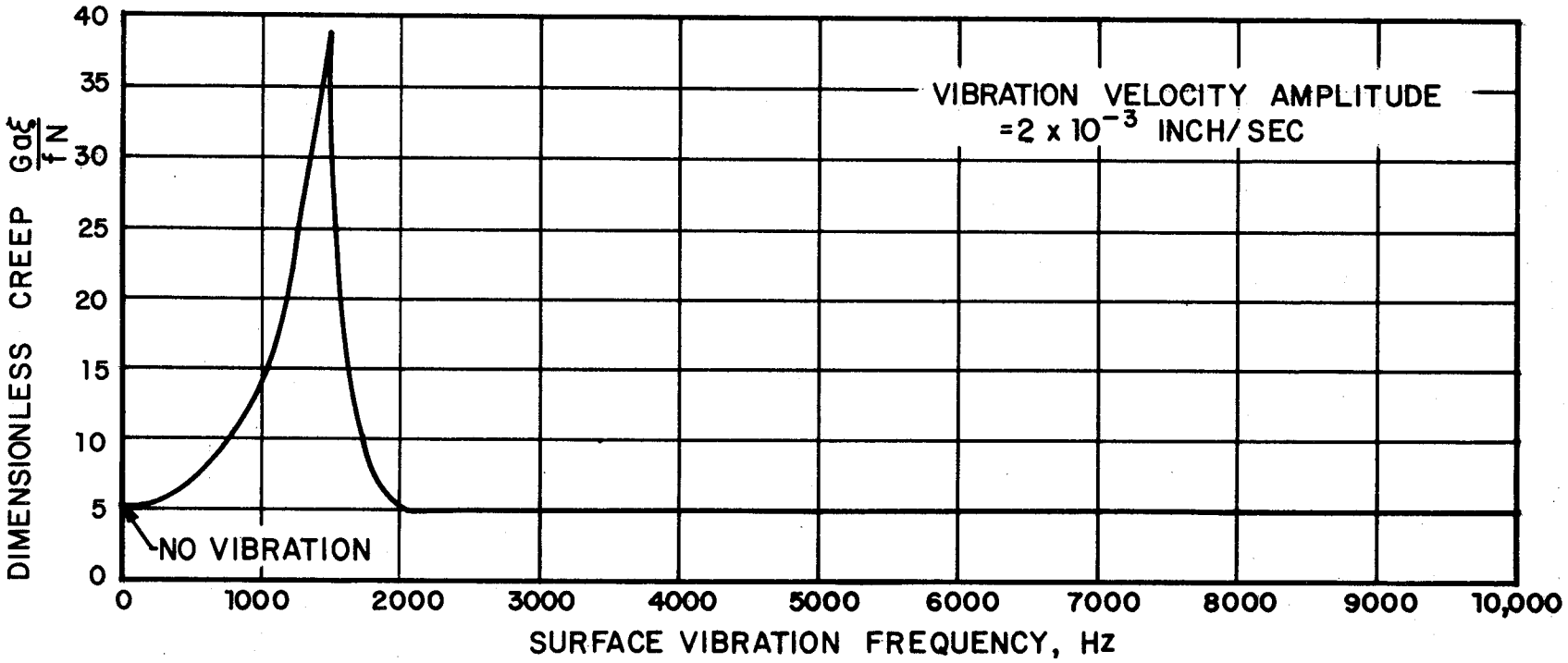


FIG. 2.10 VARIATION IN CREEP WITH FREQUENCY OF VIBRATION

Consider an elastic cylinder pressed against a rigid plane. A normal force

$$F = F_m + F_1 \sin\omega t \tag{44}$$

per unit length is applied at the axis of the sphere. If the deflection of the axis relative to the plane is u, the equation of motion of the cylinder center is, following Timeshenko [*11*],

$$m_c \ddot{u} = F - P \ , \tag{45}$$

where m_c is the cylinder mass per unit length and P is the force at the contact per unit length. Assuming that no structural resonances are excited, the force P is given by the Hertzian theory of contact [*3*]

$$P = Eu/2(1-\nu^2) \ . \tag{46}$$

Combining Eqs. 44-46, we obtain

$$m_c \ddot{u} + k_c u = F_m + F_1 \sin\omega t \ , \tag{47}$$

where

$$k_c = E/2(1-\nu^2) \tag{48}$$

is the "contact stiffness". The contact resonance frequency is

$$f_c = \frac{1}{2\pi} (k_c/m_c)^{1/2} \ . \tag{49}$$

If the excitation frequency ω were to coincide with f_c, large deflections would result, their amplitude being limited by frictional interface damping, internal damping in the contacting elements, or by loss of contact.

by a rail irregularity wavelength of about 1 ft. It is thus important to determine track roughness spectra in the region of wavelengths from a few inches to a few feet.

2.7 Discussion

A tentative theory has been presented which might account for some of the discrepancies observed between the smooth-surface theory of rolling contact and experimental data. However, these discrepancies are explained only in those cases where, roughly speaking (no pun intended!), the normal load is low, or the surface roughness is high. The theory also explains why Halling and Brothers [*18*], using steel balls on a steel surface, found surface roughness to have no effect on their creep measurements. We may note that the smallest normal load they used gave a maximum pressure of $p(0) = 168{,}000$ psi. This might presumably have led to fairly low values of D for even their lowest load, in which case no surface roughness effects would be noticed. A similar explanation might account for the very good agreement that Johnson [*19,20*] found between the smooth-surface theory and experiment.

From a purely computational point of view, smooth surfaces may be approximated by surfaces with a small value of D, leading to a far simpler numerical scheme than does the smooth-surface theory [*21*], even for three-dimensional problems [*22*]. In three-dimensional problems involving spin, however, additional complexities are introduced by the necessity of considering nonmonotonic deformation of the friction junctions as they traverse the contact region. Models for force-deformation laws for junctions undergoing reversed deformation have been proposed [*2*], and could, presumably, be used in such problems. This possibility has not been explored.

For our experimental conditions, we find $k_c \approx 2.1 \times 10^6$ lbf/in.^{2}, $m_c = 4$ lbm/in. Then, using appropriate units, we find from Eq. 49 that $f_c \approx 2260$ Hz. If we now account for the fact that the plane surface is not rigid (thus halving the contact stiffness) and that part of the plate mass is also involved in the contact vibrations, we obtain a calculated frequency that is quite close to the observed resonance frequency of 1500 Hz.

Having shown that contact resonance is, indeed, an important phenomenon, it now remains to determine what the effect of this resonance is on the friction characteristics of the contact. The analysis of Chap. 4 seems to indicate that at the high frequencies at which this resonance occurs, the effective reduction in the normal (or "adhesive") load due to dynamic load effects is insignificant. Thus it appears that the effect of the contact resonance vibration must be to change the behavior of the friction junctions in the interface. This possibility needs to be investigated in detail.

Another problem that needs to be studied is the effect of broadband random forcing on contact resonance, as opposed to the harmonic excitation studied here.

An attempt was made to study the effects of vibration at high normal loads. No effects were found, mainly due to limitations on the shaker input power. It was concluded that such a study should be deferred to the future, as it necessitates the design and construction of a separate experimental apparatus.

In wheel-rail contact, the contact stiffness will be approximately 10^7 lbf/in., ten times the value for our experiment. However, the participating masses may be as much as 300-400 times as large. Thus the contact frequency may be in the neighborhood of 200-300 Hz. At a speed of 150 mph, this frequency is generated

However, we believe that the major advantage of the theory proposed here is that it uses a more realistic model of the interface than does the smooth surface theory. This makes it easier to incorporate effects such as those of surface chemistry, boundary lubrication and surface vibration into the theory. The potential importance of such effects has been described by Wickens [*23*]. As a particular example, consider the rolling contact of a wheel and a rail when the rail surface has been treated with a plasma torch. It is clear that the effect of the plasma is to cause surface contaminants to dissociate, thereby increasing the adhesion between the contacting surfaces [*24*]. In this case, the friction junctions formed are of the "strong" type. The behavior of these junctions is markedly different from that of weak or lubricated junctions, in that over part of their life cycle, they are in tension instead of compression. It is precisely this behavior that causes the overall friction coefficient to increase. In view of the likelihood that plasma torches will be used in the future, it appears highly desirable that a theoretical and experimental investigation be made of the rolling contact of surfaces cleaned with such torches. This line of research is all the more attractive because of a recent advance in the theory of junction deformation [*25*], which provides most of the inputs necessary for a study of the effect of strong junctions in rolling contact.

Some effects of sinusoidal surface vibration on creep and friction in rolling contact were described above. Further experimental and analytical studies are recommended in order to determine

a) the effect of random vibration, and

b) the importance of these effects in wheel-rail contact.

REFERENCES

1. F.W. Carter, "On the Action of a Locomotive Driving Wheel," *Proc. Roy. Soc.* (London) 112, 151 (1926).

2. F.P. Bowden and D. Tabor, *The Friction and Lubrication of Solids, Vols. I and II*, Oxford University Press, England (1950, 1964).

3. H. Poritsky, "Stresses and Deflections of Cylindrical Bodies in Contact," *J. Appl. Mech., ASME Trans.* 17, 191 (1950).

4. J.A. Greenwood and J.B.P. Williamson, "Contact of Nominally Flat Surfaces," *Proc. Roy. Soc.* (London) 295, 300-319 (1966).

5. J.A. Greenwood, "The Area of Contact of Rough Surfaces and Flats," *J. Lub. Tech., ASME Trans.* 89, 81-91 (1967).

6. J.A. Greenwood and J.H. Tripp, "The Elastic Contact of Rough Spheres," *J. Appl. Mech., ASME Trans.* 89, 153-159 (1967).

7. J.A. Greenwood and J.H. Tripp, "The Contact of Two Nominally Flat Surfaces," University of Liverpool Report, Sept. 1967.

8. C.C. Lo, "Elastic Contact of Rough Cylinders," *Int. J. Mech. Sci.* 11, 105-115 (1969).

9. B.G. Rightmire, "Solid-Solid Interfaces," Unpublished Massachusetts Institute of Technology Notes (1965).

10. J.A. Greenwood and D. Tabor, "Deformation Properties of Junctions," *Proc. Phys. Soc.* (London) 58, 609-619 (1955).

11. S. Timoshenko and J.N. Goodier, *Theory of Elasticity*, McGraw-Hill Book Co., Inc., New York, p. 91 (1951).

12. R.D. Mindlin, W.P. Mason, J.F. Osmer, and H. Deresiewicz, "Effects of an Oscillating Tangential Force on the Contact Surfaces of Elastic Spheres," *Proc. First U.S. National Congress of Applied Mechanics*, 203

13. D.J. Haines and E. Ollerton, "Contact Stress Distribution on Elliptical Contact Surfaces Subject to Radial Forces," *Proc. Inst. Mech. Engrs.* 177, 95-114 (1963).

14. D.J. Haines, "Contact Stress in Flat Elliptical Contact Surfaces which Support Radial and Shearing Forces During Rolling," *Proc. Inst. Mech. Engrs.* 179, 154-165 (1964-1965).

15. S.Y. Poon, "An Experimental Study of the Shear Traction Distribution in Rolling with Spin," *Wear* 10, 61-69 (1967).

16. D.M. Tolstoi, "Significance of the Normal Degree of Freedom and Natural Normal Vibrations in Contact Friction," *Wear* 10, 199-213 (1967).

17. A.M. Mitskevich, "Motion of a Body Over a Tangentially Vibrating Surface, Taking Account of Friction," *Soviet Phys.-Acoust.* 13, 348-351 (1968).

18. J. Halling and B.G. Brothers, "The Effect of Surface Finish on the Creep and Wear of a Rolling Ball Subject to Normal and Tangential Surface Tractions," *Wear* 9, 199-208 (1966).

19. K.L. Johnson, "The Effect of Spin Upon the Rolling Motion of an Elastic Sphere on a Plane," *J. Appl. Mech., ASME Trans.* 80, 332-338 (1958).

20. K.L. Johnson, "The Effect of a Tangential Force Upon the Rolling Motion of an Elastic Sphere on a Plane," *J. Appl. Mech., ASME Trans.* 80, 339-346 (1958).

21. P.R. Nayak, "A New Theory of Rolling Contact," PhD Thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology (April 1968).

22. J.J. Kalker, "On the Rolling Contact of Two Elastic Bodies in the Presence of Dry Friction," Doctoral Dissertation, Technische Hogeschool, Delft, Netherlands (1967).

23. A.H. Wickens, "Vehicle Dynamics and Wheel-Rail Interface Problems," in *High Speed Ground Transportation*, Transportation Research Institute, Carnegie-Mellon University (1969).

24. H.T. Albachten, "Locomotive Wheel-to-Rail Traction," *ASME Paper* No. 66-WA/RR-7 (1966).

25. P.K. Gupta, "Topographic Analysis of Friction Between a Pair of Rough Surfaces," ScD Thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology (September 1970).

3. EXPERIMENTS ON JUNCTION DEFORMATION

Summary

The purpose of the experiments described here was to gain insight into the behavior of friction junctions in the interface between rolling elements. The asperities forming a junction were modeled by two steel spheres sliding past each other. Though certain experimental difficulties were encountered in determining the effects of relative velocity and of superposed vibration, the following conclusions were reached.

1. The measured friction and normal forces (in the direction of sliding and perpendicular to it) at each point in the life of the junction are not proportional to each other.

2. However, the "true" friction and normal forces (tangential and perpendicular to the spherical surfaces) are proportional to each other.

3. During repeated contact of dry asperities, the friction coefficient remains constant for high normal loads, as the number of traversals increases. For lighter loads, on the other hand, the friction coefficient decreases. Loads are classified as heavy or light according as plastic deformation is taking place or not.

4. With the surfaces contaminated by 20-20 motor oil, the friction coefficient is initially much lower than for dry contact. However, with repeated contact (the surfaces being continuously flooded with oil), the friction coefficient increases. This is due to the progressive roughening of the surfaces, and the increasing probability of dry contact. This result underlines the importance of surface roughness when the contacting surfaces are heavily contaminated.

5. As the sliding velocity is increased, the normal force increases and the friction coefficient decreases for dry junctions. The increase in normal force is probably connected with play in the fixtures holding the spheres. There are two mechanisms that can account for the decrease in the friction coefficient:

   a) As the sliding velocity increases, the adhesive strength of the junction decreases, due to the decreased time of interaction.

   b) Assuming elastic deformation, and a proportionality between friction force and real area of contact, the friction coefficient varies as $(\text{Normal Force})^{-1/3}$. As the normal load increases, the friction coefficient decreases.

6. Further experimental studies of the effects of sliding velocity and asperity vibration are desirable.

Based on observations of the behavior of a single junction and on analysis [*10*], the following hypotheses on the *behavior of surfaces* are offered. These hypotheses need to be investigated.

7. Surface roughness effects are least important for boundary-lubricated surfaces (those with an extremely thin film of contaminant), and most important for either very clean surfaces or flooded surfaces.

8. For surfaces with a high degree of adhesion (i.e., resistance to normal separation), the friction coefficient increases as the surface roughness increases.

9. As the rolling velocity of bodies in rolling contact increases, the friction coefficient decreases, not due to a reduction in the normal load due to dynamic load effects, but due to any of the following:

   a) decrease in the adhesive strength of the junctions,

   b) asperity vibrations,

   c) elastohydrodynamic effects.

   In any case, the friction coefficient will depend to some degree on the surface roughness, the dependence being the largest when mechanism (c) operates. With mechanism (b) operating, the effects of asperity vibration will be the smallest when asperity deformation is plastic, i.e., when the surfaces are rough. When mechanism (a) operates, the size of the contact region should be important. Thus, it would be worthwhile to investigate the effects on the friction coefficient of the wheel radius, the normal load on the wheel, and the elastic moduli of the wheel and rail materials.

3.1 Introduction

The investigation of the mechanism of friction and wear between two sliding surfaces has been of great interest to a number of research workers, particularly during the past two decades. In dealing with a given surface, the first problem is to describe the surface in terms of significant parameters which are necessary in any theoretical analysis. From the results of some of the most recent investigations [*10,6,7*], and those reported in Chap. 5, it seems plausible that a given rough surface can be characterized in terms of some statistical properties, e.g., distributions of heights, peaks, radius of curvature at the peaks, etc. Furthermore, an experimental determination of such statistical properties is not very cumbersome using high speed computing machines and modern analog to digital conversion facilities. One such investigation [*7*] shows that the distribution of asperity heights is approximately Gaussian.

Thus, the study of friction and wear may be broken up into two steps, (a) the study of a single junction formed by interaction of two asperities, and (b) an extrapolation from a single junction to the entire surface, knowing the statistical characterization of the surface. One of the basic aims of the present work is to study step (a), i.e., the behavior of a single junction.

Various attempts have been made during the past few years to study the mechanism of junction deformation. A plastic junction was first analyzed by Green [*4*]. The slip line field solution proposed by Green allows the calculation of normal and tangential forces during contact between two wedge shaped asperities. An experimental investigation was carried out with plasticine junction models to demonstrate the mechanism of deformation of the junction.

Experiments simulating the deformation of a junction have been reported by Greenwood and Tabor [5]. Some of these experimental results are in a good qualitative agreement with the analytical results obtained by Green [4].

Recently Edwards and Halling [3] have proposed a solution for the interfacial forces during the complete life cycle of the junction. They consider a slip line field similar to the one suggested by Green and obtain a solution for wedge shaped asperities after determining the junction angles by minimizing the work dissipated in deformation and frictional effects.

The mechanism of the formation and destruction of adhesive junctions has been experimentally studied by Ainbinder and Prancs [1]. Different material combinations, e.g., both materials of same hardness but similar and dissimilar work hardening characteristics, hexagonal materials, etc., have been discussed in this work.

Cocks [2] performed an experimental investigation concerning the initial stages of deformation of a junction during sliding. It is shown that the relative tangential displacement is accommodated by the plastic shearing of the underlying material in a direction slightly inclined to the surfaces, and that the junctions do not shear at the original interface, when adhesion is strong.

Some of the dynamic effects of frictional contacts have been investigated by Tolstoi [9]. He experimentally demonstrates the connection between a friction-velocity curve with negative slope, frictional self-excited vibrations, and the freedom of normal displacement of sliding bodies. The conditions for the self-excited vibrations to vanish are stated. A reduction in friction due to microvibrations of the slider (due to asperity collisions) under certain conditions is demonstrated experimentally.

The main objectives of the present work were to study in terms of the following:

1. The effect of repeated sliding of two asperities on the ultimate frictional behavior of the junction formed by their interaction.
2. The effect of sliding velocity on friction produced in case of asperity interaction during sliding.
3. The influence of the presence of a contaminant at the interface.
4. The importance of normal vibrations.

3.2 Experimental Procedure

A schematic diagram of the experimental set-up used for all junction experiments is shown in Fig. 3.1. The behavior of a weak junction was simulated by sliding two steel balls past each other. One of the steel balls was fastened to the column of a milling machine, while the other was held in a vee-block which was fixed on a 3-D dynamometer, which in turn was bolted on the milling machine table. Thus by traversing the table, it was possible to slide one sphere past the other at a fixed velocity.

The contact forces during sliding were measured with a 3-D dynamometer. Contact forces in the direction of sliding, transverse to it, and perpendicular to it were measured. By adjusting the height of the center of the spheres in such a way that both centers lay on a line perpendicular to the direction of sliding, it was possible to eliminate the contact force in the transverse direction. Thus only two force components, both lying in the vertical plane, were measured. These two component forces represent the normal and friction force during the contact.

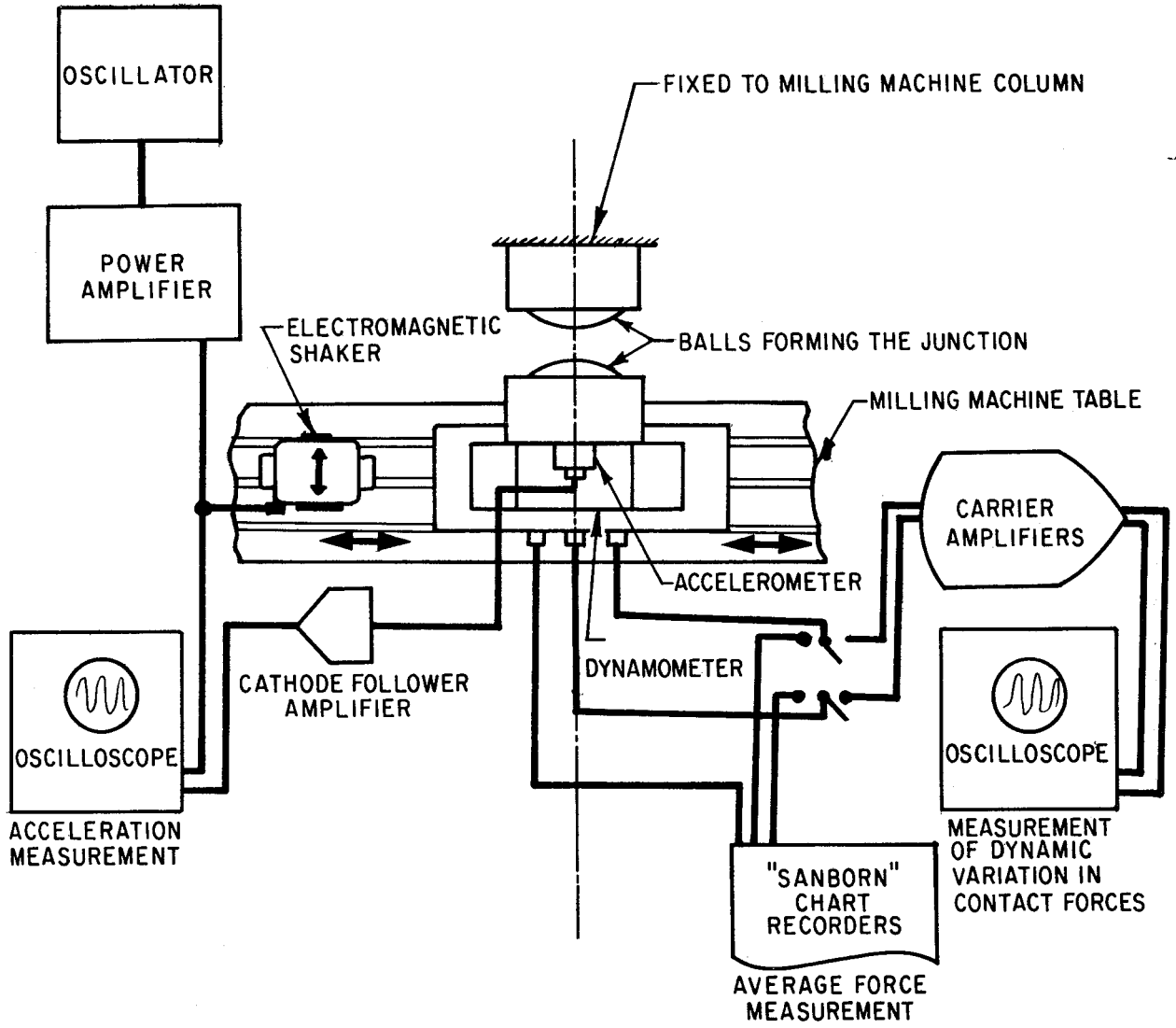


FIG. 3-1 SCHEMATIC DIAGRAM OF THE EXPERIMENTAL SET-UP USED FOR JUNCTION STUDIES.

A "Sanborn" chart recorder was used to record the magnitude of both normal and friction forces during contact. The force traces so obtained give the magnitude of the forces transmitted at any point in the life of the junction.

A suitable electrical circuit was also set up to display the high frequency variation in the contact forces. This was done by using two carrier amplifiers, the output of which was displayed on an oscilloscope, as shown in Fig. 3.1. A band-pass filter (not shown in Fig. 3.1) was sometimes used to observe components of a certain frequency.

The sliding velocity was changed by changing the traversing speed of the milling machine table. At any setting of transverse feed, the time required for a certain length of traverse was electrically measured by installing a microswitch (not shown in Fig. 3.1) on the lead screw of the machine. The velocity measurements are precise to within 1%. As the feed mechanism (motor, gears etc.) was always in operation, there were some inherent vibrations during the contact of the two spheres.

In order to study the effect of superposed vibrations on the junction, an electromagnetic shaker was installed adjacent to the dynamometer, on the milling machine table. It was thus possible to produce vibratory motion in the normal direction. The input to the shaker was from an oscillator, with suitable power amplification.

Vibration levels were measured with an accelerometer installed on the vee-block, holding one of the spheres. A cathode follower amplifier was used to provide suitable impedance matching. The acceleration level was measured directly on an oscilloscope. Assuming a sinusoidal variation in the acceleration, it is possible to determine the displacement amplitude of the vibrations (see Appendix).

The 3-D force dynamometer was calibrated by applying dead weights. Calibrations on both the Sanborn chart recorder and the oscilloscope were obtained. Some details of the calibration set-up are described in the Appendix.

Figure 3.2 shows an overall pictorial view of the experimental set-up used for all junction experiments.

Microscopic examination of the contact surfaces was made with a "Unitron" general purpose optical microscope.

When studying the effect of contaminants, a steady flow of lubricant at the interface was maintained by a suitable siphon arrangement.

3.3 Test Conditions

The interacting asperities forming the junction were two spherical bodies, of equal radius of curvature, with the following specifications:

1. Size: 3/4 in. diameter.
2. Material: Hardened steel ("Hoover" grade 25).
3. Finish: Highly polished, as in ball bearings (about 1μ-in. rms roughness).
4. Environment: The experiments were performed at room temperature. The contact surfaces were frequently cleaned with acetone.
5. Contaminant: When studying the effect of contamination, 20-20 motor oil was used.

Suitable fixtures for holding the balls were designed. The experimental results obtained are summarized in the following Section.

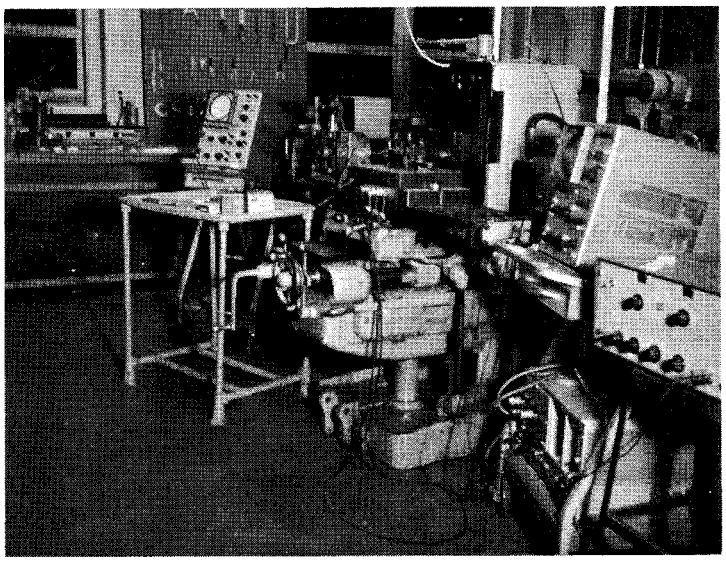
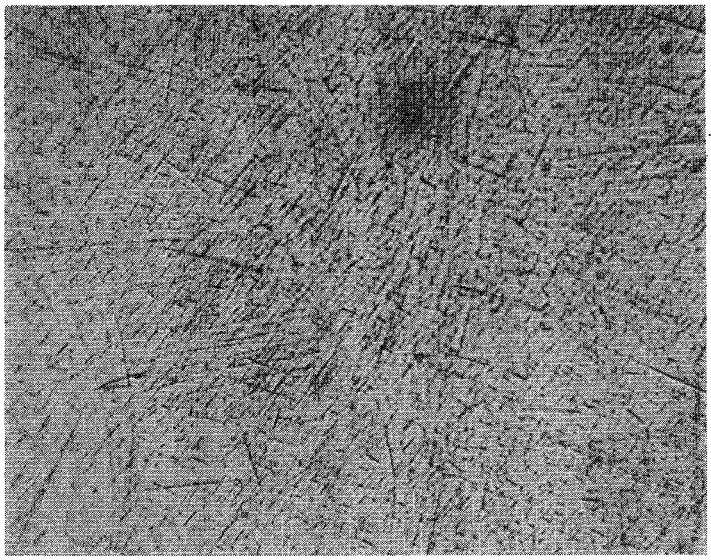


FIG. 3.2 PICTORIAL VIEW OF THE EXPERIMENTAL SET-UP USED FOR JUNCTION STUDY EXPERIMENTS

3.4 Results and Discussion

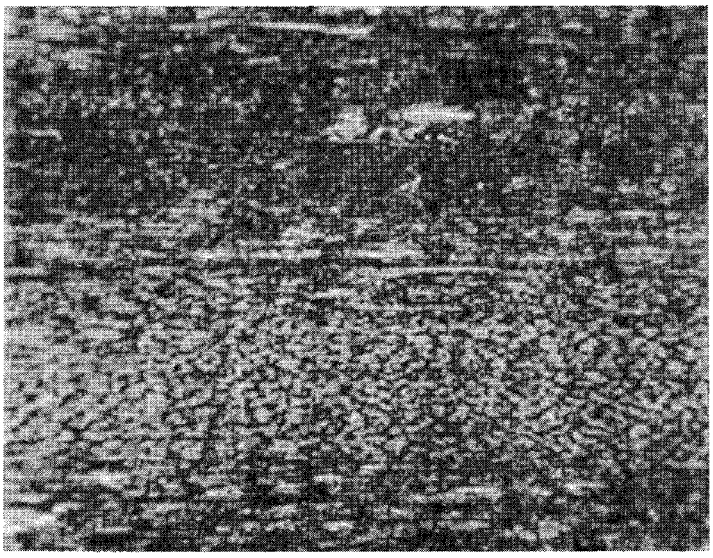
Most of the experiments consisted essentially of a measurement of the friction and normal forces during the life of a junction. Deformation of the contact surface was studied with the help of a microscope. Typical photomicrographs of the spheres are shown in Fig. 3.3. Fig. 3.3(a) shows the initial ball surface before contact, as used in most experiments. Surface deterioration due to sliding contact is shown in Figs. 3.3(b) – 3.3(d). It is clear that the surface deformation is plastic; some adhesion effects were also probably encountered. At the beginning of the life of the junction, damage to the contact surface is not very severe, Fig. 3.3(c). However, towards the end of the junction life, it seems that rupture of the junction takes place, and this results in the more severe damaged surface shown in Fig. 3.3(d). If a weld is formed during contact due to a high normal pressure, fracture at the weld will result in a tensile loading; the tension which may be supported by the junction will depend upon the adhesion strength of the junction. Thus, with a low adhesion strength, the junction will fracture immediately as it is subjected to any tension, and deformation as observed in Fig. 3.3(d) might possibly result.

Typical force traces are shown in Fig. 3.4. The variation in both forces is rather symmetrical. The relation between the friction and normal forces may be obtained by taking the ratio of these forces at each point in the life of the junction. Such a relation is shown in Fig. 3.5(a). It is clear from the contact geometry shown in Fig. 3.6 that the measured forces F_n and F_t are different from the "true" normal and tangential forces, N and T (i.e., forces normal and tangential to the spherical surfaces). Furthermore, since the direction of the friction force is always opposite to that of relative motion, the components F_n and F_t



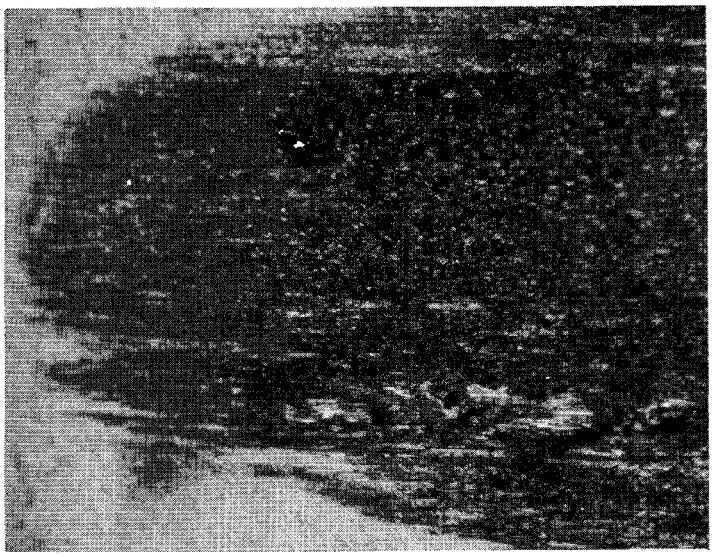
x100

FIG. 3.3(a) INITIAL BALL SURFACE, BEFORE CONTACT



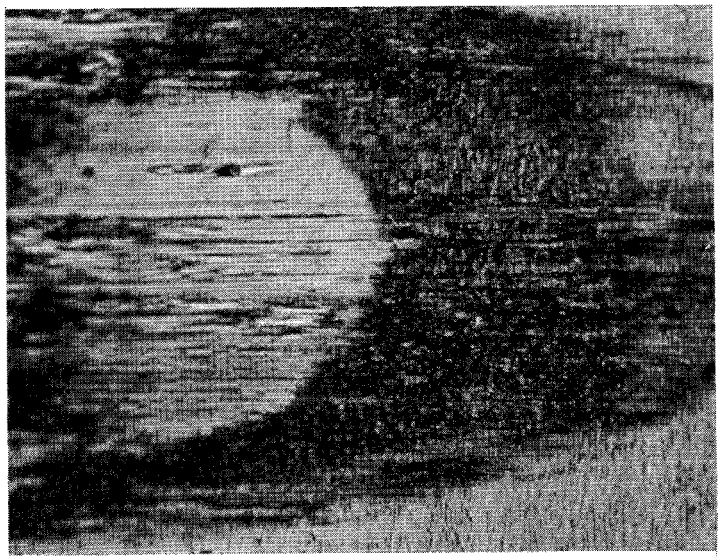
x100

FIG. 3.3(b) PHOTOMICROGRAPH OF THE BALL SURFACE AFTER SLIDING CONTACT. NORMAL CONTACT LOAD ≈150 LBS., SLIDING VELOCITY ≈½ INCHES/MIN. (THIS IS A PHOTOMICROGRAPH OF THE CENTRAL REGION OF THE TOTAL CONTACT ZONE)



x100

FIG. 3.3(c) PHOTOMICROGRAPH SHOWING THE BALL SURFACE AT THE BEGINNING OF CONTACT



x100

FIG. 3.3(d) PHOTOMICROGRAPH SHOWING THE BALL SURFACE AT END OF JUNCTION LIFE

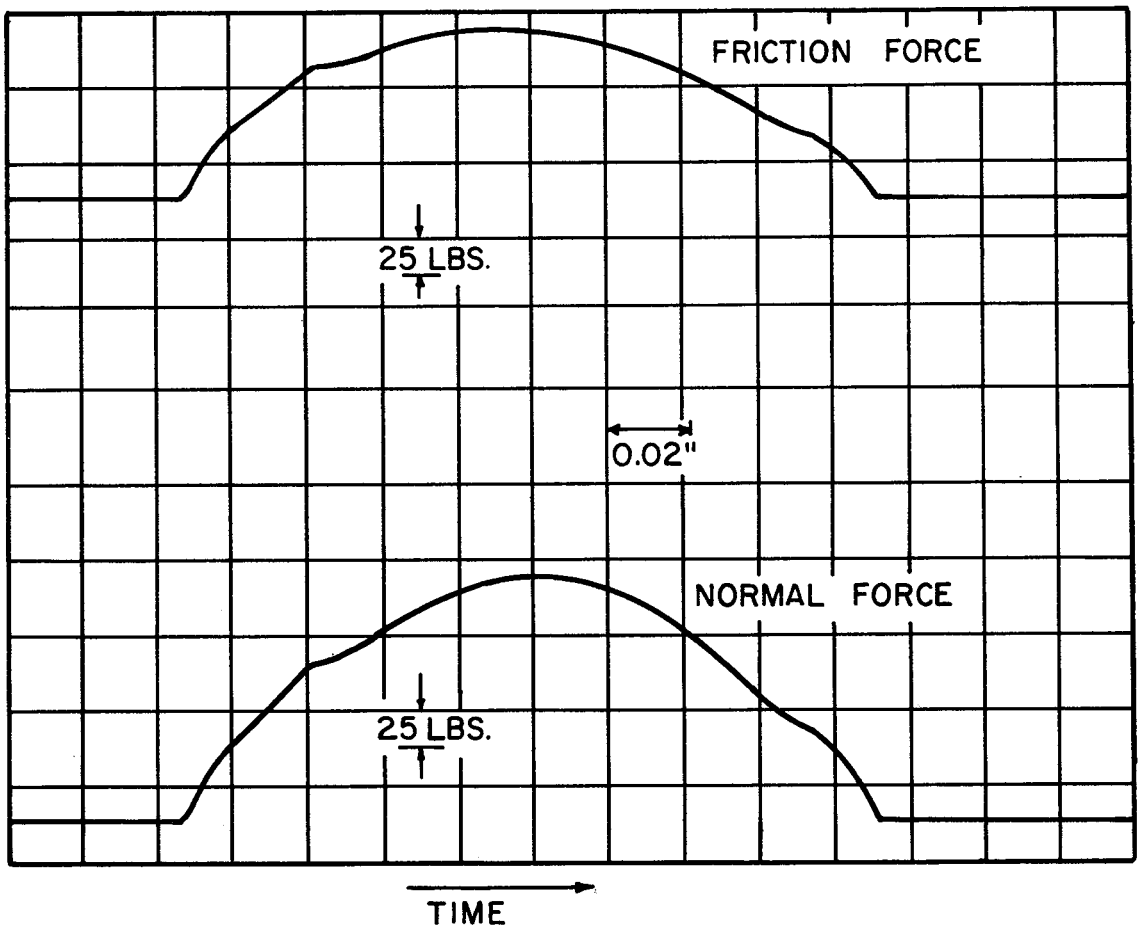


FIG. 3.4 VARIATION OF AVERAGE NORMAL AND TANGENTIAL LOADS DURING SLIDING. SLIDING VELOCITY = 0.539 INCHES/MIN.

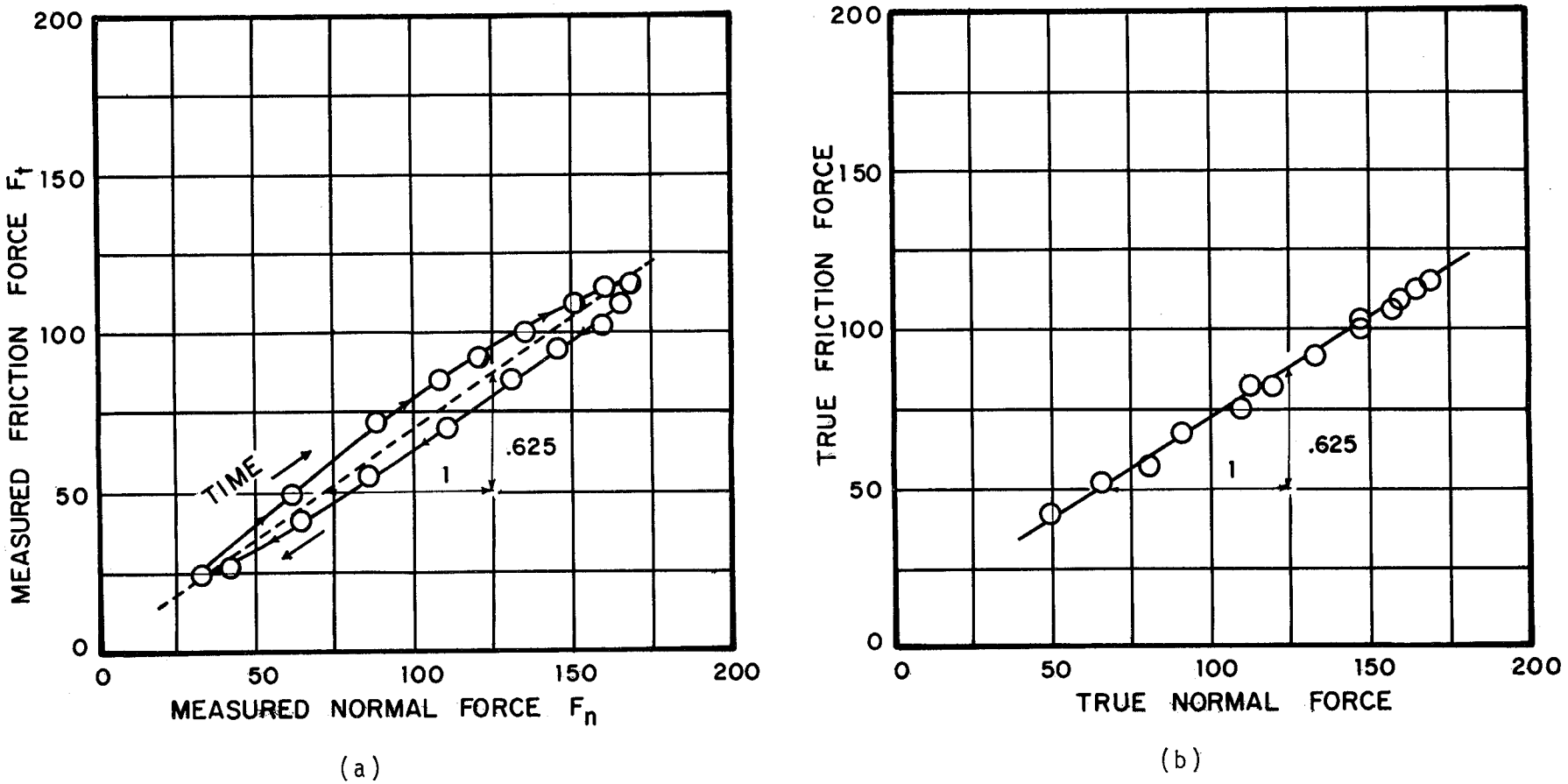


FIG. 3.5 RELATIONSHIP BETWEEN THE NORMAL AND FRICTION FORCES. DIRECT PROPORTIONALITY BETWEEN THE "TRUE" NORMAL AND FRICTION FORCE IS CLEARLY SEEN.

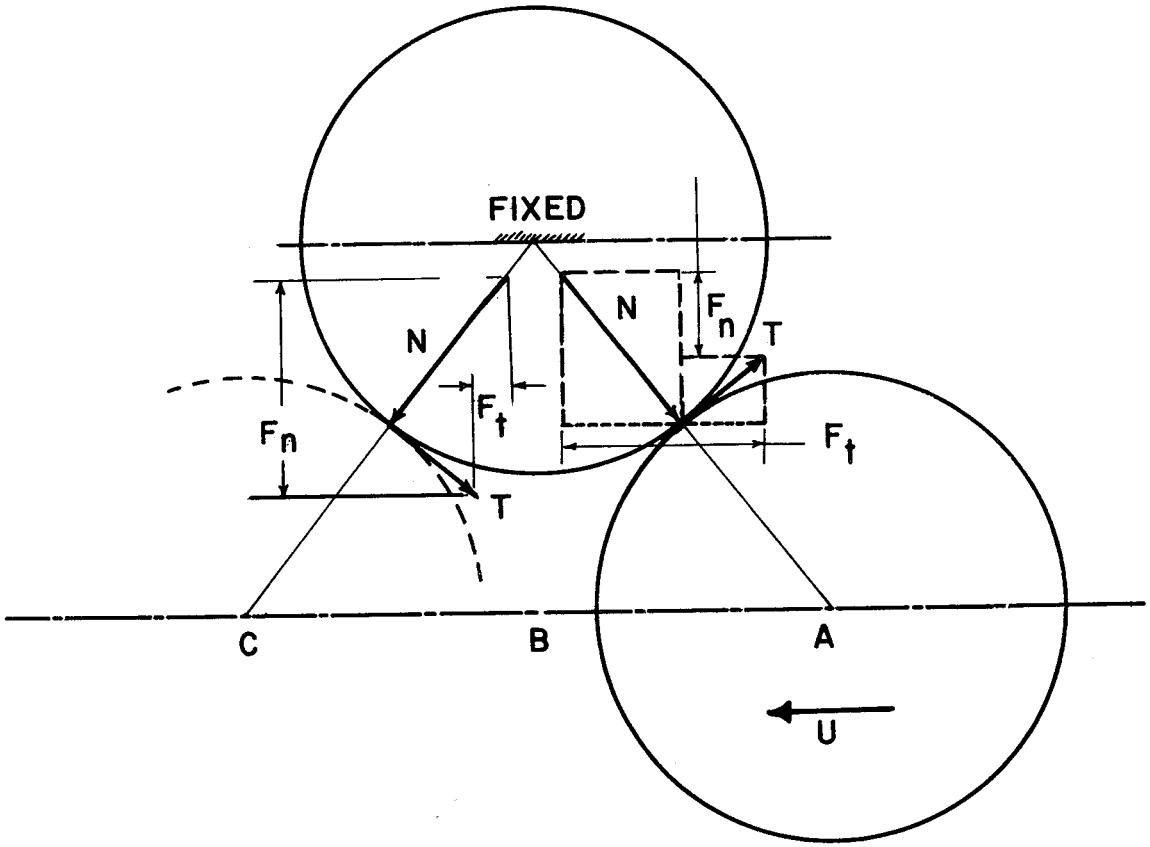


FIG. 3.6 EXAGGERATED GEOMETRY OF THE JUNCTION FORMED BY THE INTERACTION OF SPHERICAL BODIES.

will have different values at symmetrical positions between AB and BC (Fig. 3.6). However, the actual forces N and T may be computed by knowing the total contact length, by assuming that the center of the moving sphere moves in a straight line, and by assuming that the geometric interaction between the two spheres may be determined by the geometry shown in Fig. 3.6. The ratio of the "true" normal force N to the "true" tangential force T is plotted in Fig. 3.5(b). It may be seen that the "true" friction force is directly proportional to the "true" normal force at any time during the complete life cycle of the junction, though this is not true of the measured normal and tangential forces.

A difficulty experienced in all the experiments was due to the limited rigidity of the machine. Some part of the deformation took place in the structure of the machine instead of at the junction. Due to this effect the motion of the center of moving sphere deviated from a straight line, particularly with high contact loads. Thus the actual geometric interaction between the two spheres was indeterminate. An attempt was made to measure this interaction experimentally during the life cycle of a junction. A dial gauge was used to record the deflection of the milling machine table with respect to the column; however, the accuracy in such a measurement was very limited. An attempt was also made to design a suitable deflection measuring device using an L.V.D.T., but due to the limitations of time and some experimental difficulties this technique was not used during the course of the present work.

Most of the other results obtained are in terms of the measured normal and friction forces. Traces such as those shown in Fig. 3.4 were obtained. Since the maximum normal and tangential forces measured are equal to the "true" forces despite the geometry of contact and the indeterminate geometrical interaction,

these forces will be taken as characteristic of the junction in question and we shall plot these forces as dependent variables in reporting some of the other effects, which now follow.

3.4.1. The effect of repeated sliding

As sliding of one sphere past the other is repeated, both the maximum normal and friction forces tend to decrease. Some of the experimental data are plotted in Fig. 3.7. The reduction in the normal load and the friction force as the number of traversals increases is due to flattening of the contact surfaces due to deformation during sliding, and the resultant reduction in the geometric interaction between the spheres. The variation in friction coefficient is not significant in this case.

At an initially lower normal load, however, the normal load remains practically constant as the number of traverses increases. This is shown in Fig. 3.7 by a dashed line. The friction force continues to decrease, however, and a continuous reduction in friction coefficient is observed. This result shows that the proportionality between friction and normal forces does not always hold. The reason for this behavior is not known.

In the case of very high loads, deformations are highly plastic and hence the stresses are almost constant. Thus both the friction force and the total normal load are proportional to the area of contact and hence the two forces are proportional to each other. At light loads the elastic deformation which takes place in the substrate of the body in contact is probably not negligible compared to the plastic deformation at the interface and hence the relationship between friction and normal forces is not simple. The assumption of constant shear stress at the interface might be more realistic in this case. In the case of contaminated surfaces, of course, this assumption is always acceptable.

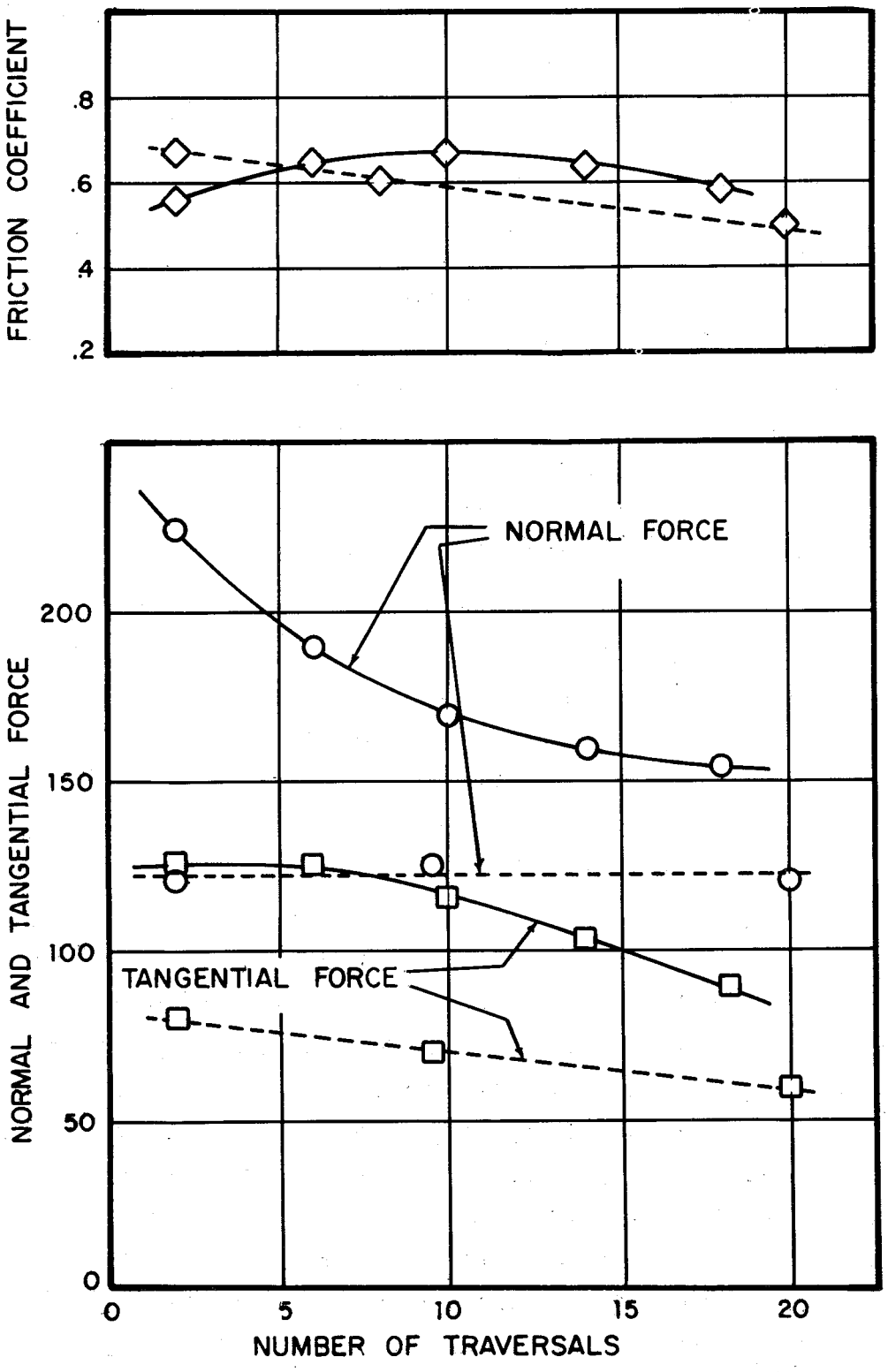


FIG. 3.7 THE EFFECT OF REPEATED TRAVERSALS.

3.4.2 The effect of sliding velocity

Experiments were performed at different sliding velocities. Some of the results, showing the variation of maximum normal and tangential forces during sliding, are shown in Fig. 3.8 and Fig. 3.9. In both of these plots a rapid increase in the normal contact force, and therefore a reduction in friction coefficient, is clearly observed.

As explained earlier the experimental set-up suffers from the problem of limited rigidity. Thus the amount of geometric interaction is not known. It was hypothesized that at high sliding speeds (and therefore reduced time of contact), the magnitude of rigid body rotation of the milling machine table (carrying one of the contacting spheres) is reduced. This would result in an increased geometric interaction and hence the normal contact load would increase. To check this hypothesis, dynamic response calculations were performed. The junction, along with the milling machine was considered to be a second-order system and the normal force was assumed to be of the form $F_0(1 - \cos\omega t)$, the frequency ω depending on the observed variation of normal force during contact. As the duration of contact was quite large compared to the period of any mode of vibration, it is unlikely that the observed increase in normal force is due to dynamic response effects. A Hertzian analysis was also made to determine the importance of impact effects, and they were found to be insignificant. The only remaining mechanism which can account for the observed behavior is that there is play in the machine, and that the rigid-body deflection of the milling machine table is larger at low speeds than at high speeds. This would lead to a smaller interference between the spheres at low speeds than at high speeds.

An interesting feature observed in both Figs. 3.8 and 3.9 is that there is a slight increase in the friction coefficient before it finally decreases. This increase is observed below a speed of about 4 in./min. At higher speeds, speeds owing to a reduced time of interaction, the strength of the adhesive bond might decrease. An alternative explanation is based on the hypothesis that the tangential (friction) force F_t is proportional to the area of contact, A. Thus $F_t \sim A$. However, $A \sim F_n^{2/3}$ (for elastic contact), where F_n is the normal load. Thus, the friction coefficient $f = F_t/F_n \sim F_n^{-1/3}$. As the normal load F_n increases with increasing sliding velocity, the friction coefficient f decreases.

3.4.3 Importance of contamination at the interface

The interface was contaminated with 20-20 motor oil and the variation of normal and friction forces as sliding continued repeatedly was measured. Some of the results are shown in Fig. 3.10. It is seen that the normal force decreases as the number of traversals increases as was observed in the case of dry contact. The friction force, however, increases as sliding continues.

As discussed earlier, it is probably more relevant to assume that the shear stress at a contaminated interface is constant and is equal to the ultimate shear strength of the contaminant. Thus, if lubricant is present at the interface, the shear stress will be considerably less than in dry contact. As repeated sliding of the two spheres continues, the contact zone is deformed plastically and the surfaces become rough. In this kind of rough surface, a dry contact between two microasperities becomes more probable, as the lubricant film, despite the continuous flow,

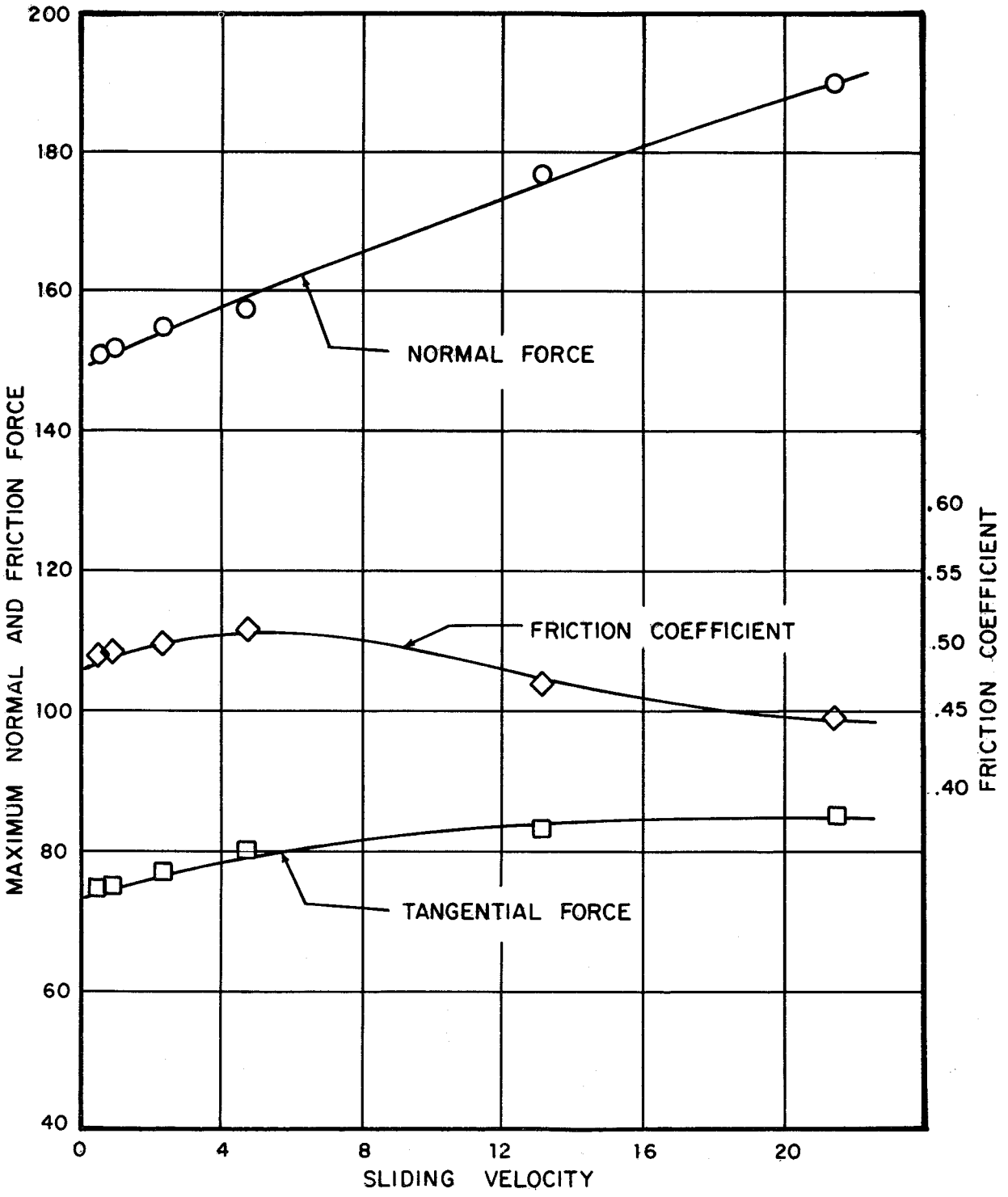


FIG. 3.8 VARIATION OF NORMAL AND FRICTION FORCES WITH SLIDING VELOCITY. (THE CONTACTING SURFACES WERE WELL RUN-IN TO EXCLUDE THE EFFECTS OF REPEATED CONTACT).

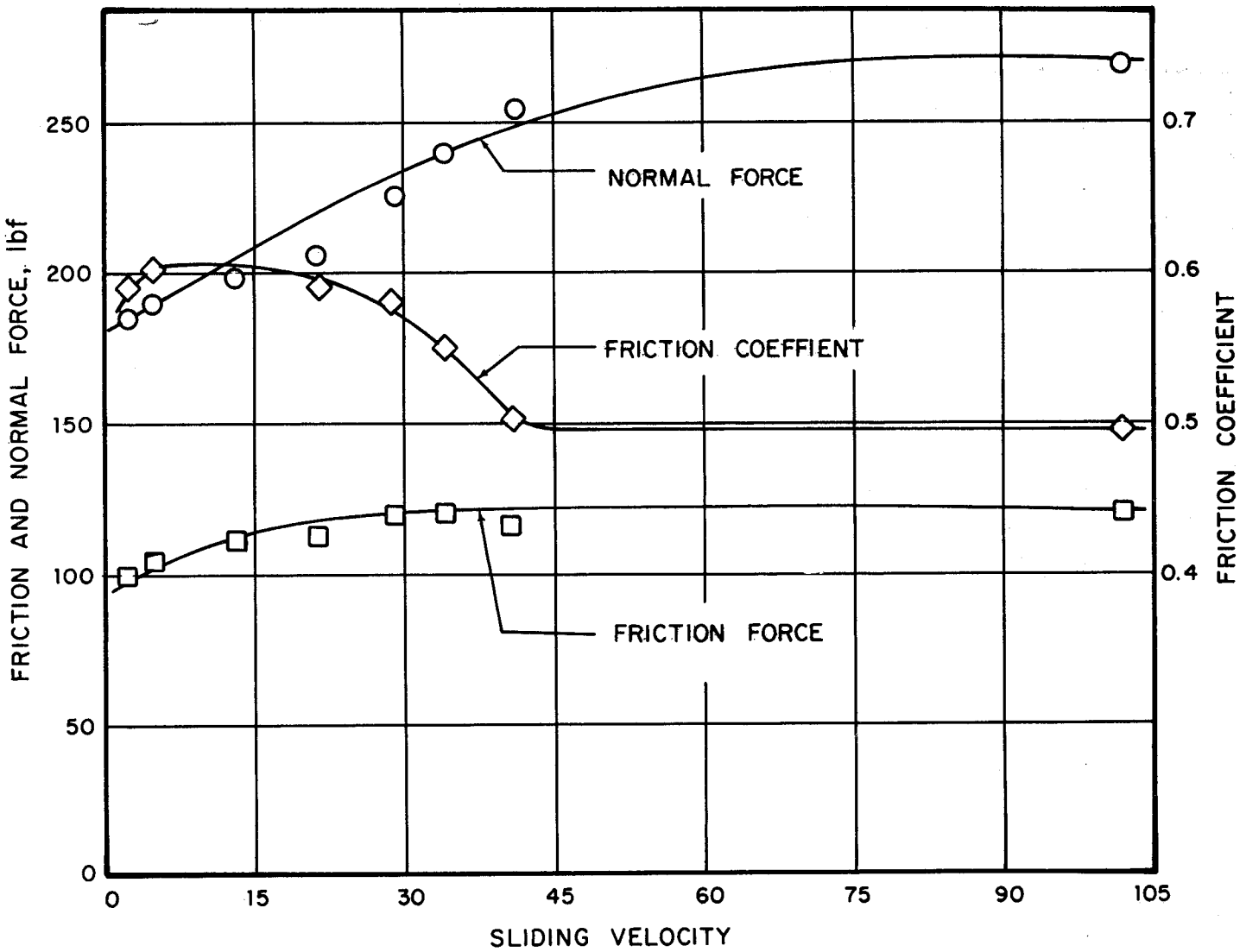


FIG. 3.9 FRICTION AND NORMAL FORCE VARIATION AT HIGH VELOCITIES OF SLIDING. CONTACT SURFACES ARE WELL RUN-IN TO EXCLUDE THE EFFECTS OF REPEATED CONTACT.

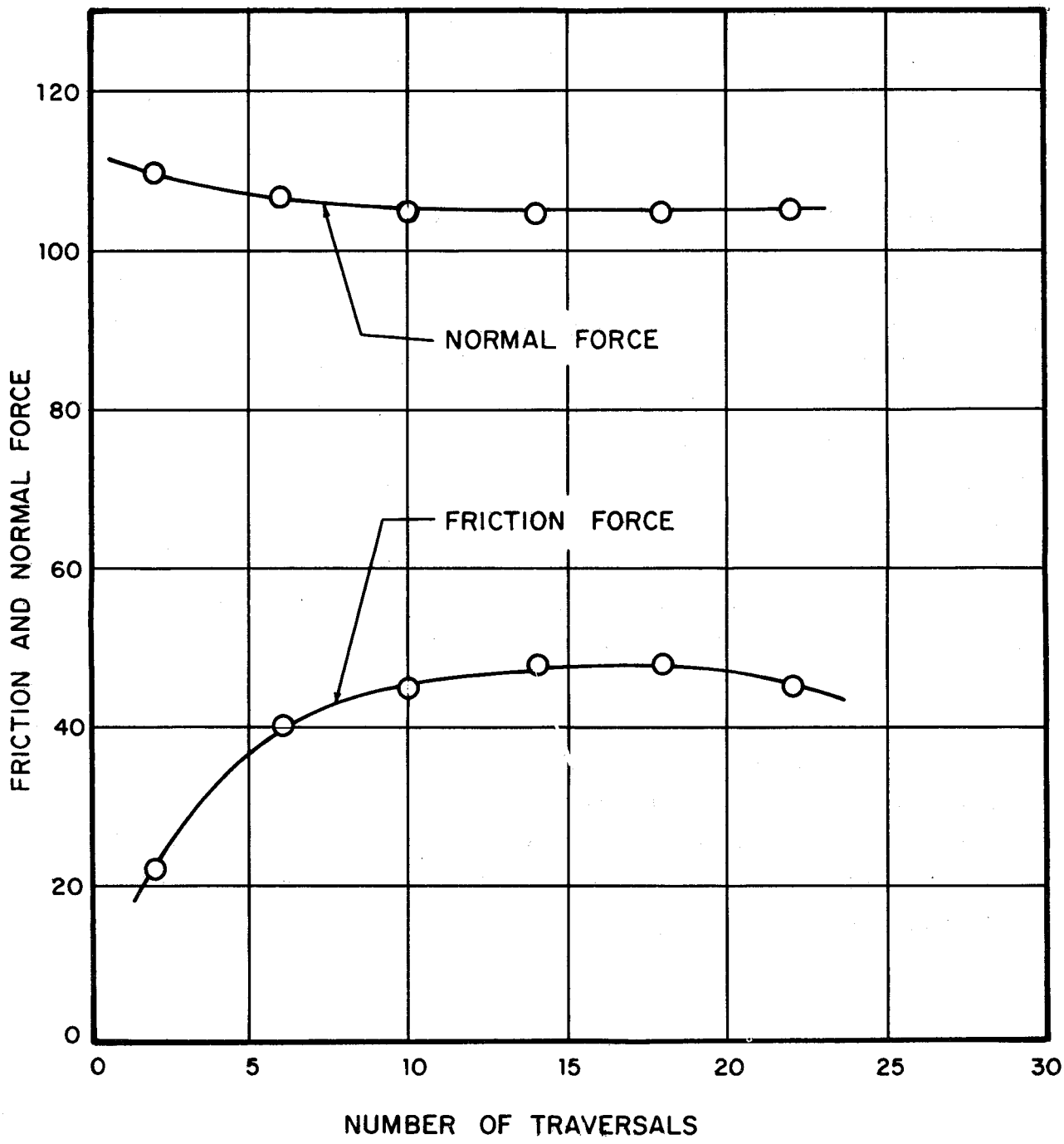


FIG. 3.10 EFFECT OF CONTAMINATING THE JUNCTION WITH 20-20 MOTOR OIL.

can be easily broken through. An increase in the interfacial shear stress then takes place, and the friction coefficient increases. This is one possible explanation for the behavior shown in Fig. 3.10.

3.4.4 The effect of superposed vibrations

An attempt was made to study the effect of asperity vibrations on the behavior of junctions by shaking one of the spheres with an electromagnetic shaker. However, difficulties were encountered in trying to obtain sufficiently high acceleration levels, and it was concluded that it is necessary to redesign the experimental apparatus for such a study. Moreover, it would be preferable to measure loads with a piezoelectric crystal instead of a strain gauge dynamometer, in order to obtain greater accuracy at high frequencies. Nevertheless, one qualitative conclusion was reached: asperity vibrations are more important at low loads than at high loads. Furthermore, the dependence is not on load *per se*, but on the magnitude of the surface deformations. With large plastic deformations, surface vibrations are less important than with elastic deformations. Thus it is conceivable that if the extent of wear is small, the asperities would shakedown to an elastic state with repeated contact, and that vibration effects would be important, even with high loads.

APPENDIX

Calibration of the 3-D Dynamometer:

Fig. 3.11 shows the experimental set-up used for calibrating the dynamometer. A suitable fixture was bolted on the dynamometer and dead weights were applied, as shown in Fig. 3.11. The output of the dynamometer was fed either to the Sanborn chart recorder or to a carrier amplifier. Output was displayed on an oscilloscope in the later case.

The following calibrations were obtained:

TABLE 3.1 Calibration of 3-D Dynamometer of Sanborn Model 321

Force	Calibrate #mm at Att.		5 lbs equals #mm at Att.		Remarks
Normal F_n	24	x1	1	x1	
Tangential F_t	24	x1	1	x1	

TABLE 3.2 Calibration of 3-D Dynamometer for Carrier Amplifiers

Force	At Strain Sensitivity	10 lbs equals #volt output	Remark
Normal F_n	x20	0.75 volts	The amplifier gain was adjusted to obtain these calibrations.
Tangential F_t	x20	1.50 volts	

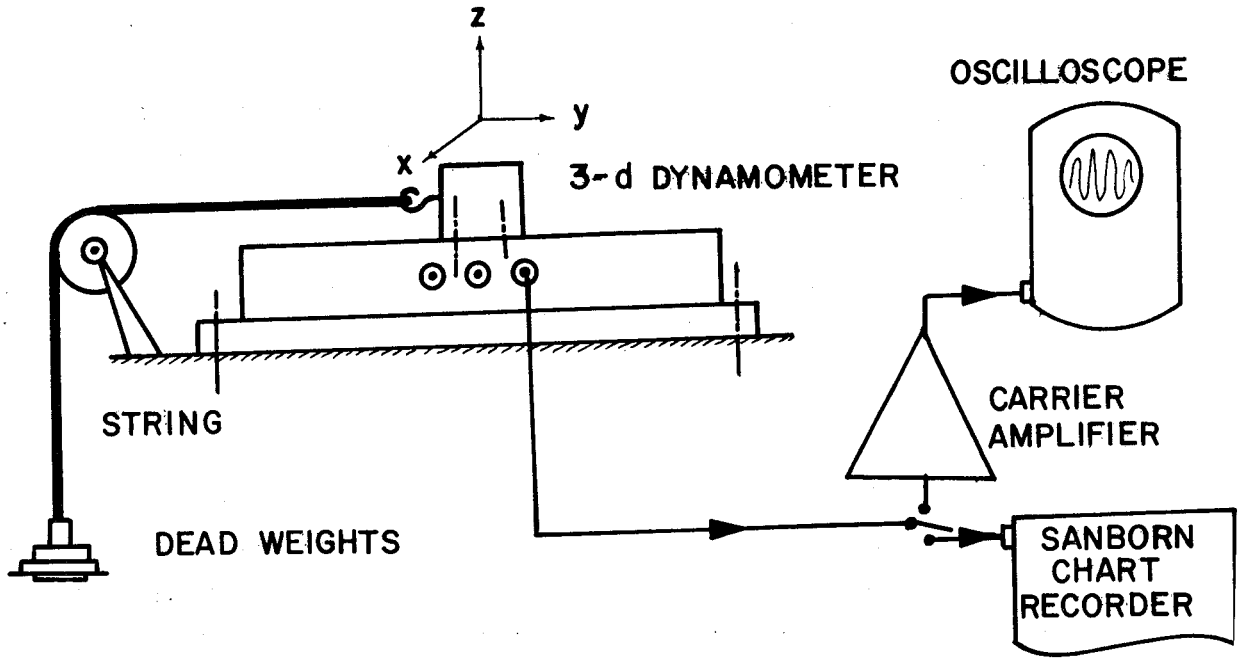


FIG. 3.11 EXPERIMENTAL SET-UP FOR CALIBRATION OF THE 3-d DYNAMOMETER.

Accelerometer Sensitivity:

The given sensitivity for the accelerometer was 390 mV/g and the cathode follower amplifier employed had a gain of 0.7. Thus for a sinusoidal displacement input we have,

$$X = X_0 \sin(wt) \quad .$$

Acceleration may be obtained by differentiating the above expression two times, thus

$$X = -X_0 w^2 \sin(wt) \quad .$$

Thus if the measured peak to peak voltage in mV is V_p and the frequency is f in cycles/sec, then

$$\frac{V_p}{0.70 \times 390} \cdot g = 2X_0 w^2$$

or $$X_0 = (V_p \cdot g)/(546 \cdot w^2)$$

or $$X_0 = 456(V_p/f^2) \text{ microns.}$$

Thus the amplitude of vibration in microns can be estimated by knowing the output peak to peak voltage, V_p in mV and the frequency of vibration f in cycles/sec.

REFERENCES

1. S.B. Ainbinder and A.S. Pranco, "On the Mechanism of the Formation and Destruction of Adhesion Junctions Between Bodies in Frictional Contact," *Wear* 9, 209-227 (1966).

2. M. Cocks, "Shearing of Junctions Between Metal Surfaces," *Wear* 9, 320-328 (1966).

3. C.M. Edwards and J. Halling, "An Analysis of the Plastic Interaction of Surface Asperities and Its Relevance to the Value of the Coefficient of Friction," *J. Mech. Engg. Sc.* 10, 2, 101-110 (1968).

4. A.P. Green, "The Plastic Yielding of Metal Junctions Due to Combined Shear and Pressure," *J. Mech. Phys. Solids* 2, 197-211 (1954). "Friction Between Unlubricated Metals: A Theoretical Analysis of Junction Model," *Proc. Roy. Soc.* (London) 228, 191-204 (1955).

5. J.A. Greenwood and D. Tabor, "Deformation Properties of Junctions," *Proc. Phys. Soc.* (London) 58, 609-619 (1955).

6. J.A. Greenwood and J.H. Tripp, "The Contact of Two Nominally Flat Surfaces," (to Be Published).

7. J.A. Greenwood and J.B.P. Williamson, "Contact of Nominally Flat Surfaces," *Proc. Roy. Soc.* (London) 295, 300-319 (1966).

8. I.V. Kragelskii, *Friction and Wear,* Butterworth and Co. Ltd., England (1965).

9. D.M. Tolstoi, "Significance of the Normal Degree of Freedom and Natural Normal Vibrations in Contact Friction," *Wear* 10, 199-213 (1967).

10. P.K. Gupta, "Topographic Analysis of Friction Between a Pair of Rough Surfaces," ScD Thesis, Department of Mechanical Engineering, MIT (September, 1970).

4. THE DYNAMICS OF A SIMPLE VEHICLE ON RANDOM TRACK

Summary

A lumped six-degree-of-freedom vehicle moving on random track is analyzed. The inputs from the track are a mean rail height and a superelevation, both considered to be random variables. The power spectral densities used for these inputs are reasonable approximations of measured data. The rails are considered to be beams on elastic foundations. The power spectral densities and standard deviations of the dynamic wheel load and the dynamic axle load are obtained. An investigation of the effects on these loads of the truck and axle mass, the foundation stiffness and the primary suspension damping is reported. The main conclusions of this part of the investigation are:

1. The main peaks in the power spectra are those due to (a) the car body bouncing on the secondary suspension, at about 1 Hz, and (b) the wheelset bouncing on the track, at about 25-30 Hz. For the track power spectral densities considered here, the effects on the dynamic loads of angular roll of either the car body, truck or wheelset are negligible.

2. The standard deviation of the dynamic loads is due almost entirely to response at the wheelset bounce frequency. Decreasing the truck or wheelset mass has a salient effect on the dynamic loads. A decrease in the wheelset and truck mass of 20%, amounting to a decrease in the vehicle mass of about 3%, decreases the standard deviation of the dynamic loads by about 10% at 150 mph.

3. The foundation stiffness has a profound effect on the dynamic loads — the stiffer the foundation, the higher the dynamic loads. However, reductions in the foundation stiffness are

constrained, among other things by (a) the bending stress in the rail, and (b) the effect of large deflections of the rail on the supporting structure. With a soft foundation, the wheel load spreads to a smaller degree than it does with a stiff foundation. If the supporting structure were ballast, for example, this could have a detrimental effect on the degradation of the ballast.

4. The optimum value of the primary suspension damping ratio for the vehicle studied here from the viewpoint of dynamic loads is

$$\zeta_1 \simeq 0.75 \times \zeta_{critical}$$

$$= 0.75 \times 2 \times (\text{Truck Mass} \times \text{Primary Suspension Stiffness})^{1/2}.$$

Variations from this value have a large effect on the peak values of the power spectra of the dynamic loads, but a smaller effect on their standard deviations.

The results of this preliminary investigation are then applied to two phenomena of interest. These are plastic flow in the rail, and the braking of a wheelset. The main conclusions are as follows:

Plastic Flow

1. Plastic collapse of the rail due to dynamic loads, i.e., the formation of a plastic hinge is a highly improbable event.

2. Localized plastic flow, around the contact region, begins to be significant above speeds of about 40 mph. The length of these damaged zones increases as the vehicle speed increases, as does their frequency of occurrence. The most important ways in which the frequency of damage can be reduced are (a) a reduction in the foundation stiffness, (b) a reduction

in the wheelset mass and (c) the introduction of optimum damping into the primary suspension and the track. As an example, for the baseline vehicle studied here, at 150 mph, there is one damaged zone every 30 ft; the average length of these zones is about 1 ft. At 90 mph, there is one damaged zone every 1000 ft, the average length being about 7 in. Thus it is manifestly unsafe to extrapolate from data obtained at lower speeds to significantly higher speeds, in predictions of rail damage.

Braking of a Wheelset

1. If data on the friction of steel on steel as a function of sliding velocity obtained in sliding contact experiments are also valid in rolling contact, dynamic loads have a negligible effect in decreasing the effective normal load on a wheelset for tractive purposes. The important factor here is the rate at which the friction coefficient decreases as the slip velocity between wheel and rail increases.

2. There is a need for an experimental investigation of friction in rolling contact. Experiments hitherto have been concerned solely with creep effects. Instead of having a constant torque being transmitted, it is necessary to study rolling contact with the angular velocity of the rolling elements being controlled to give a steady slip velocity. The torque would then be measured as a function of the slip velocity. Factors that would have to be considered in such an experiment are the normal load, the surface contamination, surface deterioration and surface temperature.

3. Should the above experiments show that dynamic loads are important, then it appears that low-frequency dynamic loads,

at the car bounce frequency, are more effective in reducing tractive capacity than high-frequency loads at the wheelset bounce frequency. It thus appears possible that the design of the secondary suspension is more important than the design of the unsprung mass and the primary suspension for the control of tractive capacity. This reasoning would also explain why there is a swift decrease in tractive capacity at low speeds: car bounce occurs at all but the lowest speeds, whereas wheelset bounce becomes important only at higher speeds. As soon as the speed is large enough for car bounce to become important, the tractive capacity begins to decrease.

4. Other possible reasons for the decrease in tractive capacity that need to be investigated in detail are: (a) As the rolling velocity increases, the time available for the formation of friction junctions decreases, causing a decrease in the strength of these junctions and in their ability to transmit tractive forces. (b) Increased surface vibration due to rail irregularities and increased impact of asperities entering the contact region decrease the strength of the junctions. (c) Elastohydrodynamic effects due to surface contaminants become important at higher speeds.

4.1 Introduction

When a train moves over tangent track that is not perfectly flat, the load between any wheel and its rail fluctuates in time. The deviation of the wheel load from its mean value is termed the dynamic load. In this paper, the following problem is analyzed: A wheelset carrying a truck and a car body moves over tangent tract at some speed S. The rails have random profiles which are described by the power spectral densities (PSD's) of the mean height and the angle of superelevation of the rails. An analytical scheme is sought whereby the statistical properties of the dynamic wheel loads may be predicted reasonably accurately.

The present paper differs from past work [*1,2,3*] in one of two respects: (1) No simulation techniques are used here; analytical methods are used throughout, leading to closed form expressions for quantities of interest. (2) The rails are not considered to be rigid, they are, instead, modeled as beams on an elastic foundation, and the resilience of the foundation is found to have a strong influence on the dynamic loads. The vertical profile of the rails is also considered to be a random process.

The importance to a designer of being able to predict dynamic loads is evident. The noise and vibration in the car (and therefore the passenger comfort) depend on the dynamic load inputs at the wheels. The probability of derailment due to wheel hop or of damage to the rail due to excessive loading can be estimated only when the statistics of the dynamic load are known. Furthermore the tractive capacity of a wheelset during acceleration or braking may be limited by the dynamic load it exerts on the track. In this paper, we confine ourselves to a discussion of the PSD's of the dynamic loads, and to an analysis of both the tractive capacity of the wheelset and the probability of damage to the rails.

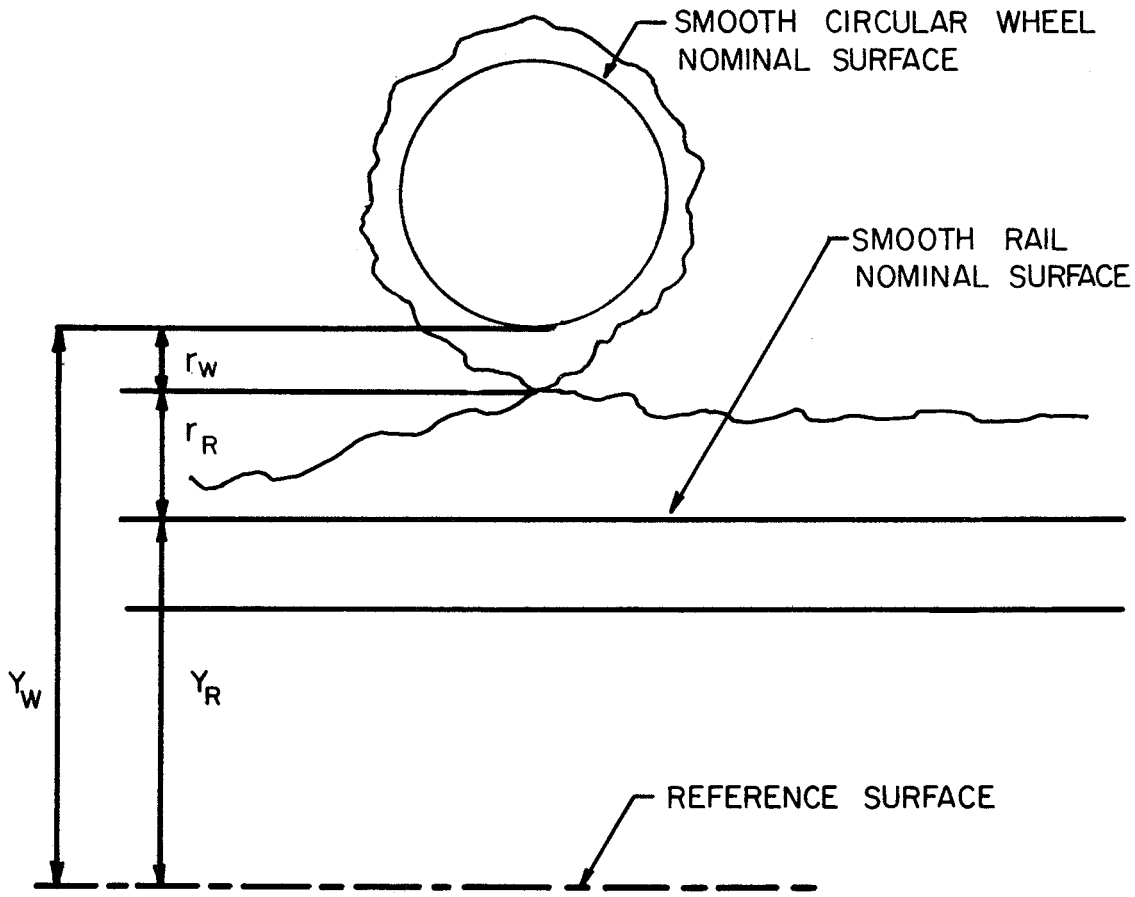
4.2 The Model

The vehicle considered in the following analysis is shown in Fig. 4.1. It is characterized by the masses and rotary inertias (about an axis parallel to the rails) of the wheelset, the truck and the car body, and by the stiffness and damping of the primary and secondary suspensions.

The stiffness and damping are both considered to be linear. Specifically, each spring-dashpot combination is characterized by a constant stiffness and a linear viscous damping. It is recognized that the latter is not an accurate representation of material, viscoelastic or interface damping. Nevertheless, it is adopted here for the sake of simplicity, and in the belief that qualitative comments that we make regarding the efficacy of damping will be valid despite the inaccuracy of the representation. All nonlinearities, such as those arising from a clearance between the axle and its bearings, are neglected. The wheels are considered to be rigid masses, rigidly attached to the axles. This assumption is justifiable for excitation frequencies below the first bending resonance frequency of the wheel acting as a flat plate, which may be shown to be of the order of 350 Hz [*4*].

The rails are considered to be beams on an elastic foundation. The foundation is assumed to be of the point-reacting or Winkler type, characterized by a foundation stiffness k_f (lbf/ft/ft). Damping in the foundation is not taken into account.

In most practical cases (i.e., except when a very high excitation frequency is coupled with a high amplitude), the elastic deflections due to the Hertzian contact between the wheel and



$$Y_W = Y_R + r_R + r_W$$

FIG. 4.1 WHEEL-RAIL GEOMETRY

the rail may be neglected. This implies that, at the contact, the rail and the vehicle each look into linear, load-independent impedances.

Coupling between the bounce (up-and-down) and roll (rotation about a longitudinal axis) modes is assumed to be absent. This permits a significant simplification in the analysis.

4.3 Analysis

4.3.1 Contact geometry

Consider a wheel moving over a rail with a random profile, as shown in Fig. 4.2. Then the following geometric relation holds [4]:

$$y_W = y_R + r_R + r_W \ , \tag{1}$$

where r_W = wheel irregularity

r_R = rail irregularity

y_R = height of nominal rail surface from a fixed reference surface

y_W = height of nominal wheel surface from the same reference surface.

As the wheel moves along the rail, Eq. 1 holds, but differentiation with respect to time also yields

$$v_W(t) + v_R(t) = \dot{r}_R + \dot{r}_W \ , \tag{2}$$

where v_W and v_R are the upward and *downward* velocities, respectively, of the wheel and rail nominal surfaces.

The Fourier Transform of Eq. 2 gives a relationship between the spectral components of the velocities:

$$V_W(\omega) + V_R(\omega) = R_R(\omega) + R_W(\omega) \equiv V(\omega) \ , \tag{3}$$

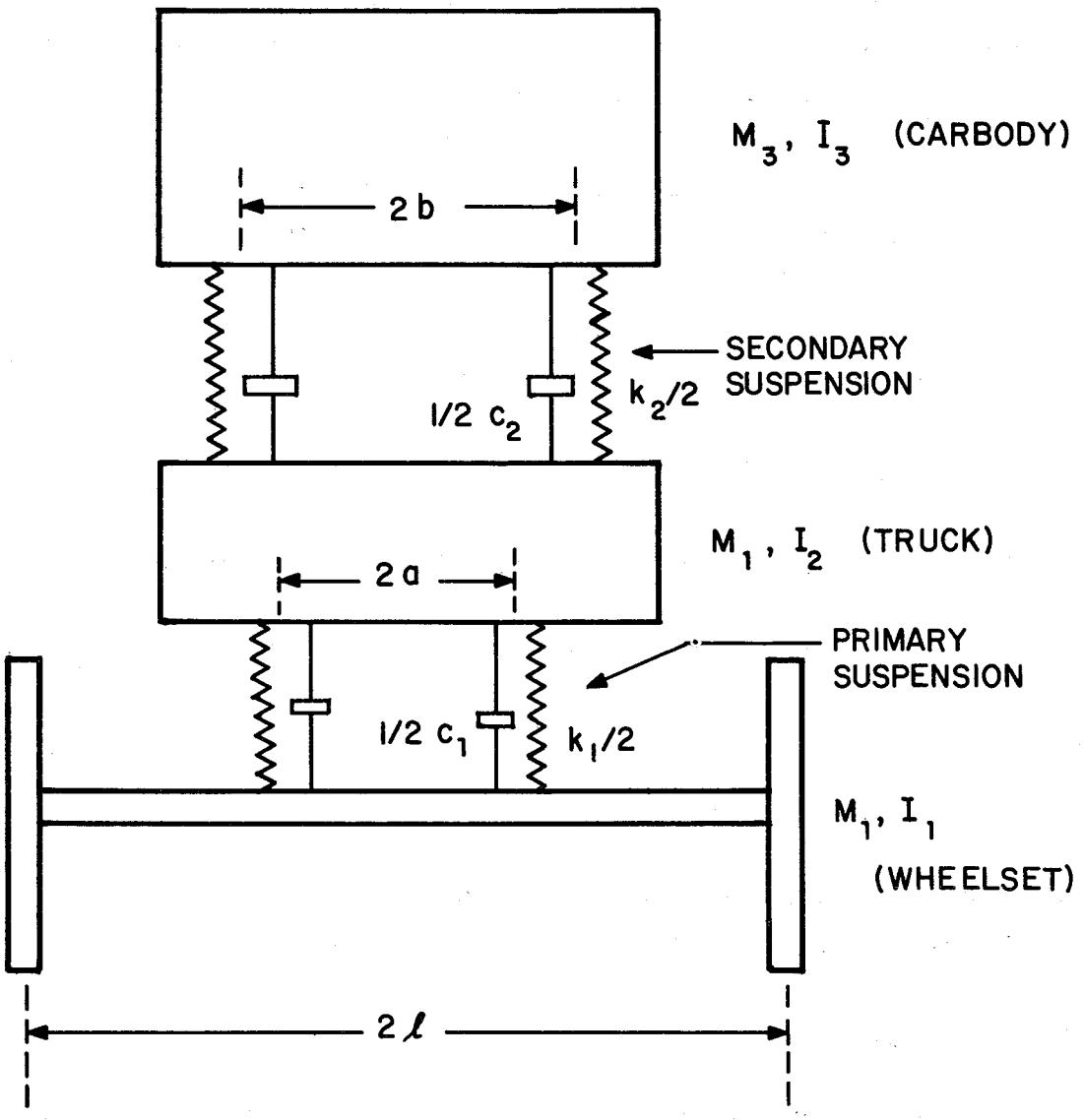


FIG. 4.2 VEHICLE MODEL

where $V(\omega)$ is a composite input velocity at the contact between wheel and rail, due to both rail and wheel irregularities.

The connection between the velocity input $V(\omega)$ and the geometrical irregularities of wheel and rail, which are measured as displacement inputs, may be obtained as follows. Let $r_R(x)$ and $r_W(x)$ be the displacement irregularities of the rail and wheel. If the wheel that senses these irregularities is moving at a speed S, the velocity due to these irregularities is obtained by means of a total derivative:

$$v(t) = \frac{d}{dt}(r_R + r_W) = \frac{\partial}{\partial t}(r_R + r_W) + S\frac{\partial}{\partial x}(r_R + r_W)$$

$$= S(r_R' + r_W') \quad . \tag{4}$$

Taking the Fourier Transform of Eq. 4, we obtain

$$V(\omega) = - ik[R_R(k) + R_W(k)] \quad , \tag{5}$$

where

$$k = \omega/S \quad , \tag{6}$$

and $R_W(k)$ and $R_R(k)$ are the spectral components of the wheel and rail irregularities with wavelength $\lambda = 2\pi/k$.

4.3.2 Dynamics

Now consider the vehicle of Fig. 4.1 to be moving over two random rails. Equations 3 and 5 then apply at both wheel-rail contacts. Using the subscripts 1 and 2 to denote the left and right contacts, respectively, we may write the following relations for the left contact:

$$V_{1W}(\omega) = F_1(\omega)Z_1(\omega) + F_2(\omega)/Z_{12}(\omega) \quad , \tag{7}$$

and

$$V_{1R}(\omega) = F_1(\omega)/Z_{1R}(\omega) \quad , \tag{8}$$

where $F_1(\omega)$ = spectral component of the dynamic wheel load at location 1

$Z_1(\omega)$ = input impedance of the vehicle at location 1

$Z_{12}(\omega)$ = cross-impedance of the vehicle between locations 2 and 1;

also, $Z_{1R}(\omega)$ = input impedance of the rail.

In addition, Eq. 3 gives

$$V_{1W}(\omega) + V_{1R}(\omega) = V_1(\omega) \quad . \tag{9}$$

Eliminating V_{1W} and V_{1R} from Eqs. 7, 8 and 9, we obtain

$$F_1(Z_1 + Z_{1R})/Z_1Z_{1R} + F_2/Z_{12} = V_1 \quad . \tag{10}$$

A similar equation may be written for location 2 on the vehicle:

$$F_1/Z_{21} + F_2(Z_2 + Z_{2R})/Z_2Z_{2R} = V_2 \quad . \tag{11}$$

If we now assume

$$\begin{aligned} Z_1 &= Z_2 = Z_V \quad , \\ Z_{12} &= Z_{21} = Z_C \quad , \\ Z_{1R} &= Z_{2R} = Z_R \quad , \end{aligned} \tag{12}$$

Eqs. 10 and 11 may be solved to obtain

$$F_1 = \frac{ZZ_c}{Z^2 - Z_c^2}(V_2 Z - V_1 Z_c) \quad , \tag{13}$$

where $Z \equiv Z_V Z_R/(Z_V + Z_R)$.

A similar equation holds for F_2. Furthermore, the total dynamic axle load is found to be

$$F_T \equiv F_1 + F_2 = \frac{ZZ_c}{Z + Z_c}(V_1 + V_2) \quad . \tag{14}$$

Before proceeding further, we introduce the mean bounce and roll input velocities, defined by

$$V_m = \frac{1}{2}(V_1 + V_2) \quad , \tag{15}$$

and

$$\dot{\theta} = \frac{1}{2\ell}(V_1 - V_2) \quad , \tag{16}$$

where 2ℓ is the distance between the rails. Introducting Eqs. 15 and 16 into Eqs. 13 and 14, and using the Wiener-Khintchine relationship for power spectral densities [5], we obtain

$$\phi_{F_1F_1}(\omega) = \phi_{V_mV_m}|Z_+|^2 + \ell^2\phi_{\dot{\theta}\dot{\theta}}|Z_-|^2 + \text{terms in } \phi_{V_m\dot{\theta}} \text{ and } \phi_{\dot{\theta}V_m} \quad . \tag{17}$$

Here, $\phi_{F_1F_1}$ is the power spectral density (PSD) of the force F_1; $\phi V_m V_m$ the PSD of the mean bounce input velocity; $\phi_{\dot\theta\dot\theta}$ the PSD of the mean roll input velocity; and $\phi_{V_m\dot\theta}$ and $\phi_{\dot\theta V_m}$ are the cross-PSD's indicating coherence between V_m and $\dot\theta$. We shall henceforth neglect the cross-PSD's of the input velocities. Equation 17 thus becomes

$$\phi_{F_1F_1}(\omega) = \phi_{V_mV_m}|Z_+|^2 + \ell^2\phi_{\dot\theta\dot\theta}|Z_-|^2 \ , \tag{18}$$

where

$$\frac{1}{Z_+} = \frac{1}{Z_V} + \frac{1}{Z_c} + \frac{1}{Z_R} \ , \tag{19}$$

and

$$\frac{1}{Z_-} = \frac{1}{Z_V} - \frac{1}{Z_c} + \frac{1}{Z_R} \ . \tag{20}$$

In a similar fashion, the PSD of the total axle load F_T is found to be

$$\phi_{F_TF_T} = 4\ \phi_{V_mV_m}|Z_+|^2 \ . \tag{21}$$

4.3.3 Roughness spectra

In order to proceed with the analysis, it is necessary to obtain expressions for $\phi_{V_mV_m}$ and $\phi_{\dot\theta\dot\theta}$ in Eqs. 18 and 21.

To do this, we proceed as follows. Combining Eqs. 15 and 16, and two such as Eq. 5, we obtain

$$V_m = - ik[R_{mR}(k) + R_{mW}(k)] \quad , \tag{22}$$

and

$$\dot{\theta} = - ik[\theta_R(k) + \theta_W(k)] \quad , \tag{23}$$

where

$$\left.\begin{aligned} R_{mr} &= \frac{1}{2} (R_{1R} + R_{2R}) \\ R_{mW} &= \frac{1}{2} (R_{1W} + R_{2W}) \\ \theta_R &= \frac{1}{2\ell} (R_{1R} - R_{2R}) \\ \theta_W &= \frac{1}{2\ell} (R_{1W} - R_{2W}) \end{aligned}\right\} . \tag{24}$$

Applying the Wiener-Khintchine relationship to Eqs. 22 and 23, and assuming the wheel and rail roughnesses to be uncorrelated with each other, we obtain

$$\phi_{V_m V_m} = k^2 S \left[\phi_{R_{mr} R_{mr}}(k) + \phi_{R_{mW} R_{mW}}(k)\right] \quad , \tag{25}$$

and

$$\phi_{\dot{\theta}\dot{\theta}} = k^2 S \left[\phi_{\theta_R \theta_R}(k) + \phi_{\theta_W \theta_W}(k)\right] \quad . \tag{26}$$

The PSD's for the wheel and rail random inputs are not known separately, but usual measuring techniques give the sums of the PSD's needed in Eqs. 25 and 26.

Measurements indicate that the following expressions hold for the sums of the PSD's of the rail and wheel inputs [6]:

$$\phi_{R_{mr}R_{mr}} + \phi_{R_{mw}R_{mw}} = 2\pi\beta\times10^{-7} \times \begin{cases} 1 \text{ , } ft^2/(1/ft) & , \quad k \leq k_o = 1.256 \\ (k_o/k)^2 & , \quad k > k_o \end{cases} \tag{27}$$

and

$$\phi_{\theta_R\theta_R} + \phi_{\theta_W\theta_W} = 2\pi\beta\gamma\times10^{-8}\times \begin{cases} 1, \text{ rad}^2/(1/ft) & , \quad k \leq k_o = 1.256 \\ (k_o/k)^2 & , \quad k > k_o \end{cases} . \tag{28}$$

More accurate representations are of the form

$$\frac{c_1}{c_2 + k^2(c_3 + k^2)} ,$$

where c_1, c_2, and c_3 are constants, but the error introduced by the simplification used in Eqs. 27 and 28 is small. Equations 25 and 26 become

$$\phi_{V_mV_m} = 2\pi\beta S\times10^{-7} \times \begin{cases} 1 \text{ , } (ft/sec)^2(1/sec), & \omega \leq \omega_o = 1.256S \\ (\omega_o/\omega)^2 & , \quad \omega > \omega_o \end{cases} , \tag{29}$$

and

$$\phi_{\dot{\theta}\dot{\theta}} = \frac{2\pi\gamma S\times 10^{-7}}{\ell^2} \times \begin{cases} 1, \ (\text{rad/sec})^2/(1/\text{sec}), & \omega \leq \omega_o = 1.256S \\ (\omega_o/\omega)^2 & \omega > \omega_o \end{cases} . \qquad (30)$$

The constants β and γ indicate track quality and range from near 2 for good track to near 8 for poor track. In the sequel, we take $\beta=2$, $\gamma=4$.

Here it is worth noting that in the dynamic analysis, static (or mean) loads and deflections have been neglected, as they may be for linear systems. However, this does require that the input displacement at the wheel-rail contact due to wheel and rail irregularities be obtained with the appropriate static load being transmitted from wheel to rail. The experimental data in [6], on which Eqs. 27 and 28 are based were, indeed, obtained with this requirement in mind. The effect of making profile measurements with a constant static load on the rail is to take into account in some way local variations in rail and foundation properties.

Combining Eqs. 18, 21, 29 and 30, we obtain the following expressions for the PSD's of the wheel load F_1 and the total axle load F_T:

$$\phi_{F_1F_1}(\omega) = 2\pi S\times 10^{-7}\times\left[\beta|Z_+|^2+0.18|Z_-|^2\right] \times \begin{cases} 1 \quad , & \omega \leq \omega_o = 1.256S \\ (\omega_o/\omega)^2, & \omega > \omega_o = 1.256S \end{cases} \qquad (31)$$

and

$$\phi_{F_T F_T}(\omega) = 8\pi\beta S \times 10^{-7} |Z_+|^2 \times \begin{cases} 1 & , \omega \leq \omega_o = 1.256S \\ (\omega_o/\omega)^2 & , \omega > \omega_o = 1.256S \end{cases} \qquad (32)$$

Once the PSD's are obtained, the standard deviations of the dynamic loads may be calculated:

$$\sigma^2_{F_1} = \int_0^\infty \phi_{F_1 F_1}(\omega)\, d\omega \quad , \qquad (33)$$

and

$$\sigma^2_{F_2} = \int_0^\infty \phi_{F_T F_T}(\omega)\, d\omega \quad . \qquad (34)$$

4.3.4 Impedances

The vehicle impedances may be obtained by well-known Fourier Transform techniques and are not presented here. Expressions for the impedance of the rail, on the other hand, are less easily found in the literature; they are:

$$Z_R = -2\sqrt{2}\, i\, \frac{k_{EQ}}{\omega} \left[1 - \left(\frac{\omega}{\omega_{nr}}\right)^2\right]^{3/4} \quad , \omega \leq \omega_{nr} \quad , \qquad (35)$$

and

$$Z_R = 2(1 + i)\, \frac{k_{EQ}}{\omega} \left[\left(\frac{\omega}{\omega_{nr}}\right)^2 - 1\right]^{3/4} \quad , \omega > \omega_{nr} \quad , \qquad (36)$$

where

$$\omega_{nr} = (k_F/m_R)^{1/2} \quad , \tag{37}$$

and

$$k_{EQ} = (EI)^{1/4}(k_F)^{3/4} \quad . \tag{38}$$

In these expressions,

k_F = foundation stiffness, lbf/ft/ft

m_R = rail linear density, slugs/ft

E = Young's modulus of rail material, lbf/ft^2

I = bending moment of inertia of rail, ft^4.

4.4 Numerical Results and Discussion

The PSD's of the dynamic wheel load F_1 (or F_2) and the dynamic axle load F_T were obtained from Eqs. 31 and 32; the standard deviations σ_{F_1} and σ_{F_T} were then obtained by numerical evaluation of the integrals in Eqs. 33 and 34. A small parametric study was made in order to gain insight into the design of vehicles, by varying vehicle and rail parameters. For this parametric study, a baseline vehicle, approximately representing the New Tokaido Line cars, was studied first. The numerical values of the parameters used for this baseline vehicle are shown in Table 4.1.

Figure 4.3 shows the values of $\phi_{F_1F_1}(\omega)$ and $\phi_{F_TF_T}(\omega)$ for the baseline vehicle, as a function of (ω/ω_1), where

$$\omega_1 = (k_1/M_2)^{1/2} \quad , \tag{39}$$

for a speed S = 220 ft/sec (150 mph).

For the same base vehicle, Fig. 4.4 shows the standard deviations σ_{F_1} and σ_{F_T} as a function of vehicle speed.

The frequencies at which the peaks occur in Fig. 4.3 are the car resonance frequency (car bouncing on the secondary suspension) and the wheelset resonance frequency (wheelset bouncing on the track). These frequencies are near the roots of the Equation

$$\left(\frac{k_1+2k^1}{M_1} - \omega^2\right)\left\{\left(\frac{k_2}{M_1} - \frac{M_3}{M_1}\omega^2\right)\left(\frac{k_1+k_2}{M_1} - \frac{M_2}{M_1}\omega^2\right) - \left(\frac{k_2}{M_1}\right)^2\right\}$$

$$- \left(\frac{k_1}{M_1}\right)^2\left(\frac{k_2}{M_1} - \frac{M_3}{M_1}\omega^2\right) = 0 \quad , \tag{40}$$

where

$$k^1 = 2\sqrt{2}\, k_{EQ} \tag{41}$$

represents the track stiffness.

The swift increase in σ_{F_1} and σ_{F_T} at low speeds and at a speed of about 140 ft/sec in Fig. 4.4 can be attributed to the presence of the peaks in the spectrum. As may be seen from Eqs. 30 and 31, at any speed S there is a frequency $\omega_o \approx 1.256S$ at which the PSD begins to fall off. This fall is due to the nature of the PSD of the rail irregularity, Eqs. 29 and 30.

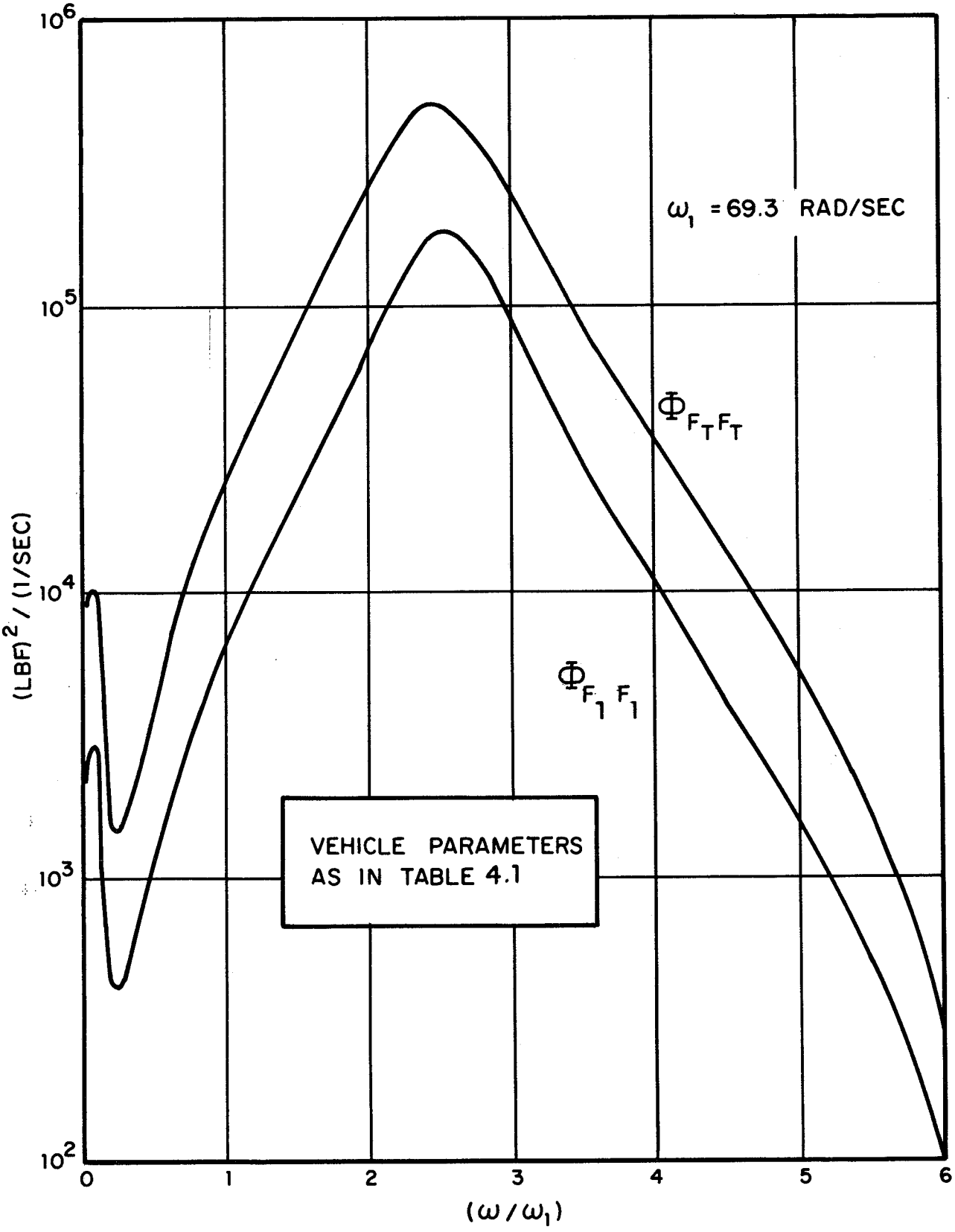


FIG. 4.3 POWER SPECTRA OF THE DYNAMIC LOADS FOR THE BASELINE VEHICLE AT S = 220 FT./SEC.

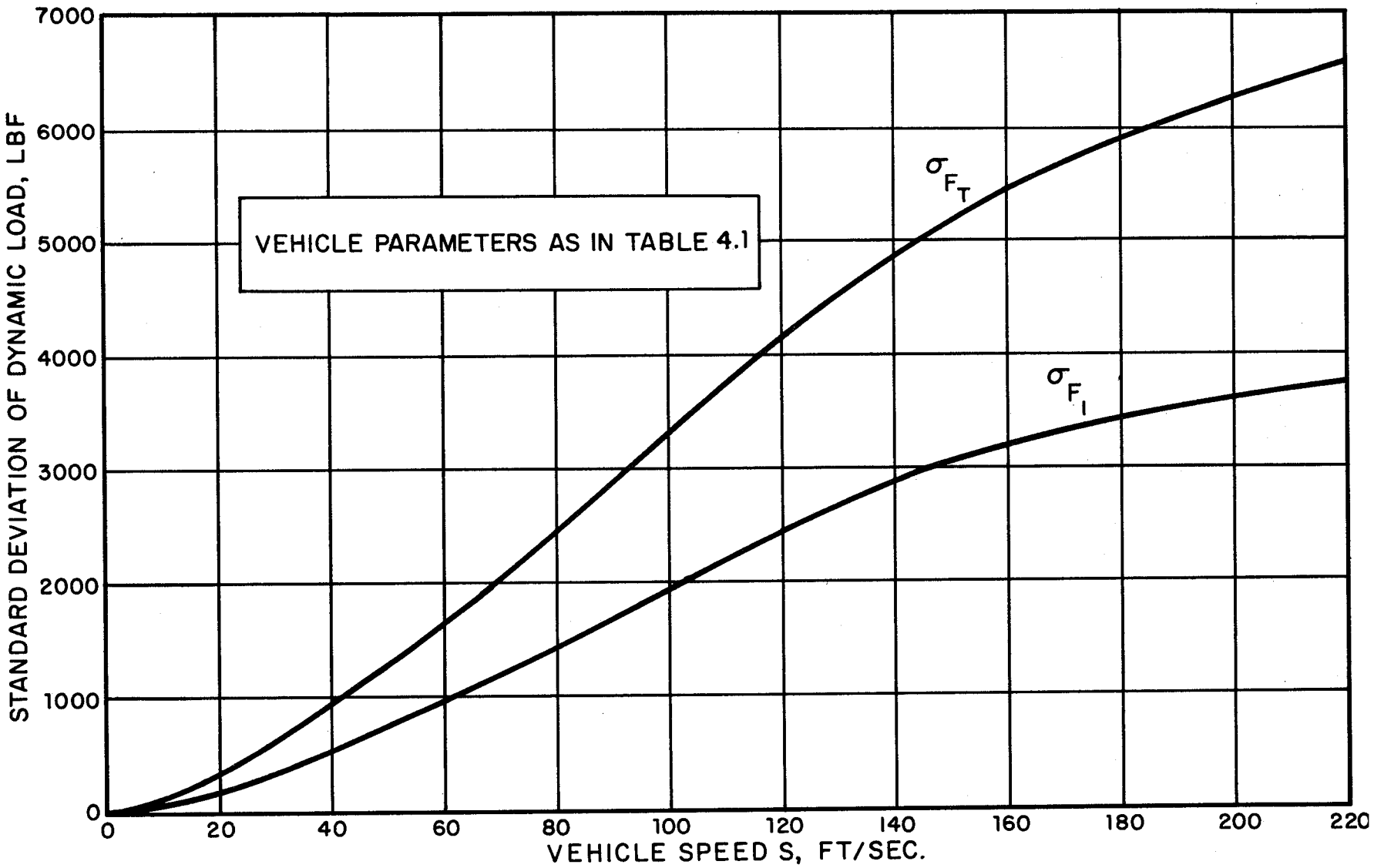


FIG. 4.4 STANDARD DEVIATIONS OF THE DYNAMIC LOADS FOR THE BASELINE VEHICLE.

As the speed S increases, the fall-off frequency ω_o crosses the resonance frequencies, and there is a large input to the σ's from the peaks in the PSD's. This first happens at very low speeds, due to an increase in ω_o beyond the car bounce frequency, and then at a fairly high speed, due to an increase in ω_o beyond the wheelset bounce frequency.

At very high speeds, the standard deviations of the wheel and axle loads increase as $\sqrt{S}$. This behavior may again be deduced from an examination of Eqs. 31-34 and the shape of the PSD's in Fig. 4.3. An increase in S causes the area under the PSD curves to increase in two ways: (a) because the PSD α S, and (b) because the fall-off frequency ω_o increases as S increases. However, at high speeds, the value of the PSD's at $\omega=\omega_o$ is small, and the increase in the area under the PSD curves due to the second cause is small. Since the area under the curves equals σ^2, the standard deviations σ increase as $\sqrt{S}$.

Effect of Truck and Wheelset Mass

As may be seen from Fig. 4.3, the main contribution to the standard deviations σ_{F_1} and σ_{F_T} comes from frequency components near the wheelset bounce frequency. It is therefore obvious that a decrease in the wheelset and truck masses should significantly decrease σ_{F_1} and σ_{F_T}, particularly at high speeds (of the order of 140 ft/sec). This reasoning is borne out by the results shown in Fig. 4.5, which are the standard deviations σ_{F_1} and σ_{F_T} corresponding to a wheelset mass M_1 = 50 slugs and a truck mass M_2 = 50 slugs (compared with $M_1 = M_2 = 62.5$ slugs for the baseline vehicle). A comparison of Figs. 4.4 and 4.5 shows that at S = 220 ft/sec, the decrease in M_1 and M_2 causes a decrease in σ_{F_1} of 10.6% and a decrease in σ_{F_T} of 10.5%. A more important

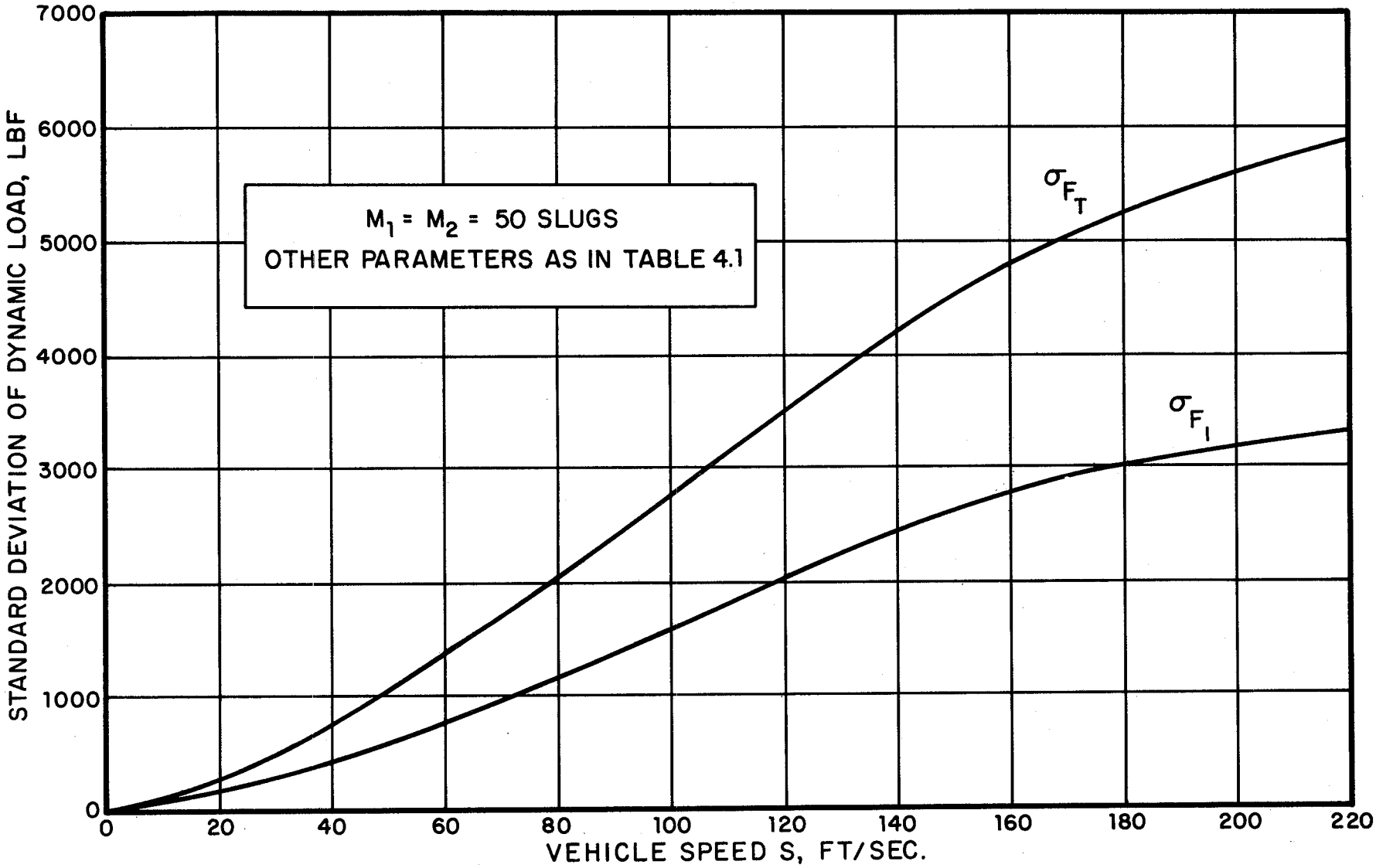


FIG.4.5 EFFECT OF TRUCK AND WHEELSET MASS ON DYNAMIC LOADS

statistic, however, is the ratio of the standard deviation of the load to the mean load. Let

$$N_{m_T} = \text{mean axle load} = 32.2\ (M_1+M_2+M_3)\ \text{lbf}$$

$$N_{m_1} = \text{mean wheel load} = N_{m_T}/2\ \text{lbf}. \tag{42}$$

For the baseline vehicle, $N_{m_T} \approx 27{,}200$ lbf; for the lighter vehicle of Fig. 4.5, $N_{m_T} \approx 26{,}400$ lbf. Thus the decrease in σ_{F_1}/N_{m_1} is about 7.95%, and in σ_{F_T}/N_{m_T}, about 7.6%

Effect of Primary Suspension Damping

At the wheelset resonance, the deflections occur mainly in the track and in the primary suspension. Thus the damping in either should have a significant effect on the standard deviations σ_{F_1} and σ_{F_T}. A study was made of the effect of the primary suspension damping ratio ζ_1 (see Table 4.1 for definition) on σ_{F_1} and σ_{F_T}. Figure 4.6 shows σ_{F_1} and σ_{F_T} as functions of ζ_1 for three different vehicle speeds. It appears from Fig. 4.6 that the baseline value of $\zeta_1 = 1.25$ is not optimal from the viewpoint of dynamic loading for any of the three speeds. A value of ζ_1 between 0.5 and 0.75 seems preferable. It is worth noting that this criterion of optimality for the damping is at variance with the classical criterion [7], which is usually concerned with minimizing the peak value of the spectrum of some quantity, such as the dynamic load, or the car body acceleration. When the spectrum has a very large peak at some frequency, however, the two criteria tend to coincide. This is true in the particular example we are considering. Figure 4.7 shows the maximum values of $\phi_{F_TF_T}(\omega)$ at S = 220 ft/sec as a function of ζ_1. The optimal value of ζ_1 from the point of view of the dynamic load spectra is again between 0.5 and 0.75.

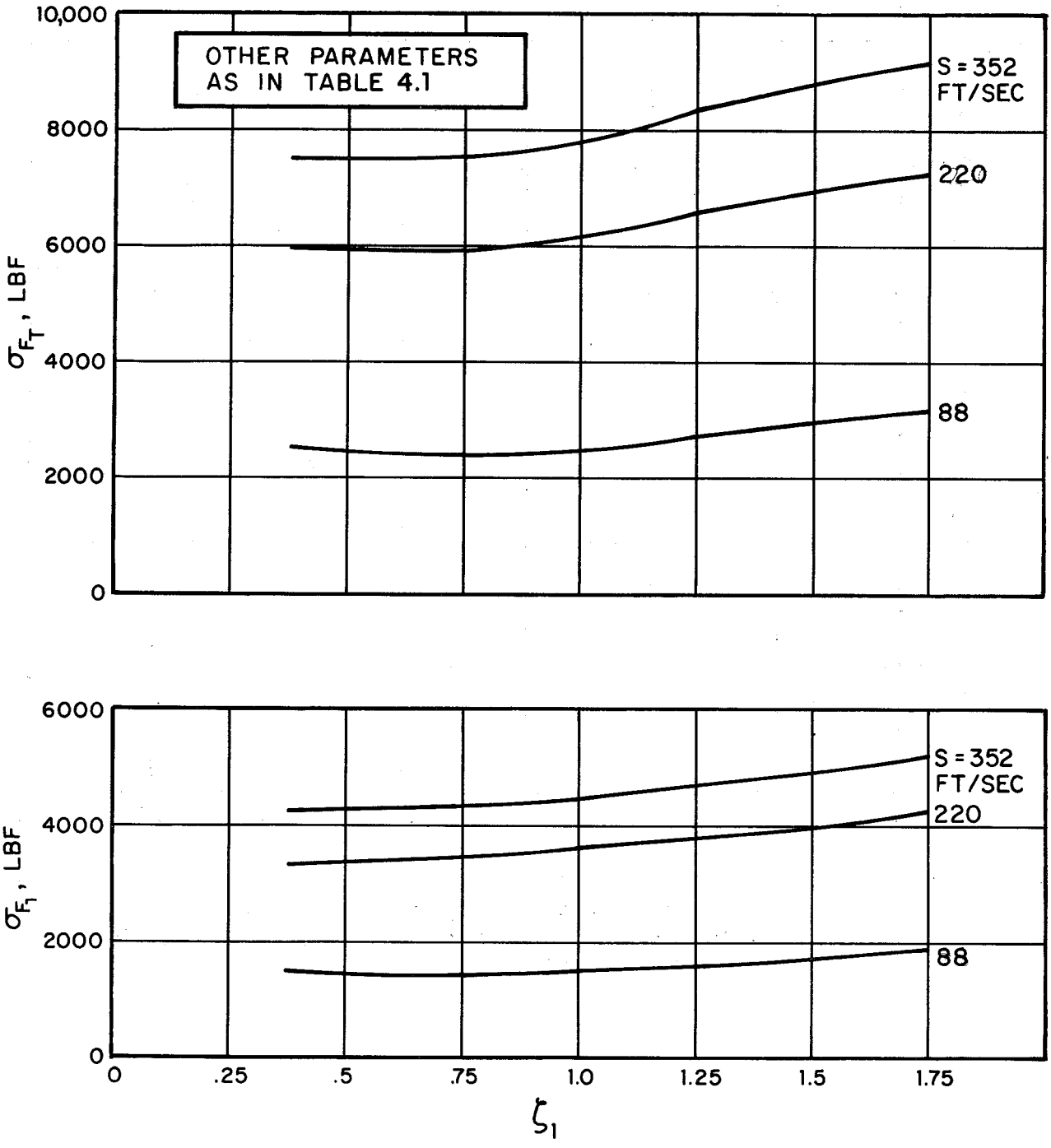


FIG. 4.6 EFFECT OF PRIMARY SUSPENSION DAMPING ON DYNAMIC LOADS.

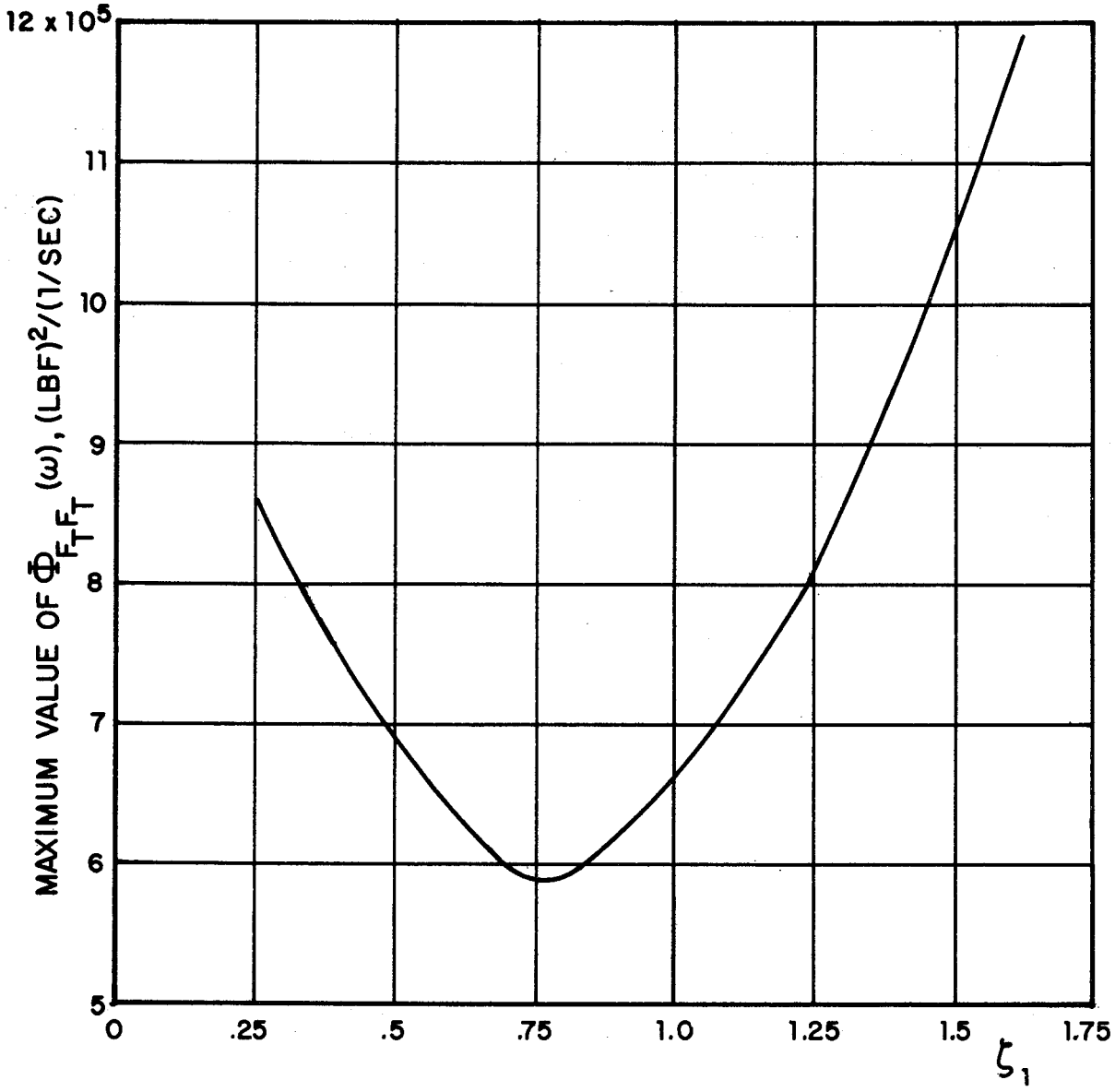


FIG. 4.7 EFFECT OF PRIMARY SUSPENSION DAMPING ON MAXIMUM VALUE OF DYNAMIC LOAD POWER SPECTRUM AT s = 220 FT./SEC.

Effect of Foundation Stiffness

Figure 4.8 shows σ_{F_1} and σ_{F_T} for the baseline vehicle, with a foundation stiffness of $k_F = 7.2\times10^5$ lbf/ft^2 (5000 lbf/in^2). A comparison with Fig. 4.4, where $k_F = 2.16$ lbf/ft^2 (1500 lbf/in^2) shows a marked increase in the dynamic loads: an increase of 88% in σ_{F_1} and 81% in σ_{F_T} at S = 220 ft/sec. This reason is obvious. Despite the increase in the wheelset bounce frequency, the input velocity spectrum due to rail irregularity is the same at this frequency for both foundation stiffnesses. With the stiffer foundation, a smaller portion of this input velocity is taken by the track, and a larger portion by the vehicle, than for the softer foundation. The impedance of the vehicle being the same in both cases, however, the dynamic force is greater for the stiff foundation than for the soft foundation.

The obvious conclusion is that the foundation should be made as soft as possible in the vertical direction. The most important constraints limiting the lowest value of k_f seem to be those imposed by bending stresses in the rail (which increase as k_f decreases), and the effect that large rail deflections may have on the design of the supporting structure.

4.5 Applications

4.5.1 Damage

There are a large variety of ways in which a rail may be damaged. Here we consider two types of damage.

a) The formation of a plastic hinge in the rail, causing a low spot which then leads to high impact loads.

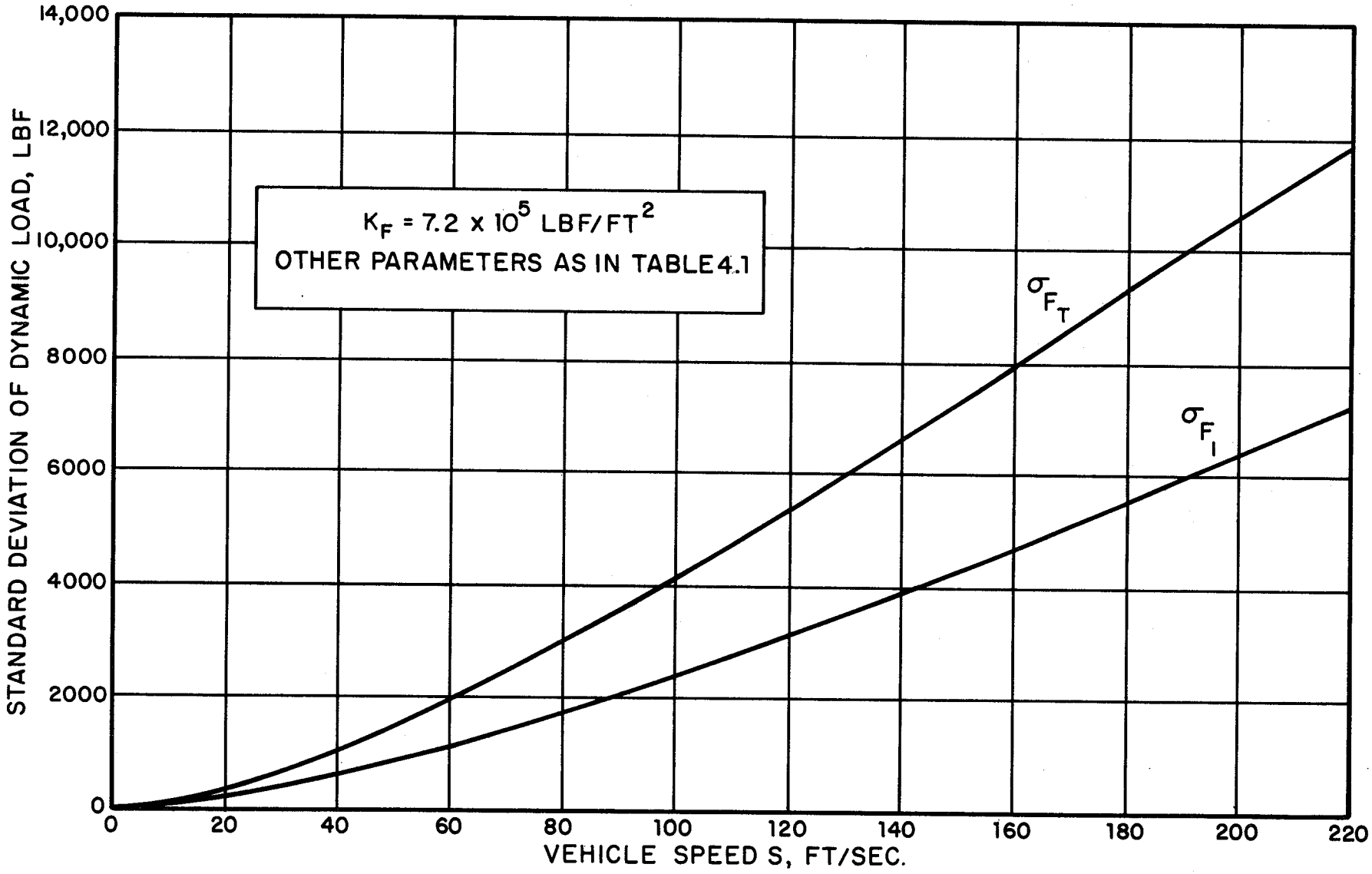


FIG. 4.8 EFFECT OF FOUNDATION STIFFNESS ON DYNAMIC LOADS.

b) Plastic flow in the rail head due to the Hertzian stresses in the contact region between the wheel and the rail.

In both cases, the stress analysis will perforce have to be a static one, since the dynamic solutions to the problems are, to our knowledge, yet to be obtained.

Plastic Hinge

When a transverse load is applied to a beam such that the bending stresses across some section are everywhere equal to the yield stress in tension Y, the curvature at that section increases indefinitely; the beam is said to develop a plastic hinge at the section. The limiting bending moment transmitted across the hinge may be called the plastic moment, M_P. The plastic moment was calculated for 109 AREA rail, and was found to be approximately

$$M_P = 7Y/3 \text{ lbf-ft} \tag{43}$$

where Y is the yield stress in tension, in psi.

When a load N_1 is applied to a beam on an elastic foundation, the maximum bending moment occurs under the load and has the value [8]

$$M_{max} = \frac{N_1}{4} (4EI/k_F)^{1/4} \quad . \tag{44}$$

The meaning of the symbols in Eq. 44 is given in the text immediately following Eq. 38.

For a hinge to form, $M_{max} \geq M_P$, or

$$N_1 \geq 28Y\ (k_F/4EI)^{1/4}/3 \quad . \tag{45}$$

Inserting values for the parameters in Eq. 45 from Table 4.1, and taking Y = 70,000 psi, we obtain the following criterion for formation of a plastic hinge:

$$N_1 \geq 175{,}000 \text{ lbf} \quad . \tag{46}$$

As can be seen from the numerical results for the standard deviation of the wheel load in Fig. 4.4, the formation of a plastic hinge caused by the passage of the example vehicles considered in this paper is an extremely unlikely occurrence.

Plastic Flow

When a rigid flat punch presses against an elastic-perfectly plastic solid, plastic flow will not occur under the punch if

$$\bar{p} < 5k \quad , \tag{47}$$

where $\bar{p}$ is the average pressure and k the yield stress in shear of the material [*9*].

As we are not aware of a lower bound to the limit pressure for the contact of elastic solids, we use the following, slightly conservative criterion: When a wheel presses against a rail, plastic flow in the material around the contact region will not occur if

$$p_{max} < 5k \quad , \tag{48}$$

where p_{max} is the maximum pressure in the contact region, and k the yield stress in shear of the rail.

In order to relate the maximum pressure in the contact region to the wheel load, we consider a wheel with a radius of 15.8 in., and rail with a radius of curvature of 12 in. Then, if N_1 is the wheel load, the following relation may be obtained [*10*]:

$$p_{max} = 6360\ (N_1)^{1/3}\ \text{psi} \quad . \tag{49}$$

The wheel load N_1 is the sum of the mean wheel load N_{m_1} and the dynamic wheel load F_1:

$$N_1 = N_{m_1} + F_1 \quad . \tag{50}$$

The mean wheel load for our baseline vehicle is $N_{m_1} \approx 13{,}600$ lbf. Combining Eqs. 48, 49 and 50, and using the value $k = 35{,}000$ psi, the following criterion is obtained for the occurrence of plastic flow:

$$F_1 \geq F_P = 7000\ \text{lbf} \quad . \tag{51}$$

If we now hypothesize that the dynamic load F_1 has a Gaussian probability density,

$$p(F_1) = \frac{1}{\sigma_{F_1}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\,(F_1/\sigma_{F_1})^2\right] , \tag{52}$$

the probability that F_1 will exceed F_P may be calculated:

$$P(F_1 \geq F_P) = \int_{F_P}^{\infty} p(F_1)dF_1 = \frac{1}{2}\ \text{erfc}\ (F_P/\sigma_{F_1}\sqrt{2})$$

$$= \frac{1}{2}\ \text{erfc}\ (4950/\sigma_{F_1}) \quad . \tag{53}$$

The complementary error function is defined as follows:

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} \exp(-x^2)dx \quad . \tag{54}$$

A shortcoming of the above analysis is that Eq. 53 gives a finite probability of damage whenever $\sigma_{F_1} \neq 0$. This makes it difficult to attach a physical significance to small probabilities of damage.

One way to circumvent this difficulty is to attach a penalty function to the damage. In other words, an exceedance (i.e., an increase of F_1 above F_P) that lasts a small time is less damaging than one that lasts a long time. To be specific, let us assume that an exceedance that is shorter in duration than the time to roll a distance equal to the length of the contact region between the wheel and rail causes no damage.

Let d be the length of contact region; then the time to roll this distance is

$$\tau = d/S \quad . \tag{55}$$

Let us use a penalty function of the form

$$\delta = \begin{cases} (t_e - \tau)/\tau & , \quad t_e \geq \tau \\ \\ 0 & , \quad t_e < \tau \end{cases} , \tag{56}$$

where t_e is the time of an exceedance.

What is needed now is the probability density for exceedance times; though theoretically possible, this, in practice, is difficult to obtain [*11,12*]. We shall therefore deal only with the mean exceedance time $\bar{\tau}_e$.

If we assume again that F_1 has a Gaussian probability distribution, the number of exceedances/unit time is [*13*]

$$n = \frac{1}{2\pi} \frac{\sqrt{m_2}}{\sigma_{F_1}} \exp\left[-\frac{1}{2}(F_P/\sigma_{F_1})^2\right] , \tag{57}$$

where

$$m_2 = \int_0^\infty \phi_{F_1 F_1}(\omega)\omega^2 d\omega . \tag{58}$$

Over a time T, there are nT exceedances. The total time of these exceedances is $T \times P(F_1 \geq F_P)$, where $P(F_1 \geq F_P)$ is given by Eq. 54. Thus the mean exceedance time is

$$\bar{\tau}_e = \frac{P(F_1 \geq F_P)}{n} . \tag{59}$$

Introducing Eqs. 56 and 58 into Eq. 57, we obtain

$$\delta = \frac{SP(F_1 \geq F_P)}{nd} - 1 . \tag{60}$$

The function δ is an estimate of the extent to which the mean length of the damaged zones exceeds the length of the contact

region. The rationale behind the penalty function δ is that very short zones of damage are not likely to be felt by the wheel, as it averages the input irregularity over the length of the contact region, d.

The quantity m_2 (Eq. 58) was evaluated numerically for the baseline vehicle. The number of exceedances/sec, n, was then obtained from Eq. 57, and is shown in Fig. 4.9 as a function of the vehicle speed, S. The penalty function δ was then obtained from Eq. 60, and is also shown in Fig. 4.9. The contact width d may be estimated approximately from the following equation [*10*]

$$d \approx 1.57\times10^{-3} \, (Nm_1)^{1/3} \quad \text{ft.}$$

For the baseline vehicle, $d \approx 3.78\times10^{-2}$ ft.

It may be seen from Fig. 4.8 that for speeds above $S \approx 60$ ft/sec, the penalty function δ is greater than zero; the average length of the damaged zone is greater than the contact width d. Also, the length of the damaged zone increases almost linearly with the vehicle speed S. Moreover, the number of damaged zones per second increases rapidly as S increases. Another interesting statistic is the number of damaged zones/unit length of track, which is n/S. This quantity is also shown in Fig. 4.9. Thus, it may be seen that at S = 220 ft/sec, there is approximately one damaged zone every 30 ft.

The most important ways in which both the frequency and length of the damaged zones may be reduced are (a) by decreasing the mass of the wheelset and truck and (b) by decreasing the

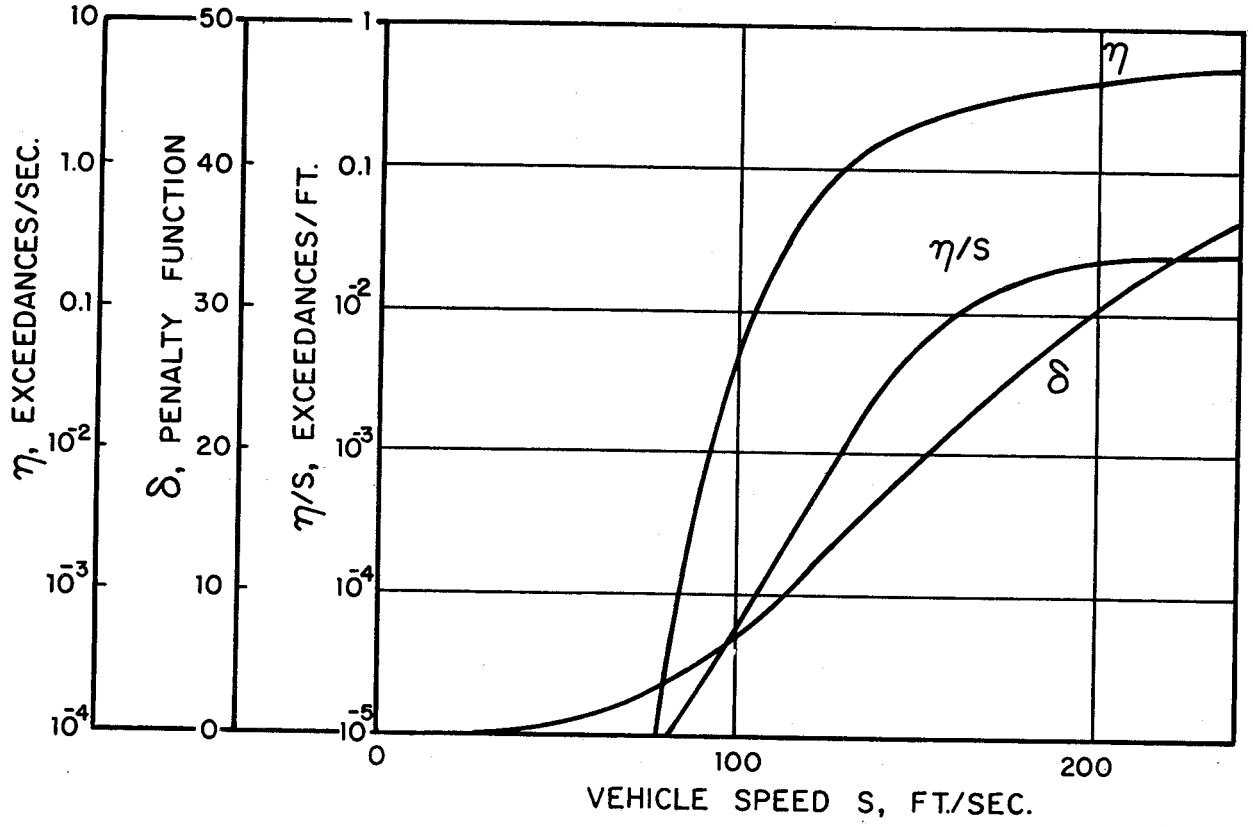


FIG. 4.9 STATISTICS OF RAIL DAMAGE FOR BASELINE VEHICLE, PARAMETERS AS IN TABLE 4.1

foundation stiffness, k_f. The quantity m_2 may be approximated by

$$m_2 \approx \sigma_{F_1}^2 \omega_{RES}^2 \quad ,$$

where ω_{RES} is the wheelset resonance frequency. Therefore,

$$n \approx \frac{\omega_{RES}}{2\pi} \exp\left[-\frac{1}{2}\left(\frac{F_P}{\sigma_{F_1}}\right)^2\right] \quad .$$

Thus, n may be reduced either by reducing σ_{F_1} or by reducing σ_{RES}.

4.5.2 Tractive capacity

It has been observed [*14*] that the tractive capacity of a wheel, i.e., its ability to transmit a tractive force without gross slip decreases as the rolling velocity of the wheel increases. The usual explanation offered for this behavior is that as the speed increases, the dynamic load between the wheel and rail also increases. This dynamic load is assumed to decrease the effective normal load transmitted by the wheel, the reasoning being as follows: suppose the friction coefficient between the wheel and rail is f_R, and the mean normal load on the wheelset N_{m_T}; then at low speeds, the tractive force that may be transmitted is $T_R = f_R N_{m_T}$. At high speeds, however, the normal load varies about N_{m_T}. When the normal load is less than N_{m_T}, an attempt to transmit a tractive force T_R will cuase the axle to slip. In the following, we present an analysis of the braking of a wheel which seems to show the above explanation to be incorrect. Alternate mechanisms are also postulated.

Consider an axle moving at speed S_o over a track. Brake shoes are pressed against the wheel rims with a normal force F_B, producing a frictional force $T_B = f_B F_B$, where f_B is the friction coefficient between the brake block and the wheel rim, Fig. 4.10. The mass of the wheelset is M_1, and that of the whole vehicle, M_T. The equations of motion for the wheelset and the vehicle are

$$I\ \ddot{\theta} = R(T_R - f_B F_B) \tag{61}$$

and

$$M_T\ \ddot{X} = -\ T_R \quad , \tag{62}$$

where I is the rotary inertia of the wheelset. Furthermore, if no slip occurs,

$$R\ \ddot{\theta} = \ddot{X} \quad . \tag{63}$$

Equations 61-63 may be solved to obtain T_R and F_B in terms of the deceleration, $a = |\ \ddot{X}\ |$:

$$T_R = M_T\ a\ ,\ F_B = \frac{a}{f_B}\left(M_T + \frac{I}{R^2}\right) \quad . \tag{64}$$

Also,

$$F_B = \frac{1}{f_B}\left(1 + \frac{I}{M_T R^2}\right) T_R \quad . \tag{65}$$

We have assumed that the wheelset is capable of transmitting the force T_R without gross slip. Suppose now that the force T_R is the maximum available while the normal load between the wheelset

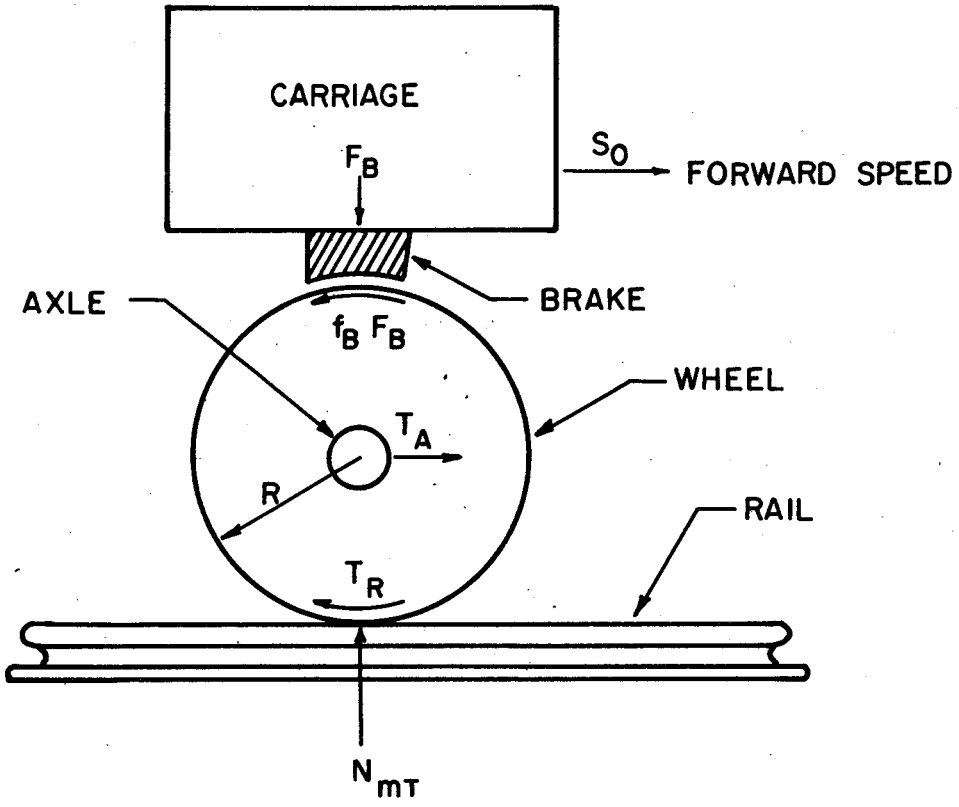


FIG. 4.10 FORCES ACTING ON A WHEELSET BEING BRAKED

and the track is a constant. Then

$$T_R = f_R N_{m_T} \quad , \tag{66}$$

and from Eq. 65, assuming $I << M_T R^2$,

$$F_B = \frac{f_R N_{m_T}}{f_B} \quad . \tag{67}$$

Suppose now that the same brake force is applied while the wheel-set load fluctuates about its mean value:

$$N_T = N_{m_T} - F_T \sin\omega t \quad . \tag{68}$$

The maximum tractive force that can be transmitted at any instant is

$$T_{R,max} = f_R\left[N_{m_T} - F_T \sin\omega t\right] \quad . \tag{69}$$

There are now periods of time when the required tractive force T_R, given by Eq. 66, exceeds the maximum available tractive force, given by Eq. 69, and the wheel slips. As soon as the wheel slips, its motion is no longer governed by Eq. 61, but by

$$I\ddot{\theta} = R(f_R N_T - f_B F_B) \quad . \tag{70}$$

We shall now attempt to determine the value to which the brake force F_B should be reduced from that given in Eq. 67 in order that the wheel slip averaged over time be zero.

It may be noted that the wheelset decelerates when $\sin\omega t>0$ in Eq. 69, but it accelerates again when $\sin\omega t<0$; the wheelset would not slip on the average were it not for the fact that the friction coefficient between the wheel and the rail is a function of the slip velocity in the interface [*15*]. Assuming that over a short time, the forward speed of the wheelset remains at its initial value S_o, the slip velocity is

$$V_S = S_o - R\dot{\theta} \quad . \tag{71}$$

The friction coefficient may then be approximated by [*15*]

$$f_R = f_{R_\infty}\left[1 + \exp\left(-\frac{S_o - R\dot{\theta}}{S^*}\right)\right] \quad , \tag{72}$$

where S^* is a reference velocity. Initially, $S_o = R\theta$, and $f_R = 2f_{R_\infty}$. Thus the brake force without dynamic loading is, from Eq. 67,

$$F_B = \frac{2f_{R_\infty} N m_T}{f_B} \tag{73}$$

On the other hand, experimental data seem to indicate that the friction coefficient between the brake block and the wheel rim is a constant [*16*]:

$$f_B = \text{constant} \quad . \tag{74}$$

Suppose now that the actual brake force is reduced from the value given by Eq. 74 to the following value:

$$F_B = \frac{2f_{R_\infty} N_{m_T}(1-\varepsilon)}{f_B} \quad . \tag{75}$$

We shall now determine the value of ε such that no gross slip occurs on the average. Then the effective normal load is $N_{m_T}(1-\varepsilon)$. Combining Eqs. 68, 70, 72 and 74, we obtain

$$I\ddot{\theta} = Rf_{R_\infty} N_{m_T} \left\{ \left[1 - \frac{F_r}{N_{m_T}} \sin\omega t \right] \left[1 + \exp\left(- \frac{S_o - R\dot{\theta}}{S^*} \right) \right] - 2(1-\varepsilon) \right\} . \tag{76}$$

As it does not appear that this nonlinear equation can be solved with ease, we replace the sine-wave by a rectangular wave. Assuming that the wheel is not slipping at t=0, we see from Eq. 76 that the wheel will not begin to decelerate until

$$\sin\omega t > \frac{\varepsilon N_{m_T}}{F_T} \quad ,$$

or

$$t > t_o = \frac{1}{\omega} \sin^{-1} \left(\frac{\varepsilon N_{m_T}}{F_T} \right) \quad . \tag{77}$$

Thus, the quantity $1 - (F_T/N_m)\sin\omega t$ is replaced by the rectangular wave shown in Fig. 4.11; the time during which it decelerates is less than the time during which it accelerates, thus making it possible for the slip to be zero on the average.

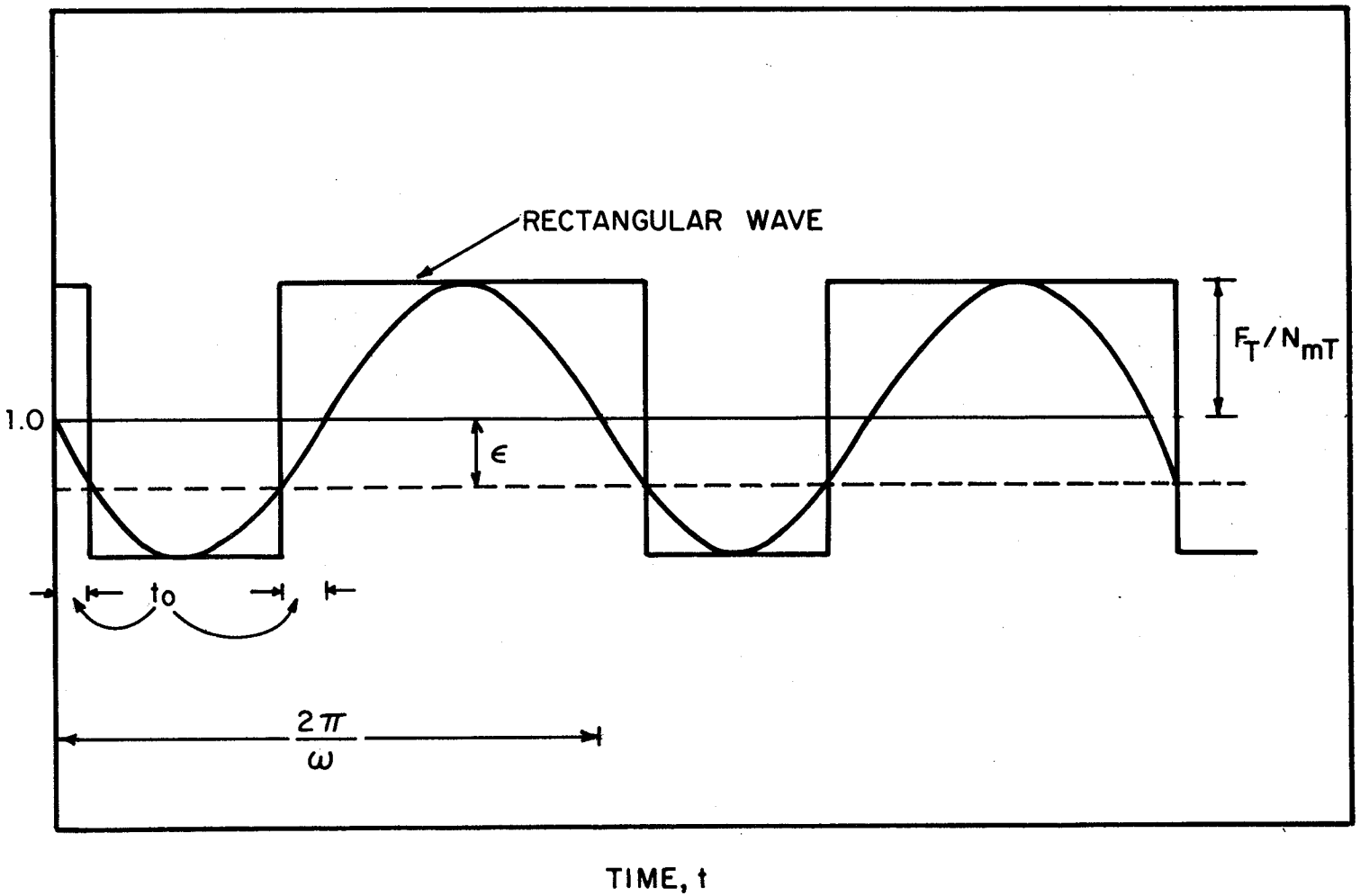


FIG. 4.11 REPLACEMENT OF A SINE-WAVE DYNAMIC LOAD BY A RECTANGULAR WAVE FOR THE ANALYSIS OF THE BRAKING OF A WHEELSET.

Requiring

$$R\dot{\theta} = S_o \text{ at } t = t_o \text{ and at } t = t_o + 2\pi/\omega \quad , \tag{78}$$

the equation of motion can now be solved to obtain ε as a function of ω and of

$$\left.\begin{aligned} a_1 &\equiv IS^*/\, R^2 f_{R_\infty} N_{m_T} \quad , \\ \text{and} \\ a_2 &\equiv F_T/N_{m_T} \quad . \end{aligned}\right\} \tag{79}$$

It is worth noting that because of the form of the initial and final conditions, Eq. 78, ε does not depend on S_o. Thus S_o acts only through its influence on the load ratio a_2.

Figure 4.12 shows ε as a function of $f = (\omega/2\pi)$ for $a_1 = .25,1$ and $a_2 = 0.25,1$. As a_1 is increased or a_2 decreased, ε decreases at any frequency.

It may be seen that except when the dynamic load frequency is very low, and the dynamic load very high — a combination that almost never occurs — the decrease in the effective normal load is of the order of 5%. In order to interpret these results, however, we must examine the constant a_1. When a_1 is very small, it is possible for the decrease in the effective normal load to be large, especially at low frequencies.

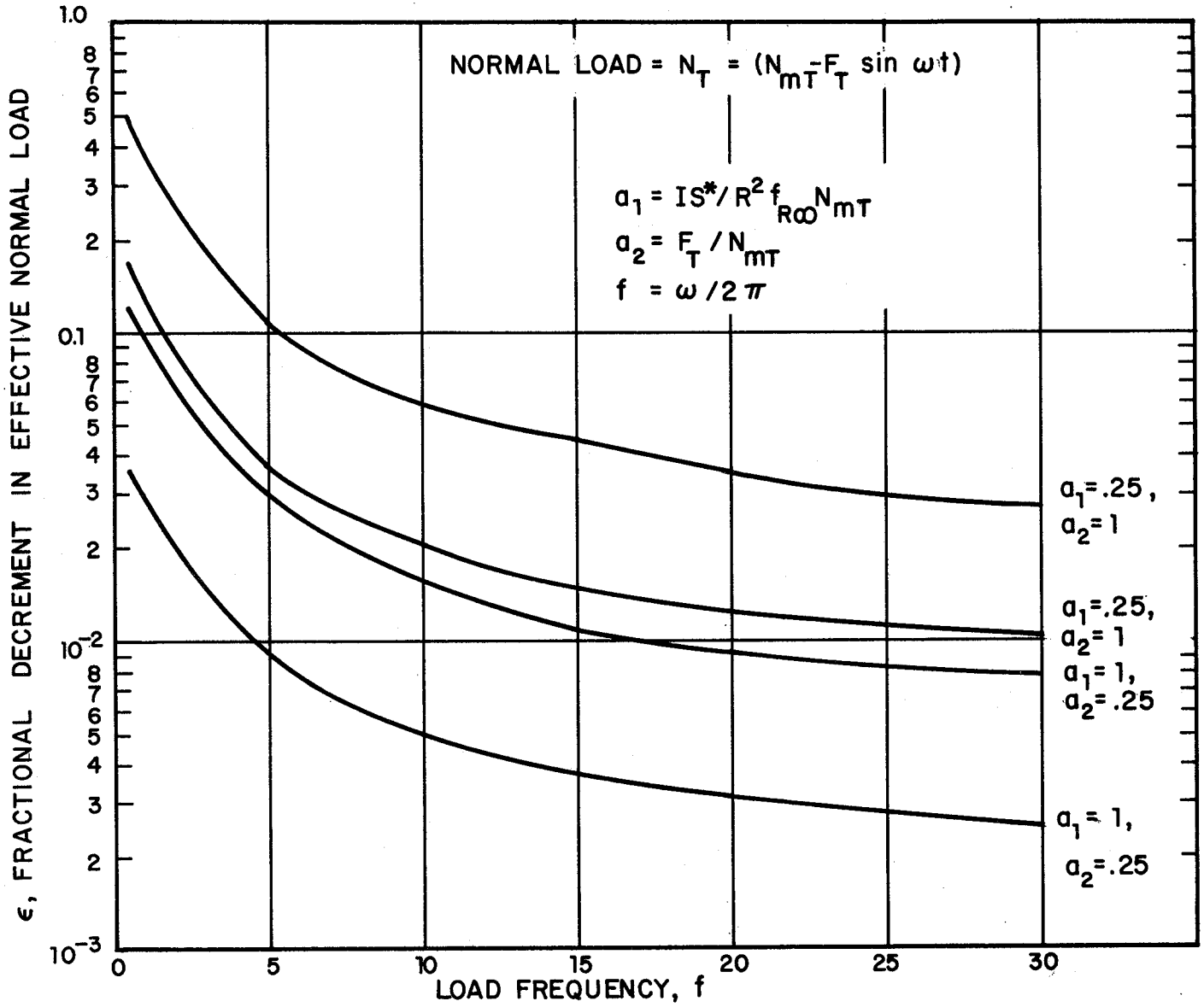


FIG. 4.12 CHANGE IN EFFECTIVE NORMAL LOAD DUE TO DYNAMIC EFFECTS.

Estimating $I \approx 50$ slug-ft^2, $R = 1.5$ ft, $f_{R_\infty} \approx 0.15$ and $N_{m_T} = 27{,}200$ lbf, we have, from Eq. 79,

$$a_1 \approx 1.48\times10^{-2}S^* \quad . \tag{80}$$

For a_1 to be 0.25, we must have $S^* \approx 17$ ft/sec. Values observed for S^* in experimental investigations of sliding friction, however, are on the order of hundreds of feet per second [*15,17*]. However, the possibility remains that in rolling contact, the value of S^* is much smaller than in sliding contact, and this possibility deserves investigation. If S^* values in rolling contact are large, however, it appears that the direct effect of dynamic loads on the effective normal load for tractive purposes is negligible.

Before making this a conclusive statement, we shall briefly examine the case when the dynamic load F_T is a random variable, instead of a sinusoid as we had assumed in the foregoing analysis. Suppose that the brake force F_B is that given by Eq. 67; this is the largest force that could conceivably be applied under any conditions. Suppose now that at $t=0$ the normal load on the wheelset drops from $N_T=N_m$ to $N_T=0$. Then the time for the wheelset to lock may be obtained from Eq. 70:

$$T_{lock} = \frac{IS_o}{R^2 f_B F_B} \quad , \tag{81}$$

where S_o is the forward speed of the vehicle. Noting from Eq. 72 that the no-slip friction coefficient is $f_R = 2f_{R_\infty}$, we have

$f_B F_B = 2 f_{R_\infty} N_{m_T}$. Using the same values as before, Eq. 81 gives

$$T_{lock} \approx 7.4 \times 10^{-3} \, S_o \quad \text{seconds} \tag{82}$$

If the normal load were to fall to a nonzero level, the time to lock would be greater than that given by Eq. 82. The question now arises as to whether the total wheelset load,

$$N_T = N_{m_T} + F_T \quad ,$$

makes excursions to low values, the length of the excursion being of the order of T_{lock}. This is the same problem as that of exceedance times, examined in our analysis of rail damage. As remarked at the end of that analysis, the number of excursions/sec of F_T below a value -F may be approximately written

$$n \approx \frac{\omega_{res}}{2\pi} \exp\left[- \frac{1}{2} \left(\frac{F}{\sigma_{F_T}} \right)^2 \right] \quad , \tag{83}$$

where ω_{res} is the wheelset resonance frequency. We have, approximately, $\frac{\omega_{res}}{2\pi} \approx 30$. The average time of an excursion below the value N is

$$\bar{t}_e = \frac{\text{Probability } [F_T < -F]}{n} \quad . \tag{84}$$

The quantities n and $\bar{t}_e$ are shown in Fig. 4.13 for various values of F/σ_{F_T}. For the baseline vehicle studied earlier, N_{m_T} = 27,200 lbf; also, σ_{F_T} = 1050 lbf at S_o = 44 ft/sec and

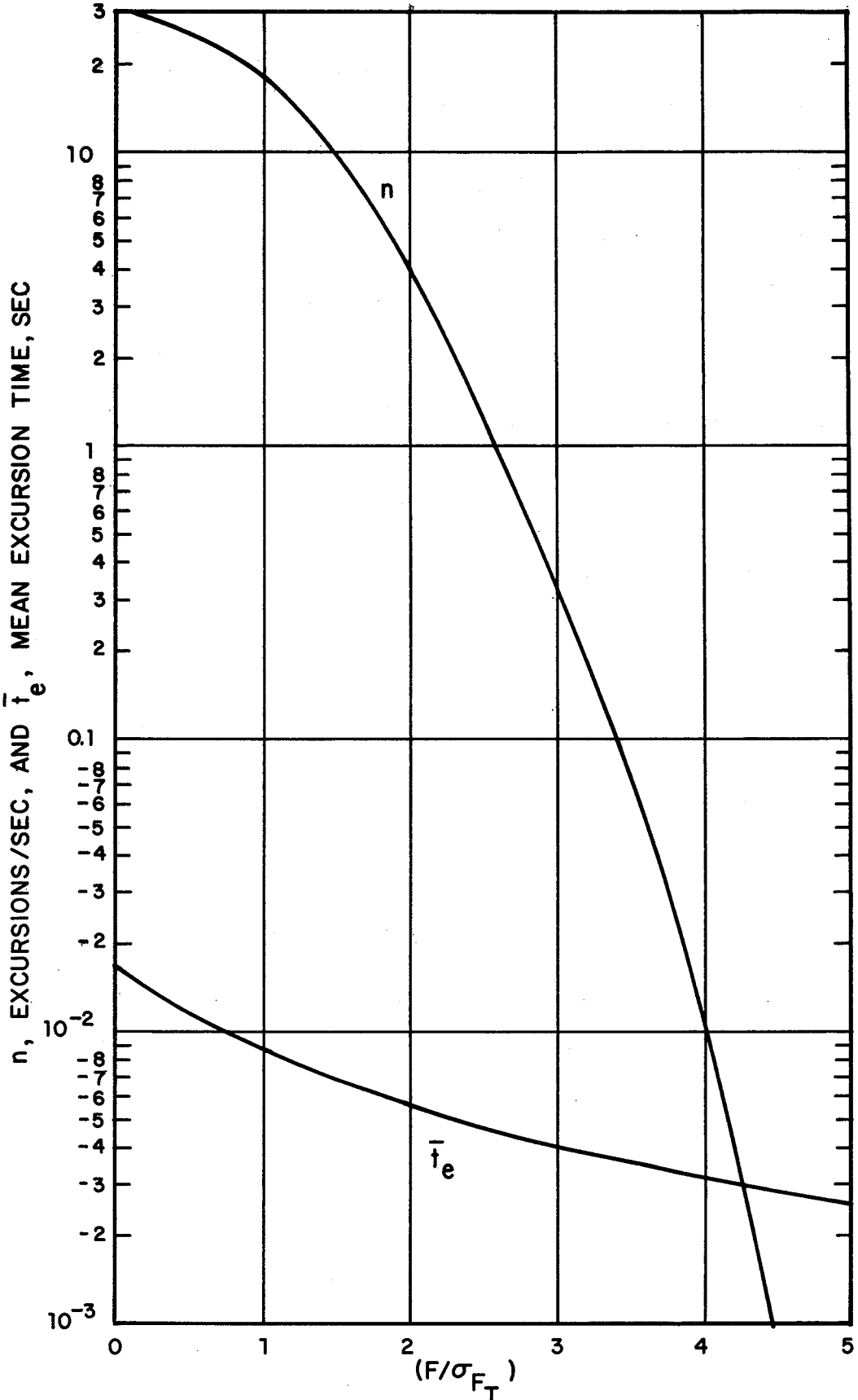


FIG. 4.13 DYNAMIC LOAD STATISTICS FOR ANALYSIS OF BRAKING ON RANDOM TRACK

σ_{F_T} = 6550 lbf at S_o = 220 ft/sec. Thus wheel hop occurs at F/σ_{F_T} = 25.1 at 44 ft/sec and F/σ_{F_T} = 4.15 at 220 ft/sec. The time to lock with wheel hop is found from Eq. 82:

$$T_{lock} = 0.326 \text{ sec at } 44 \text{ ft/sec}$$

$$= 1.63 \text{ sec at } 220 \text{ ft/sec.}$$

A comparison with Fig. 4.13 shows that wheel locking is very unlikely to occur. For $F > 2.5\ \sigma_{F_T}$, there is less than 1 excursion per second, and its mean length is less than 0.005 sec. It must be remembered that we have discussed only *mean* excursion times; the possibility remains that there will be excursions much larger than the mean. The possibility seems remote, however, that there will be excursions to high values of F comparable to T_{lock}.

We conclude, therefore, that dynamic loads have a negligible effect on the effective normal load during the braking (or acceleration) of a wheelset, if sliding friction data for steel on steel are valid for rolling contact.

Thus, the question of why the tractive capacity of wheels decreases as their speed increases still remains to be resolved.

There seem to be four possible explanations, which need to be investigated in detail:

1. Based on the discussion of braking with a sinusoidal dynamic load, it is possible that the value of S* in Eq. 72 is very small. In other words, the friction coefficient between a rolling wheel and rail may decrease rapidly as the slip velocity at the interface increases. This possibility could be studied in a laboratory investigation of rolling contact, with the angular velocity of the rolling elements being controlled, rather than the torque being applied to them.

The transmitted torque would then be measured as a function of the slip velocity. It would be important to contaminate the rolling surfaces with the same contaminants that are found on rails. Other factors that might enter are the surface waviness of the rollers, the normal load being transmitted, the temperature of the surfaces and the geometry of the contact.

2. Another possibility is that as the speed of rolling increases, the time spent by surface asperities in the contact region decreases. These surface asperities form friction junctions which transmit the tractive force, and it is possible that the strength of these junctions decreases as their time of formation and separation decreases. This would result in a decreased tractive capacity. Some experimental evidence of this is presented in Chap. 3; however, it is desirable to conduct a detailed investigation of the behavior of friction junctions at high speeds.

3. Apart from the decrease in the junction strength due to the decrease in the time available for the formation of an adhesive bond as the rolling velocity increases, there remains the possibility that increased surface vibrations decrease the strength of the junctions. There are two mechanisms that might be operating: surface vibrations due to dynamic load effects, and asperity vibrations due to impact when entering the contact region. Both of these vibrations increase as the rolling speed increases. The first mechanism could be studied in conjunction with the study of friction in rolling contact described above, and the second mechanism, in conjunction with the study of the high-speed deformation of friction junctions.

4. As the rolling velocity increases, elastohydrodynamic effects due to contaminant films on the rail or wheel may become important. These effects have been studied in some detail in the laboratory, and an important parameter is the thickness of the film. It is therefore necessary to make field measurements of thicknesses of contaminant films on rails. Should the thicknesses be appropriate for elastohydrodynamic effects to be operative, it would then be important to study the viscosity of the contaminants as a function of pressure and temperature.

TABLE 4.1 Baseline Vehicle and Track Parameters

Parameter	Symbol	Value	Definition
Car Body Mass	M_3	720 slugs	
Car Body Rotational Inertia	I_3	7200 slug-ft^2	
Truck Mass	M_2	62.5 slugs	
Truck Rotational Inertia	I_2	195.3 slug-ft^2	
Wheelset Mass	M_1	62.5 slugs	
Wheelset Rotational Inertia	I_1	390.6 slug-ft^2	
	a	2.5 ft	
	b	2.5 ft	
Rail-To-Rail Distance	ℓ	2.5 ft	
Primary Suspension Stiffness	k_1	3×10^5 lbf/ft	
Primary Suspension Damping	c_1		
Primary Suspension Damping Ratio	ζ_1	1.25	$\zeta_1 = c_1/2(k_1M_1)^{1/2}$
Secondary Suspension Stiffness	k_2	3.12×10^4 lbf/ft	
Secondary Suspension Damping	c_2		
Secondary Suspension Damping Ratio	ζ_2	0.20	$\zeta_2 = c_2/2(k_2M_2)^{1/2}$
Foundation Stiffness	k_f	2.16×10^5 lbf/ft^2	
Rail Linear Density	ρ	1.13 slugs/ft	
Rail Bending Rigidity	EI	1.15×10^7 lbf-ft^2	

REFERENCES

1. A. Kaplan, L.R. Koval, and F. Choy, Jr., "Rail Vehicle Dynamic Analysis for High Speed Ground Transportation Roadbed Study," TRW Systems Report EM-17-22 (Sept. 1967).

2. A.A. Liepins, "Digital Computer Simulation of Railroad Freight Car Rocking," *J. Eng. for Ind.*, 701 (Nov. 1968).

3. J.L. Sewall, R.V. Parrish, and B.J. Durling, "Rail Vehicle Dynamic Studies," Presented at the 40th Shock and Vibration Symposium, Hampton, Virginia (Oct. 1969).

4. E.K. Bender, U.J. Kurze, P.R. Nayak, and E.E. Ungar, "Effects of Rail-Fastener Stiffness on Vibration Transmitted to Buildings Adjacent to Subways," BBN Report No. 1832, submitted to the Washington Metropolitan Transit Authority (Aug. 1969).

5. Y.W. Lee, *Statistical Theory of Communication*, John Wiley and Sons, Inc., New York (1960).

6. T.G. Pearce and B.J. May, "A Study of the Lateral Stability, Curving and Dynamic Response of the Linear Induction Motor Test Vehicle," British Railways Research Department, Advanced Projects Division, Report submitted to the U.S. Dept. of Transportation, Office of High Speed Ground Transportation (Sept. 1969).

7. J.P. Den Hartog, *Mechanical Vibrations*, McGraw-Hill Book Company, Inc., New York (1956).

8. J.P. Den Hartog, *Advanced Strength of Materials*, McGraw-Hill Book Company, Inc., New York (1952).

9. E. Levin, "Indentation Pressure of a Smooth Circular Punch," *Quarterly of Applied Mathematics* 13, 133 (1955).

10. S. Timoshenko and J.N. Goodier, *Theory of Elasticity*, McGraw-Hill Book Company, Inc., New York (1951).

11. S.O. Rice, "Distribution of the Duration of Fades in Radio Transmission: Gaussian Noise Model," *The Bell System Technical Journal* 37 (3), 581 (May 1958).

12. E. Wong, "The Distribution of Intervals Between Zeros for a Stationary Gaussian Process," *SIAM J. Appl. Math.* 18 (1), 67 (Jan. 1970).

13. M.S. Longuet-Higgins, "The Statistical Geometry of Random Surfaces," in *Hydrodynamic Stability, Proc. 13th Symposium on Applied Mathematics,* American Math. Society (1962).

14. "Some Problems of Wheel/Rail Interaction Associated with High-Speed Trains," TRW Systems Report submitted to the U.S. Dept. of Transportation, Office of High Speed Ground Transportation (March 1969).

15. F.P. Bowden and D. Tabor, *The Friction and Lubrication of Solids, Vol. II,* Oxford University Press, p. 475 (1964).

16. H.R. Broadbent, *An Introduction to Railway Braking,* Chapman and Hall Ltd., London, p. 38 (1969).

17. K. Williams and E. Griffen, "Friction Between Unlubricated Steel Surfaces at Sliding Speeds up to 750 Feet per Second," *Proc. Instn. Mech. Engrs.* 178, Part 3N, 24 (1963-1964).

5. RANDOM PROCESS MODEL OF ROUGH SURFACES

Summary

Investigations to date of the importance of surface roughness in interfacial phenomena such as friction and wear have been seriously hampered by the lack of a proper characterization of surface roughness. Another problem has been the inability (except at great cost) to distinguish between the characteristics of the surface and the characteristics of a profile of the surface, such as is obtained with a profilometer.

A mathematical model of surface roughness which overcomes both difficulties is presented here. A three-parameter characterization of surface roughness is shown to be adequate for the investigation of most interfacial phenomena. Next, it is shown that a profile may seriously misrepresent the surface in certain crucial respects, as, for example, the distribution of asperity heights. As has been shown in Chaps. 2-4, such surface statistics are likely to be important in rolling contact under certain circumstances likely to be encountered in practice. For example, when a surface is heavily contaminated, its frictional characteristics depend to a large degree on the probability of high asperities breaking through the contaminant film and making dry contact. This probability, in turn, depends on the height distribution of the surface asperities.

Finally, a technique for predicting surface parameters from a profile is described. A simple electronic technique for subjecting a profile to the necessary analysis is also presented.

5.1 Introduction

The characterization of the topography of solid surfaces is of interest in the study of a number of interfacial phenomena such as friction and wear, and electrical and thermal contact resistance.

A very general typology of solid surfaces is shown in Fig. 5.1. Surfaces that are deterministic may be studied by relatively simple analytical and empirical methods; their detailed characterization is straightforward. However, many engineering surfaces are random; and it is these that have been subjected to a great deal of study in the past decade.

In this paper, attention is concentrated on random, isotropic, Gaussian surfaces, although extensions of the theory to nonisotropic surfaces are indicated. It is clear [*1*] that many surfaces are nonGaussian; but it is equally clear that many surfaces are Gaussian [*1*]. Moreover, a study of Gaussian surfaces should provide a good preparatory background for the study of nonGaussian surfaces.

Our approach is to use the techniques of random process theory. The height of a rough surface may be considered to be a two-dimensional random variable, with the Cartesian coordinates in a reference surface being the independent variables. This approach was first used by Longuet-Higgins in 1957 in two outstanding studies of random ocean surfaces [*2,3*]. Some of the results reported in the following Sections are from Longuet-Higgins' work, and, in these instances, we shall refer the reader to his work for proofs of statements.

A brief survey of the literature on surface mechanics has revealed a number of attempts to analyze solid surfaces with the

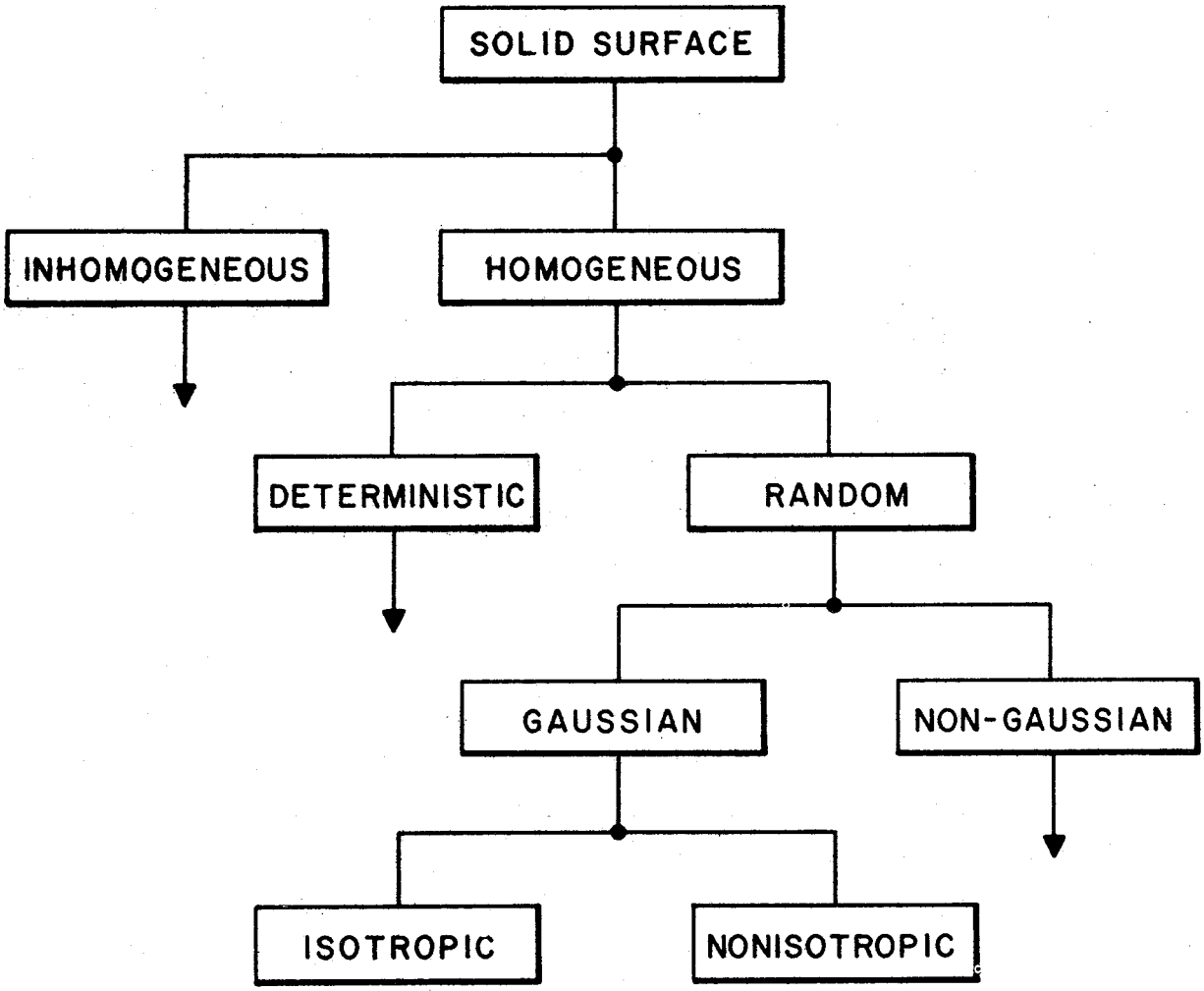


FIG. 5.1 A GENERAL TYPOLOGY OF SURFACES

techniques of random processes [4-6]. All these analyses, however, rest on two assumptions, both of which are shown to be unnecessary in the present work: (1) The statistics of the surface are the same as the statistics of a profile of the surface, and (2) the asperities have spherical caps. The first assumption is found in the following pages to lead to serious error. It is necessary to distinguish a peak on a profile from a summit on the surface, to use the terminology of Williamson and Hunt [7]. A profile will more often than not pass over the shoulder of an asperity on the surface instead of its summit. The shoulder will, nevertheless, appear as a peak on the profile, though one of reduced height. Thus, the profile indicates the presence of far fewer high peaks than actually exist on the surface. A similar error occurs in the determination of the mean surface gradient; it is not the same as the mean slope on a profile. Both of the above assumptions are dropped in the present work.

In Sec. 5.2, some of the relevant results of random process theory are summarized. Further details may be found in the work of Longuet-Higgins [2,3]. In Sec. 5.3, various results are obtained for random, isotropic, Gaussian surfaces. In Sec. 5.4, the sampling of such surfaces by profilometry is analyzed. Results are given for the statistics of profiles, and a comparison is made of surface and profile statistics. In addition, a very simple technique is described for obtaining all the parameters necessary for the analysis of Sec. 5.3. Some conclusions are presented in Sec. 5.5, together with indications of extensions of the theory to nonisotropic Gaussian surfaces.

5.2 The Characterization of Random Processes [2]

5.2.1 The autocorrelation function

Consider a rough surface whose height above a plane reference surface is $z(x,y)$, where z is a random variable and x,y are Cartesian coordinates in the reference surface, Fig. 5.2. The reference surface may be taken to be the mean plane of the rough surface. We assume that the surface is homogeneous — that is, its statistical description is invariant with respect to translation along the surface. The autocorrelation function is then defined to be

$$R(x,y) = \lim_{\substack{L_1\to\infty \\ L_2\to\infty}} \frac{1}{4L_1L_2} \int_{-L_1}^{L_1}\int_{-L_2}^{L_2} z(x_1,y_1)z(x_1+x,y_1+y)dx_1dy_1 \ . \qquad (1)$$

If the surface is isotropic, R depends only on $r = (x^2+y^2)^{1/2}$, and not on the polar angle $0 = \tan^{-1}(y/x)$.

5.2.2 The power spectral density (PSD)

The Fourier transform of R is called the power spectral density:

$$\Phi(k_x,k_y) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} R(x,y)\exp[-i(xk_x+yk_y)]dxdy \ . \qquad (2)$$

The inverse Fourier relation holds:

$$R(x,y) = \iint_{-\infty}^{\infty} \Phi(k_x,k_y)\exp[i(xk_x+yk_y)]dk_xdk_y \ . \qquad (3)$$

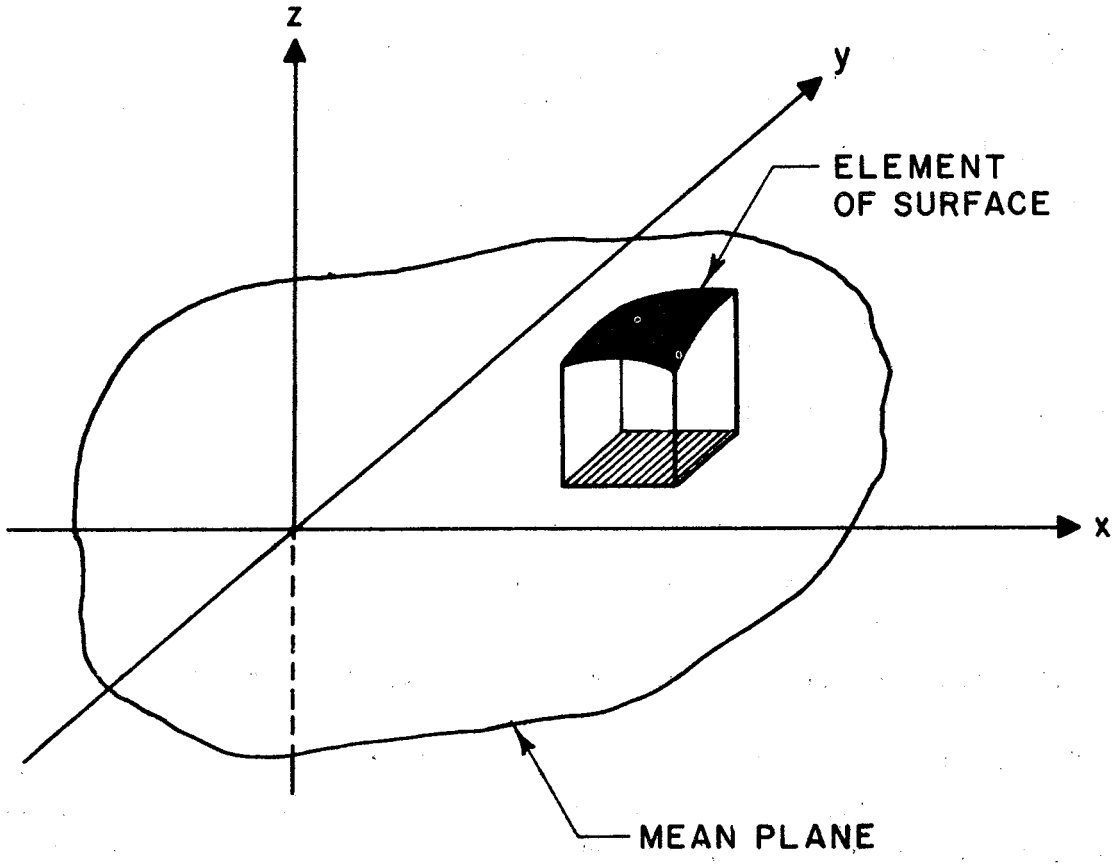


FIG. 5.2 ROUGH SURFACE COORDINATES

k_x and k_y are the components of a wave-vector $\underline{k}$. From Eq. 1, we see that $R(0,0)$ is σ^2, where σ is the rms roughness, or the standard deviation of the roughness. It then follows from Eq. 3 that

$$\sigma^2 = \iint_{-\infty}^{\infty} \Phi(k_x,k_y)dk_x dk_y \ . \tag{4}$$

Equation 4 indicates that $\Phi(k_x,k_y)$ is a decomposition of σ^2 (the *power* in electrical terminology, when the random variable is a current) into contributions from various spectral components, which are waves with wave-number $\underline{k}$. The wavelength of these waves is

$$\lambda = 2\pi/|\underline{k}| \ , \tag{5}$$

and their direction is along

$$\theta = \tan^{-1}(k_y/k_x) \ . \tag{6}$$

For isotropic surfaces, Φ depends only on $k \equiv |\underline{k}|$.

5.2.3 The autocorrelation and PSD for a surface profile

If a profile of the surface be taken in the plane $\theta = \theta_0$, the height z of the profile is a function only of the distance r from the origin along the profile. The autocorrelation and the PSD for the profile are defined to be

$$R_{\theta_0}(r) = \lim_{L\to\infty} \frac{1}{2L} \int_{-L}^{L} z(r_1)z(r_1+r)dr_1 \ , \tag{7}$$

and

$$\Phi_{\theta_0}(k') = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_{\theta_0}(r)\exp(-ik'r)dr \ . \qquad (8)$$

5.2.4 Relation between surface and profile PSD's

Longuet-Higgins [2] shows that the following relation holds between Φ and Φ_{θ_0}:

$$\Phi_{\theta_0}(k') = \int_{-\infty}^{\infty} \Phi(k_x,k_y)d\ell \ , \qquad (9)$$

where

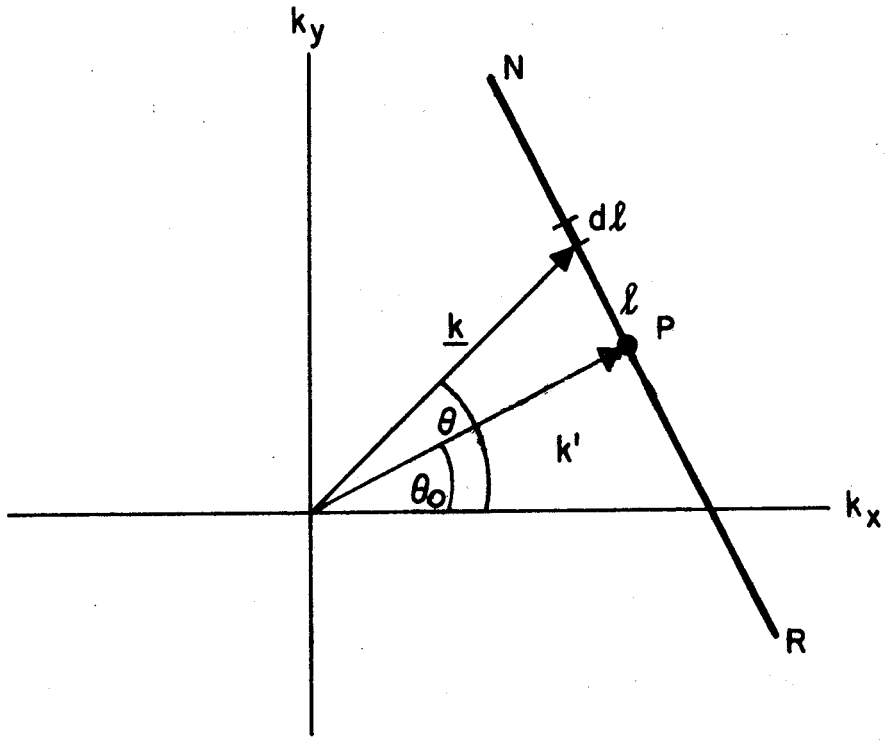
$$\ell = (k_x^2+k_y^2-k'^2)^{1/2} \ . \qquad (10)$$

The physical meaning of Eq. 9 may be clarified with the help of Fig. 5.3. The point P in the wave-number plane has coordinates $(k'\sin\theta_0, k'\cos\theta_0)$. It represents waves with wave-number k' along the profile. The line NPR, perpendicular to OP, is the locus of all wave numbers whose projection on OP is k'. Thus, any wave with a wave vector lying on NPR appears to have a wave-number k' in a section parallel to OP. The line NPR is the path of integration in Eq. 9.

5.2.5 Moments of the PSD

The moments m_{pq} of Φ are defined as follows:

$$m_{pq} = \iint_{-\infty}^{\infty} \Phi(k_x,k_y)k_x^p k_y^q dk_x dk_y \ . \qquad (11)$$



$$\Phi_{\theta_o}(k') = \int_{-\infty}^{\infty} \Phi(\underline{k}')\, d\ell$$

FIG. 5.3 GENERATION OF THE PSD OF THE θ_0 PROFILE FROM THE PSD OF THE SURFACE

It follows from Eq. 4 that

$$m_{00} = \sigma^2 \ . \tag{12}$$

The moment $m_{n\theta_0}$ of the profile is defined as follows:

$$m_{n\theta} = \int_{-\infty}^{\infty} \Phi_{\theta_0}(k')(k')^n dk' \ . \tag{13}$$

The following relation exists between the m_{pq} and the $m_{n\theta_0}$ [2]:

$$m_{n\theta_0} = m_{n0}\cos^n\theta_0 + C_1^n m_{n-1,1}\cos^{n-1}\theta_0\sin\theta_0$$

$$+ C_2^n m_{n-2,2}\cos^{n-2}\theta_0\sin^2\theta_0 + \cdots\cdots + m_{0n}\sin^n\theta_0 \ , \tag{14}$$

where

$$C_m^n = n!/m!(n-m)! \ . \tag{15}$$

Equation 14 may be derived by combining Eqs. 9, 11, and 13. When the surface is isotropic, the following relations may be derived from Eqs. 11 and 14:

$$\left.\begin{aligned} &m_{20} = m_{02} = m_2 \ ; \quad m_{11} = m_{13} = m_{31} = 0 \ ; \\ &m_{00} = m_0 \ ; \quad 3m_{22} = m_{40} = m_{04} = m_4 \ . \end{aligned}\right\} \tag{16}$$

In Eq. 16, the subscript θ_0 for the moments of the profile PSD is dropped, since isotropy implies that the profile statistics are independent of the direction θ_0 of the profile.

5.2.6 The central limit theorem

Let $\xi_1 \cdots \xi_n$ be n quantities, each of which is the sum of a large number of independent variables with zero expectation. Then, under very general conditions [8], the joint probability density of the ξ_i is Gaussian in n dimensions:

$$p(\xi_1 \cdots \xi_n) = (2\pi)^{-\frac{n}{2}} \Delta^{-\frac{1}{2}} \exp\{-\tfrac{1}{2} M_{ij}\xi_i\xi_j\} , \tag{17}$$

where the matrix M is given by

$$M = (N)^{-1} \tag{18}$$

and

$$N = \begin{pmatrix} \overline{\xi_1^2} & \overline{\xi_1\xi_2} & \cdots & \overline{\xi_1\xi_n} \\ \overline{\xi_2\xi_1} & & & \\ \overline{\xi_n\xi_1} & \cdots\cdots\cdots\cdots & & \overline{\xi_n^2} \end{pmatrix} , \tag{19}$$

and

$$\Delta = \text{Det.}(N_{ij}) . \tag{20}$$

The element $\overline{\xi_i\xi_j}$ of the matrix is defined by

$$\overline{\xi_i\xi_j} = \int_E \cdots \iint \xi_i(\varepsilon_1 \cdots \varepsilon_k)\xi_j(\varepsilon_1 \cdots \varepsilon_k)p(\varepsilon_1)\cdots p(\varepsilon_k)d\varepsilon_1 \cdots d\varepsilon_k , \tag{21}$$

where E is the probability space of the independent random variables ε_j on which the ξ_i depend.

5.3 The Statistics of Random, Isotropic, Gaussian Surfaces

Assume that the surface height $z(x,y)$ may be represented by the infinite sum

$$z(x,y) = \sum_n C_n \cos(xk_{xn}+yk_{yn}+\varepsilon_n) \ . \tag{22}$$

It is assumed in Eq. 22 that there are an infinite number of wave-vectors (k_{xn},k_{yn}) in any area $dk_x dk_y$. ε_n is a random phase, with a uniform probability of lying in the range $(0,2\pi)$. The coefficients C_n are related to the PSD Φ by

$$\Phi(k_x,k_y)dk_x dk_y = \frac{1}{2} \sum_{\Delta \underline{k}} C_n^2 \ , \tag{23}$$

the summation being over all n such that (k_{xn},k_{yn}) lies in the area $dk_x dk_y$ around (k_x,k_y). From Eqs. 11 and 23, we see that

$$m_{00} = \sigma^2 = \frac{1}{2} \sum_{\text{all } n} C_n^2 \ . \tag{24}$$

Similarly, we have

$$m_{pq} = \frac{1}{2} \sum_{\text{all } n} k_{xn}^p k_{yn}^q C_n^2 \ . \tag{25}$$

Define the variables $\xi_1 \cdots \xi_6$ as follows:

$$\left.\begin{aligned} \xi_1 &= z & \xi_4 &= \partial^2 z/\partial x^2 \\ \xi_2 &= \partial z/\partial x & \xi_5 &= \partial^2 z/\partial x \partial y \\ \xi_3 &= \partial z/\partial y & \xi_6 &= \partial^2 z/\partial y^2 \ . \end{aligned}\right\} \tag{26}$$

The variables $\xi_1 \cdots \xi_6$ satisfy the requirements of the central limit theorem, and their joint probability density is given by Eq. 17. Then the matrix N_{ij} of Eq. 19 is found to be

$$(N_{ij}) = \begin{pmatrix} m_{00} & 0 & 0 & -m_{20} & -m_{11} & -m_{02} \\ 0 & m_{20} & m_{11} & 0 & 0 & 0 \\ 0 & m_{11} & m_{02} & 0 & 0 & 0 \\ -m_{20} & 0 & 0 & m_{40} & m_{31} & m_{22} \\ -m_{11} & 0 & 0 & m_{31} & m_{22} & m_{13} \\ -m_{02} & 0 & 0 & m_{22} & m_{13} & m_{04} \end{pmatrix} . \qquad (27)$$

As an example of how the elements of N_{ij} are computed, consider the element $\overline{\xi_1 \xi_4}$. From Eqs. 22 and 26 we have

$$\overline{\xi_1 \xi_4} = - \sum_{\text{all } n} C_n^2 k_{xn}^2 \, \overline{\cos^2(xk_{xn} + yk_{yn} + \varepsilon_n)} . \qquad (28)$$

For any one value of n, the average on the right-hand side of Eq. 28 is taken over ε_n, and we have, using Eq. 25,

$$\overline{\xi_1 \xi_4} = - \frac{1}{2} \sum_{\text{all } n} C_n^2 k_{xn}^2 = - m_{20} . \qquad (29)$$

When the surface is isotropic, the matrix M_{ij} of Eq. 18 is found, using Eq. 16, to be

$$M_{ij} = \begin{pmatrix} \frac{2m_4}{\Delta_1} & 0 & 0 & \frac{3m_2}{2\Delta_1} & 0 & \frac{3m_2}{2\Delta_1} \\ 0 & \frac{1}{m_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{m_2} & 0 & 0 & 0 \\ \frac{3m_2}{2\Delta_1} & 0 & 0 & \frac{9\Delta_2}{4m_4\Delta_1} & 0 & -\frac{3\Delta_3}{4m_4\Delta_1} \\ 0 & 0 & 0 & 0 & \frac{3}{m_4} & 0 \\ \frac{3m_2}{2\Delta_1} & 0 & 0 & -\frac{3\Delta_3}{4m_4\Delta_1} & 0 & \frac{9\Delta_2}{4m_4\Delta_1} \end{pmatrix} , \tag{30}$$

where

$$\Delta_1 = (2m_0m_4 - 3m_2^2) , \tag{31}$$

$$\Delta_2 = (m_0m_4 - m_2^2) , \tag{32}$$

and

$$\Delta_3 = (m_0m_4 - 3m_2^2) . \tag{33}$$

The determinant Δ of N_{ij} is found to be

$$\Delta = \frac{4}{27} (m_2m_4)^2 (2m_0m_4 - 3m_2^2) . \tag{34}$$

Thus the joint probability distribution of the variables $\xi_1 \cdots \xi_6$ is found from Eq. 17 to be

$$p(\xi_1 \cdots \xi_6) = (2\pi)^{-3}\ \Delta^{-1/2}\ \exp(-\tfrac{1}{2}\ X)\ , \qquad (35)$$

where Δ is given by Eq. 34 and

$$X = \frac{2m_4}{\Delta_1}\ \xi_1^2 + \frac{9\Delta_2}{4m_4\Delta_1}\ (\xi_4^2+\xi_6^2) + \frac{3}{m_4}\ \xi_5^2$$

$$+ \frac{3m_2}{\Delta_1}\ \xi_1(\xi_4+\xi_6) - \frac{3\Delta_3}{2m_4\Delta_1}\ \xi_4\xi_6 + \frac{1}{m_2}\ (\xi_2^2+\xi_3^2)\ . \qquad (36)$$

We shall now obtain the following statistics of the surface:

1. the probability distribution for summit heights
2. the spatial density of summits
3. the probability distribution for the mean curvature of the summits.

5.3.1 Distribution of summit heights

The requirements for some point (x,y) in an area dA = dxdy to be a summit (i.e., a maximum) are that, at (x,y),

$$\left.\begin{aligned} &\xi_2 = \xi_3 = 0 \\ &\xi_4 < 0\ , \qquad \xi_6 < 0\ , \qquad \xi_4\xi_6 - \xi_5^2 \geq 0\ . \end{aligned}\right\} \qquad (37)$$

The probability that the variables ξ_i at (x,y) will lie in the range $(\xi_i, \xi_i + d\xi_i)$ is $p(\xi_1 \cdots \xi_6) d\xi_1 \cdots d\xi_6$. The increments $d\xi_2$ and $d\xi_3$ that take place in an area dA are given by

$$d\xi_2 d\xi_3 = \left| \frac{\partial(\xi_2, \xi_3)}{\partial(x,y)} \right| dA \ , \tag{38}$$

where the Jacobian has the value

$$\frac{\partial(\xi_2, \xi_3)}{\partial(x,y)} = \frac{\partial \xi_2}{\partial x} \cdot \frac{\partial \xi_3}{\partial y} - \frac{\partial \xi_2}{\partial y} \cdot \frac{\partial \xi_3}{\partial x} = \xi_4 \xi_6 - \xi_5^2 \ . \tag{39}$$

The point (x,y) will be a summit of height between ξ_1 and $\xi_1 + d\xi_1$ if $\xi_2 = \xi_3 = 0$, $d\xi_2$ and $d\xi_3$ satisfy Eq. 38 and ξ_4, ξ_5, ξ_6 take on arbitrary values subject to Eq. 37. Thus, if $P_{sum}(\xi_1)$ is the probability distribution for summits of height ξ_1, the probability of having a summit in the area dA with height in the range $(\xi_1, \xi_1 + d\xi_1)$ is

$$P_{sum}(\xi_1) dA d\xi_1 = d\xi_1 \iiint_V p(\xi_1, 0, 0, \xi_4, \xi_5, \xi_6) d\xi_2 d\xi_3 d\xi_4 d\xi_5 d\xi_6 \ . \tag{40}$$

The domain of integration V is defined by

$$\xi_4 < 0 \quad , \quad \xi_6 < 0 \quad , \quad \xi_4 \xi_6 - \xi_5^2 \geq 0 \ . \tag{41}$$

Using Eqs. 38 and 39 in Eq. 40, and substituting for $p(\xi_1 \cdots \xi_6)$ from Eq. 35, we obtain

$$P_{sum}(\xi_1) = \frac{\exp(-m_4\ \xi_1^2/\Delta_1)}{(2\pi)^3\ \Delta^{\frac{1}{2}}} \iiint_V |\xi_4\xi_6-\xi_5^2| \exp\left\{-\frac{1}{2}\left[\frac{9\Delta_2}{4m_4\Delta_1}\right.\right.$$

$$\times\ (\xi_4^2+\xi_6^2) + \frac{3}{m_4}\ \xi_5^2 + \frac{3m_2}{\Delta_1}\ \xi_1(\xi_4+\xi_6)$$

$$\left.\left. - \frac{3\Delta_3}{2m_4\Delta_1}\ \xi_4\xi_6\right]\right\} d\xi_4 d\xi_5 d\xi_6 \ . \qquad (42)$$

Define

$$\left.\begin{aligned} (t_1,t_2,t_3) &= \left(\frac{3}{m_4}\right)^{\frac{1}{2}}\left[\frac{1}{2}\ (\xi_4+\xi_6)\ ,\ \xi_5\ ,\ \frac{1}{2}\ (\xi_4-\xi_6)\right] \\ \text{and} \\ \xi_i^* &= \xi_1/m_0^{\frac{1}{2}} = \xi_1/\sigma \ . \end{aligned}\right\} \qquad (43)$$

Then, using the transformation $P_{sum}(\xi_1^*) = P_{sum}(\xi_1)\left|\partial\xi_1/\partial\xi_1^*\right|$, Eq. 42 may be written

$$P_{sum}(\xi^*) = \left(\frac{m_4}{m_2}\right)\ \frac{[C_1(\alpha)]^{\frac{1}{2}}}{3(2\pi)^3}\ \exp[-C_1\xi^{*2}]$$

$$\times \iiint_{V'} |t_1^2-t_2^2-t_3^2| \exp\{-\frac{1}{2}\ [C_1t_1^2+t_2^2+t_3^2+C_2t_1\xi^*]\}dt_1dt_2dt_3 \ , \qquad (44)$$

where

and

$$\left.\begin{aligned} \alpha &= m_0 m_4 / m_2^2 \\ C_1 &= \alpha/(2\alpha-3) \\ C_2 &= C_1 (12/\alpha)^{\frac{1}{2}} \end{aligned}\right\} \tag{45}$$

The domain of integration V' is defined by

$$\left.\begin{aligned} t_1 &< 0 \\ t_2^2+t_3^2 &\leq t_1^2 \end{aligned}\right\} . \tag{46}$$

The probability density for summit heights, p_{sum} is obtained by dividing P_{sum} by D_{sum}, the density of summits:

$$p_{sum}(\xi^*) = P_{sum}(\xi^*)/D_{sum} , \tag{47}$$

where

$$D_{sum} = \int_{-\infty}^{\infty} P_{sum}(\xi^*) d\xi^* . \tag{48}$$

The integrals in Eqs. 44 and 48 may be evaluated analytically to yield

$$D_{sum} = \frac{1}{6\pi\sqrt{3}} \left(\frac{m_4}{m_2}\right) . \tag{49}$$

This result agrees with the expression given by Longuet-Higgins [3], which was derived by a slightly different method.

Substituting the expression for D_{sum} into Eq. 47, we obtain

$$p_{sum}(\xi^*) = \frac{\sqrt{3}}{2\pi}\left\{e^{-C_1\xi^{*2}}\left[\frac{3(2\alpha-3)}{\alpha^2}\right]^{\frac{1}{2}} + \frac{3\sqrt{2}\pi}{2\alpha}\, e^{-\frac{1}{2}\xi^{*2}}(1+\text{erf}\beta)(\xi^{*2}-1)\right.$$

$$\left. + \sqrt{2\pi}\left[\frac{\alpha}{3(\alpha-1)}\right]^{\frac{1}{2}} \exp\{-[(\alpha\xi^{*2})/2(\alpha-1)]\}(1+\text{erf}\gamma)\right\}, \tag{50}$$

where

$$\left.\begin{aligned} \beta &= \left[\frac{3}{2(2\alpha-3)}\right]^{\frac{1}{2}}\xi^* \\ \text{and} \\ \gamma &= \left[\frac{\alpha}{2(\alpha-1)(2\alpha-3)}\right]^{\frac{1}{2}}\xi^* \end{aligned}\right\}. \tag{51}$$

The probability density $p_{sum}(\xi^*)$ is shown in Fig. 5.4 for a range of values of α. Longuet-Higgins [2] has shown that $\alpha \geq 1.5$ for a random, isotropic surface. Thus no values of α lower than 1.5 appear in Fig. 5.4. It may be seen that as α decreases to 1.5, the probability of a high peak increases. The parameter α is related to the breadth of the surface PSD. A broad spectrum is one that has waves with a large range of wavelengths; a narrow spectrum has waves of approximately equal wavelength. As $\alpha \to 1.5$, the spectrum gets narrower; and as $\alpha \to \infty$, it gets broader.

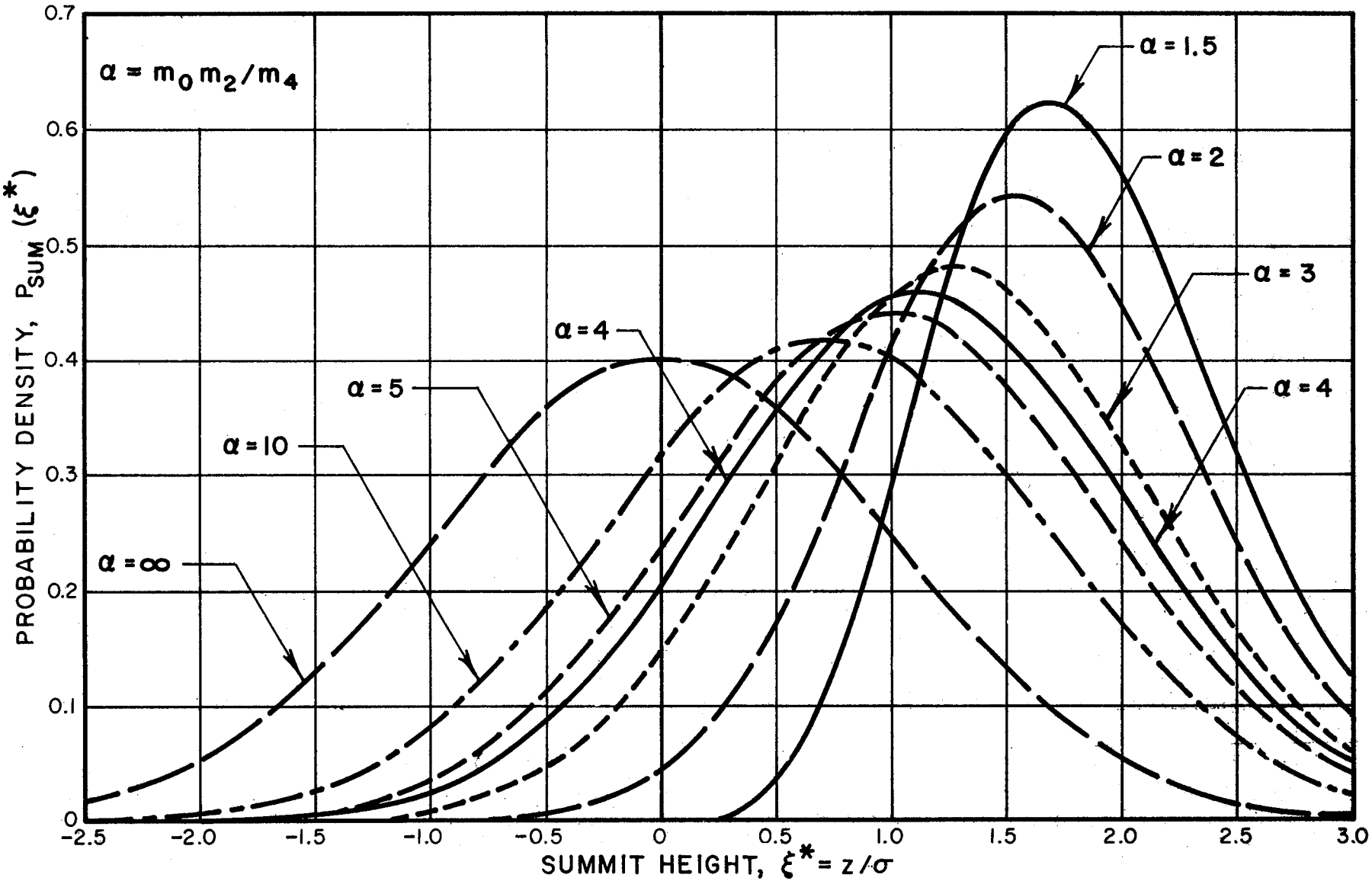


FIG. 5.4 PROBABILITY DENSITY FOR SUMMIT HEIGHTS

The two limiting forms of p_{sum} for $\alpha \to 1.5$ and $\alpha \to \infty$ are:

1. Lim $\alpha \to 1.5$

$$p_{sum}(\xi^*) = \begin{cases} \frac{2\sqrt{3}}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi^{*2}} \left[\xi^{*2}-1+e^{-\xi^{*2}}\right] , & \xi^* \geq 0 \\ 0 , & \xi^* < 0 . \end{cases} \quad (52)$$

2. Lim $\alpha \to \infty$

$$p_{sum}(\xi^*) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi^{*2}} . \quad (53)$$

Thus, when $\alpha \to \infty$, the summit heights have a Gaussian distribution. When $\alpha \to 1.5$, the distribution of summit heights for $\xi^* > 2$ may be written

$$p_{sum}(\xi^*>2) \approx \left(\frac{6}{\pi}\right)^{\frac{1}{2}} (\xi^{*2}-1) e^{-\frac{1}{2}\xi^{*2}} , \quad \alpha = 1.5 . \quad (54)$$

The most remarkable thing about the expression for $p_{sum}(\xi^*)$ is that it depends only on the parameter α, which may be obtained from a surface profile, as explained in Sec. 5.4.

An interesting observation that may be made about the expected density of summits D_{sum}(no./unit area), Eq. 49, is that when the surface PSD is flat up to fairly high wave-numbers (i.e., small wavelengths), D_{sum} will be very large. This is because one effect of the small-wavelength components is to cause large clusters of "mini-summits" of small amplitude and wavelength to appear around a "maxi-summit" of large amplitude and wavelength. If, however,

one is interested only in the maxi-summits, their density may be obtained by filtering out the high wave-number content of the surface PSD. This causes m_4 to decrease faster than m_2, and consequently causes a decrease in D_{sum}.

A further quantity of interest is the cumulative density of summits $q_{sum}(\xi^*)$, which indicates the fraction of all summits that are expected to lie below ξ^*:

$$q_{sum}(\xi^*) = \int_{-\infty}^{\xi^*} p_{sum}(\xi^*)d\xi^* \ . \qquad (55)$$

This function is tabulated in Table 5.1. It may be seen that as $\alpha \to 1.5$, the fraction of summits that lie below the $+3\sigma$ level (i.e., below $\xi^* = 3$) decreases.

5.3.2 The mean summit curvature

The mean curvature κ_m at any point on a surface is defined as the mean of the principal curvatures κ_1 and κ_2 at that point. In addition, the sum of the curvatures of a surface at a point along any two orthogonal directions is equal to the sum of the principal curvatures [9]. Furthermore, the curvatures of a surface at a summit in the x and y directions are $-\partial^2 z/\partial x^2$ and $-\partial^2 z/\partial y^2$. Thus the mean curvature at a summit is

$$\kappa_m = \frac{1}{2}(\kappa_1+\kappa_2) = -\frac{1}{2}\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right)$$

$$= -\frac{1}{2}(\xi_4+\xi_6) \ . \qquad (56)$$

TABLE 5.1 Cumulative Probability Distribution for Summit Heights

ξ^*	$q_{sum} = \int_{-\infty}^{\xi^*} P_{sum}(\xi^*)d\xi^*$						
	$\alpha = 1.5$	2	3	4	5	10	∞
0.0	0.0	.0129	.0601	.1026	.1363	.2327	.5
.25	.0001	.0306	.1103	.1686	.2118	.3263	.5987
.50	.0037	.0706	.1739	.2446	.2943	.4187	.6915
.75	.0237	.1296	.2696	.3496	.4031	.5296	.7733
1.00	.0785	.2281	.3701	.4521	.5050	.6251	.8413
1.25	.1787	.3381	.4955	.5724	.6205	.7251	.8943
1.50	.3176	.4794	.6052	.6724	.7134	.8000	.9332
1.75	.4745	.6030	.7199	.7725	.8040	.8689	.9599
2.00	.6254	.7293	.8042	.8435	.8668	.9138	.9772
2.25	.7525	.8187	.8785	.9044	.9196	.9498	.9878
2.50	.8482	.8936	.9247	.9414	.9512	.9704	.9938
2.75	.9134	.9375	.9594	.9687	.9742	.9848	.9970
3.00	.9539	.9625	.9778	.9830	.9861	.9920	.9987

Using Eq. 43, we find

$$\kappa_m = - \left(\frac{m_4}{3} \right)^{1/2} t_1 \ . \tag{57}$$

We call t_1 the *equivalent* mean curvature. The joint probability distribution for summits with height ξ_1^* and equivalent mean curvature t_1 is

$$P_{sum}(\xi_1^*,t_1) = \frac{\left(\frac{m_4}{m_2} \right)}{3(2\pi)^3} C_1^{\frac{1}{2}} \exp[- C_1\xi^{*2}]\exp\{- \frac{1}{2} [C_1(\alpha)t_1^2+C_2(\alpha)t_1\xi^*]\}$$

$$\times \iint_S |t_1^2-t_2^2-t_3^2| \exp\{- \frac{1}{2} [t_2^2+t_3^2]\}dt_2dt_3 \ . \tag{58}$$

The domain of integration S is defined by

$$t_1^2 \geq t_2^2 + t_3^2 \ . \tag{59}$$

In addition, at a summit we always have $t_1 < 0$.

$P_{sum}(\xi_1^*,t_1)$ may be normalized by dividing by D_{sum} to yield a probability density:

$$p_{sum}(\xi_1^*,t_1) \equiv P_{sum}(\xi_1^*,t_1)/D_{sum} \ . \tag{60}$$

The integral in Eq. 58 may be evaluated analytically to give

$$p_{sum}(\xi_1^*,t_1) = \frac{\sqrt{3C_1}}{2\pi} \exp(-C_1\xi^{*2})$$

$$\times \left(t_1^2-2+2e^{-\frac{1}{2} t_1^2} \right) \exp\{- \frac{1}{2} [C_1(\alpha)t_1^2+C_2(\alpha)t_1\xi^*]\} \ . \tag{61}$$

The expected value of the mean curvature for summits of height ξ_1^*, $\overline{\kappa}_m(\xi_1^*)$ is found from Eq. 61:

$$\overline{\kappa}_m(\xi_1^*) = \frac{-\left(\frac{m_4}{3}\right)^{\frac{1}{2}} \int_{-\infty}^{0} t_1 p_{sum}(\xi_1^*, t_1) dt_1}{\int_{-\infty}^{0} p_{sum}(\xi_1^*, t_1) dt_1} . \tag{62}$$

The expression for $\overline{\kappa}_m(\xi^*)$ is found analytically to be

$$\overline{\kappa}_m(\xi^*) = \left(\frac{m_4}{3}\right)^{\frac{1}{2}} \cdot \frac{I_3(\xi^*) - 2I_1(\xi^*) + 2I_5(\xi^*)}{I_2(\xi^*) - 2I_0(\xi^*) + 2I_4(\xi^*)} , \tag{63}$$

where $I_0 \cdots I_5$ are functions of ξ^* and α, given in Appendix I.

The dimensionless expected mean curvature $\overline{\kappa}_m/\sqrt{m_4}$ is shown in Fig. 5.5 for a range of values of α. It may be seen that high peaks always have a larger curvature (i.e., a smaller summit radius) than lower peaks. For $\alpha = 1.5$, $\overline{\kappa}_m$ varies linearly with ξ^*. For very large values of α, however, the curvature is very nearly constant for summits of all heights. The two limiting cases are:

1. <u>Lim $\alpha \to 1.5$</u>

$$\overline{\kappa}_m/\sqrt{m_4} = \sqrt{2/3}\ \xi^* \quad , \qquad \xi^* \geq 0 . \tag{64}$$

2. <u>Lim $\alpha \to \infty$</u>

$$\overline{\kappa}_m/\sqrt{m_4} = 8/3\sqrt{\pi} . \tag{65}$$

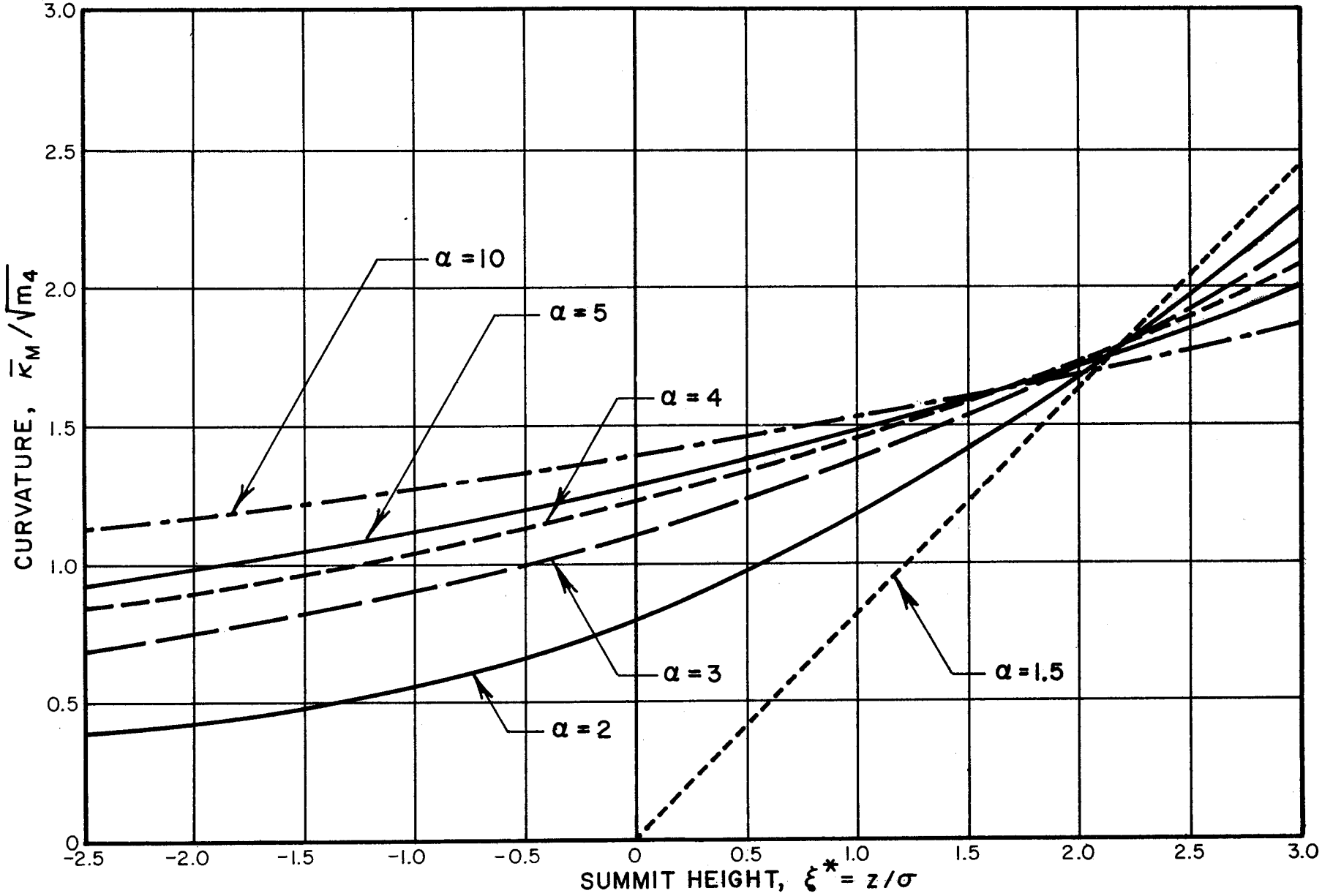


FIG. 5.5 EXPECTED MEAN DIMENSIONLESS CURVATURE OF SUMMITS, AS A FUNCTION OF SUMMIT HEIGHT

Further fairly elementary results may be derived concerning the probability distributions of the height of the surface and the surface gradient.

5.3.3 The surface height

The probability density for the surface height ξ_1 is obtained in a straightforward manner from Eq. 17 by noting that

$$M_{11} = \frac{1}{N_{11}} = \frac{1}{\overline{\xi_1^2}} = \frac{1}{m_{00}} = \frac{1}{\sigma^2} ,$$

and $\Delta = \sigma^2$. Thus,

$$p(\xi_1) = \frac{1}{\sigma\sqrt{2\pi}} \exp[-\frac{1}{2}(\xi_1/\sigma)^2] . \tag{66}$$

5.3.4 The surface gradient

The surface gradient is defined by

$$\zeta = (\xi_2^2 + \xi_3^2)^{1/2} . \tag{67}$$

Longuet-Higgins [*10*] shows that the probability density for ζ is

$$p(\zeta) = \frac{\zeta}{m_2} \exp(-\zeta^2/2m_2) . \tag{68}$$

The expected value of the gradient is found from Eq. 68

$$\overline{\zeta} = \int_0^\infty \zeta p(\zeta) d\zeta = \left(\frac{\pi m_2}{2}\right)^{1/2} . \tag{69}$$

5.4 The Sampling of Random Surfaces

In Sec. 5.3, it was shown that a large number of useful statistics of the surfaces may be found for random, isotropic, Gaussian surfaces if the parameters m_0, m_2, and m_4 are known. It was also shown that these parameters are moments of the PSD of a profile of the surface in an arbitrary direction.

Once a profile of the surface is obtained, the autocorrelation of the profile may be calculated, using Eq. 7. The profile PSD is then calculated, using Eq. 8. The moments m_0, m_2, and m_4 are obtained by using Eq. 13.

An ingenious alternative has been described by Longuet-Higgins [2]. He shows that the densities of zeroes and of extrema (maxima or minima) along a profile are given by

$$D_{zero,\theta} = \frac{1}{\pi}\left(\frac{m_2}{m_0}\right)^{\frac{1}{2}} , \tag{70}$$

and

$$D_{extrema,\theta} = \frac{1}{\pi}\left(\frac{m_4}{m_2}\right)^{\frac{1}{2}} . \tag{71}$$

The subscript θ indicates a profile statistic, not anistropy.

Once a profile is obtained, $\sigma = \sqrt{m_0}$ is easily calculated. Then m_2 and m_4 may be obtained from Eqs. 70 and 71 simply by counting the number of zeroes and extrema per unit length of the profile:

$$m_2 = \pi^2\sigma^2(D_{zero,\theta})^2 , \tag{72}$$

and

$$m_4 = \pi^4 \sigma^2 (D_{zero,\theta})^2 \, (D_{extrema,\theta})^2 \quad . \tag{73}$$

It is worth noting that by this technique, the high wave-number content of Φ_θ is filtered out if one only counts major peaks, valleys, and zero-crossings. Conversely, a low-pass filtered signal will only contain major peaks, valleys and zero-crossings.

The parameter α of Eq. 45 is

$$\alpha = \frac{m_0 m_4}{m_2^2} = \left(\frac{D_{extrema,\theta}}{D_{zero,\theta}} \right)^2 \quad . \tag{74}$$

In passing, we may note that the density of peaks (maxima) along the profile is, due to symmetry, one-half the number of extrema:

$$D_{peak} = \frac{1}{2\pi} \left(\frac{m_4}{m_2} \right)^{1/2} \quad . \tag{75}$$

Comparing this with the density of summits on the surface, D_{sum}, Eq. 49, we have

$$D_{sum} \approx 1.2 \, (D_{peak})^2 \quad . \tag{76}$$

The question arises as to whether the parameters m_0, m_2, m_4 and α can be obtained in a direct fashion by instrumental analysis of the output of a profilometer. One technique for doing this, being developed at Bolt Beranek and Newman Inc. is as follows. The signal from the profilometer is passed through a preamplifier,

low-pass filter with adjustable cutoff frequency, and an amplifier. The rms value is read directly from a quasi-rms meter. A circuit for determining the zero-level of the signal is used along with a comparator and a counter to determine the zero-crossing rate. To obtain the rate of extrema, the zero-crossing rate of the differentiated signal is measured. The cutoff frequency on the filter is adjusted according to the nature of the particular contact problem being investigated.

For further comparisons between the statistics of the surface and the statistics of the profile, we present results for (a) the height distribution of peaks of the profile, (b) the expected value of the peak curvature as a function of peak height, and (c) the distribution of slopes on the profile.

5.4.1 Heights of profile peaks

Cartwright and Longuet-Higgins [*11*], following Rice [*12*], have shown that the probability density for the heights of profile peaks is

$$p_{peak}(\xi^*) = \frac{\delta}{\sqrt{2\pi}} \{\exp[-(\xi^{*2}/2\delta^2)] + \sqrt{\pi}\,\chi \exp(-1/2\ \xi^{*2})(1+\mathrm{erf}\chi)\}\ , \tag{77}$$

where

$$\left.\begin{aligned} \alpha &= m_0 m_4 / m_2^2 \\ \delta &= [(\alpha-1)/\alpha]^{\frac{1}{2}} \end{aligned}\right\} \quad \text{and} \quad \left\{ \chi = \left(\frac{1-\delta^2}{2\delta^2}\right)^{\frac{1}{2}} \xi^* \right\} . \tag{78}$$

The function $p_{peak}(\xi^*)$ is shown in Fig. 5.6 for a range of values of α. It may be shown that for one-dimensional random processes, $1 \leq \alpha \leq \infty$. Thus, δ varies from $\delta = 0$ when $\alpha = 1$ to $\delta = 1$ when $\alpha = \infty$. The significance of the parameter α is the same as for a random surface. Small values of α indicate a narrow spectrum, and large values of α indicate a broad spectrum. When the one-dimensional random process is a profile of an isotropic, random surface, however, α can only take values greater than 1.5.

Figures 5.7- 5.12 show a comparison of $p_{peak}(\xi^*)$ and $p_{sum}(\xi^*)$. It may be seen that the profile distorts the surface in such a way as to show far fewer high peaks and far more low peaks than actually exist on the surface. The distortion is the greatest when $\alpha = 1.5$; it becomes zero when $\alpha \to \infty$. The reason for the distortion is that more often than not, the profile-measuring instrument will travel over the shoulder of an asperity on the surface, rather than over the summit. A peak will still appear on the profile, but of a smaller height than the summit being sampled.

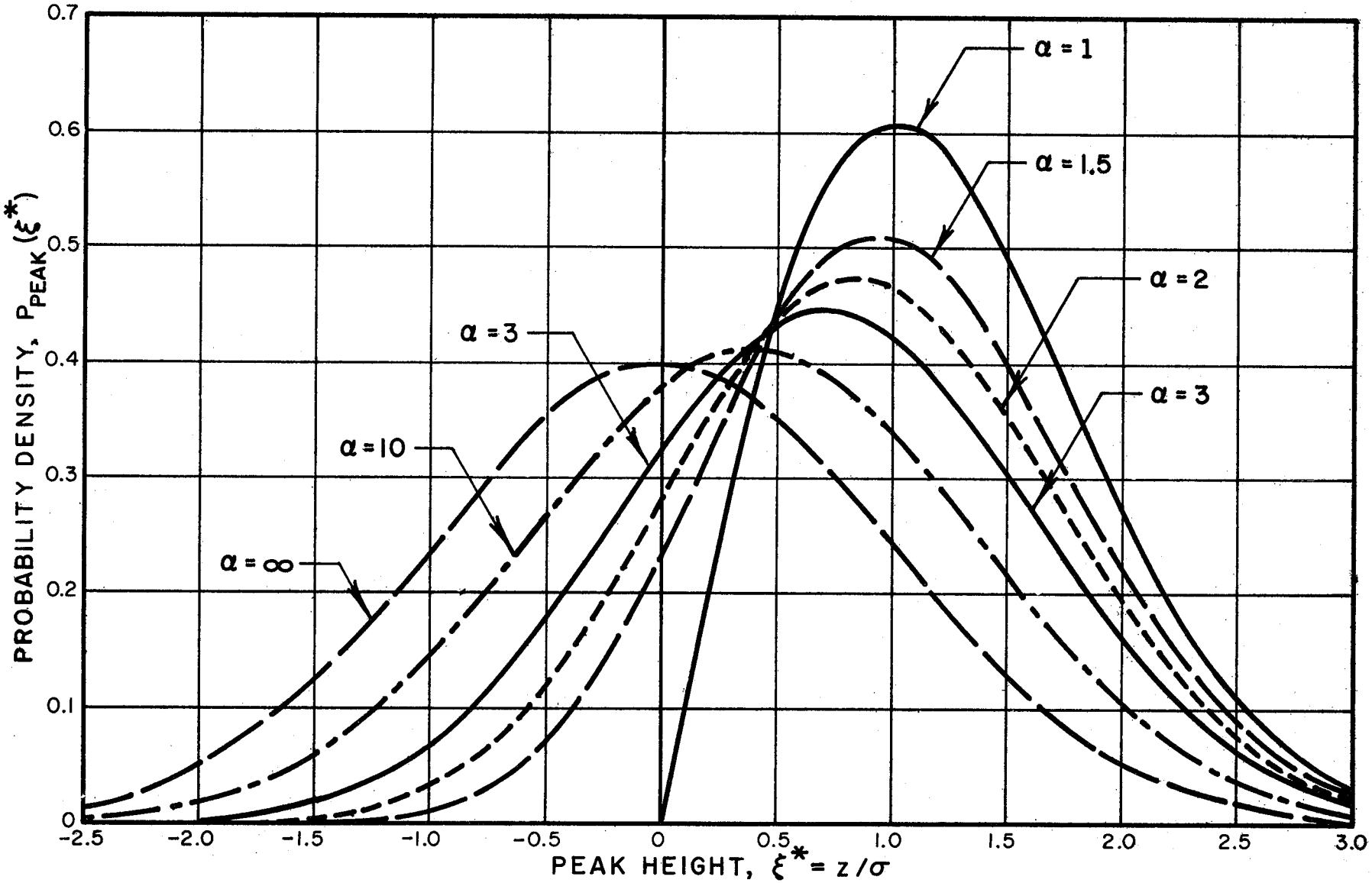
In the limiting cases of $\alpha \to 1$ and $\alpha \to \infty$, the following expressions are obtained for p_{peak} from Eq. 77:

1. Lim $\alpha \to 1$

$$p_{peak}(\xi^*) = \begin{cases} \xi^* \exp(-1/2\ \xi^{*2}) & , \quad \xi^* \geq 0 \\ 0 & , \quad \xi^* < 0 \ . \end{cases} \tag{79}$$

2. Lim $\alpha \to \infty$

$$p_{peak}(\xi^*) = \frac{1}{\sqrt{2\pi}} \exp(-1/2\ \xi^{*2}) \ . \tag{80}$$



FIG, 5.6 PROBABILITY DENSITY FOR HEIGHTS OF PEAKS ON A PROFILE

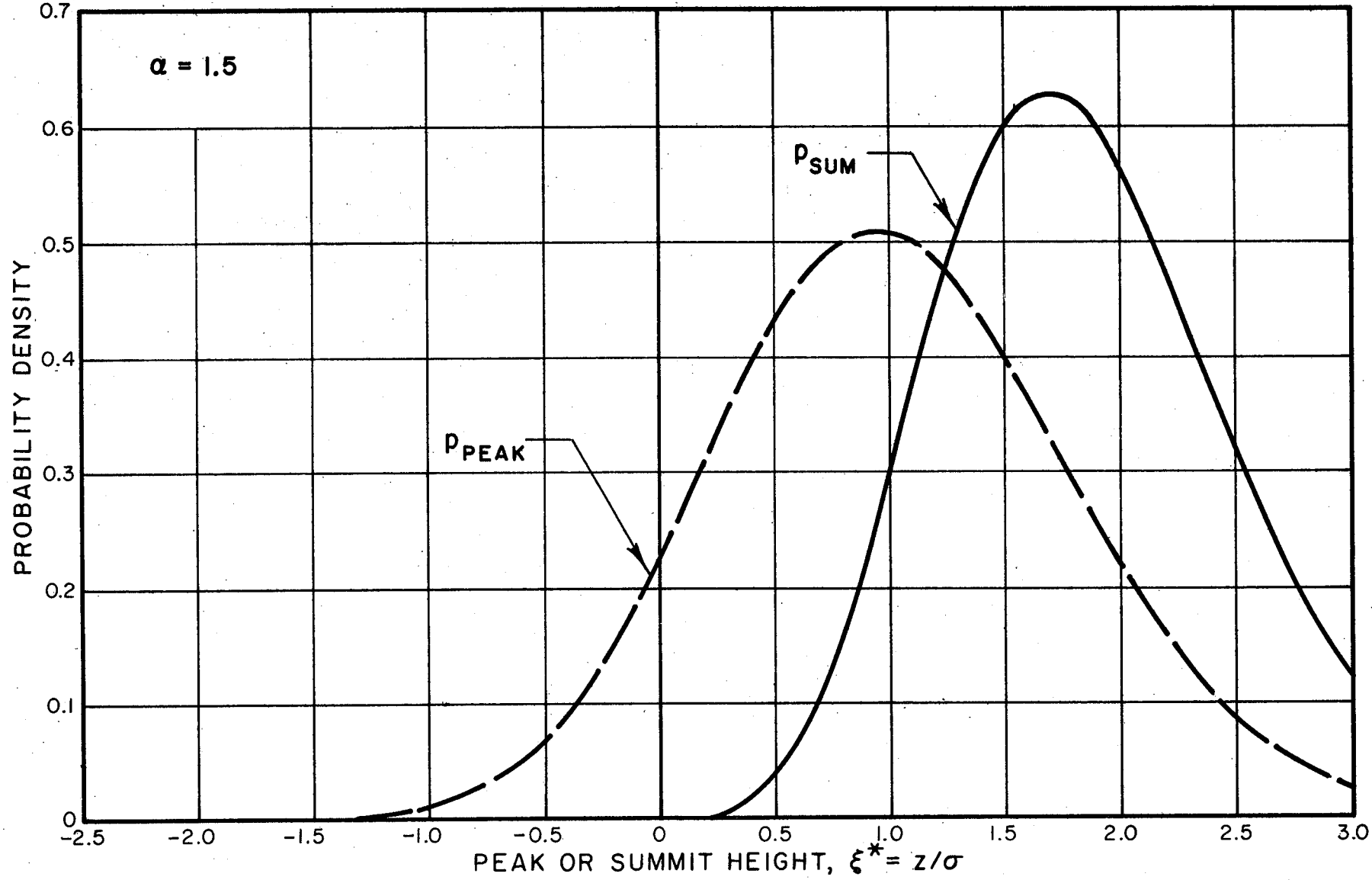


FIG. 5.7 COMPARISON OF PROBABILITY DENSITIES FOR PEAK AND SUMMIT HEIGHTS

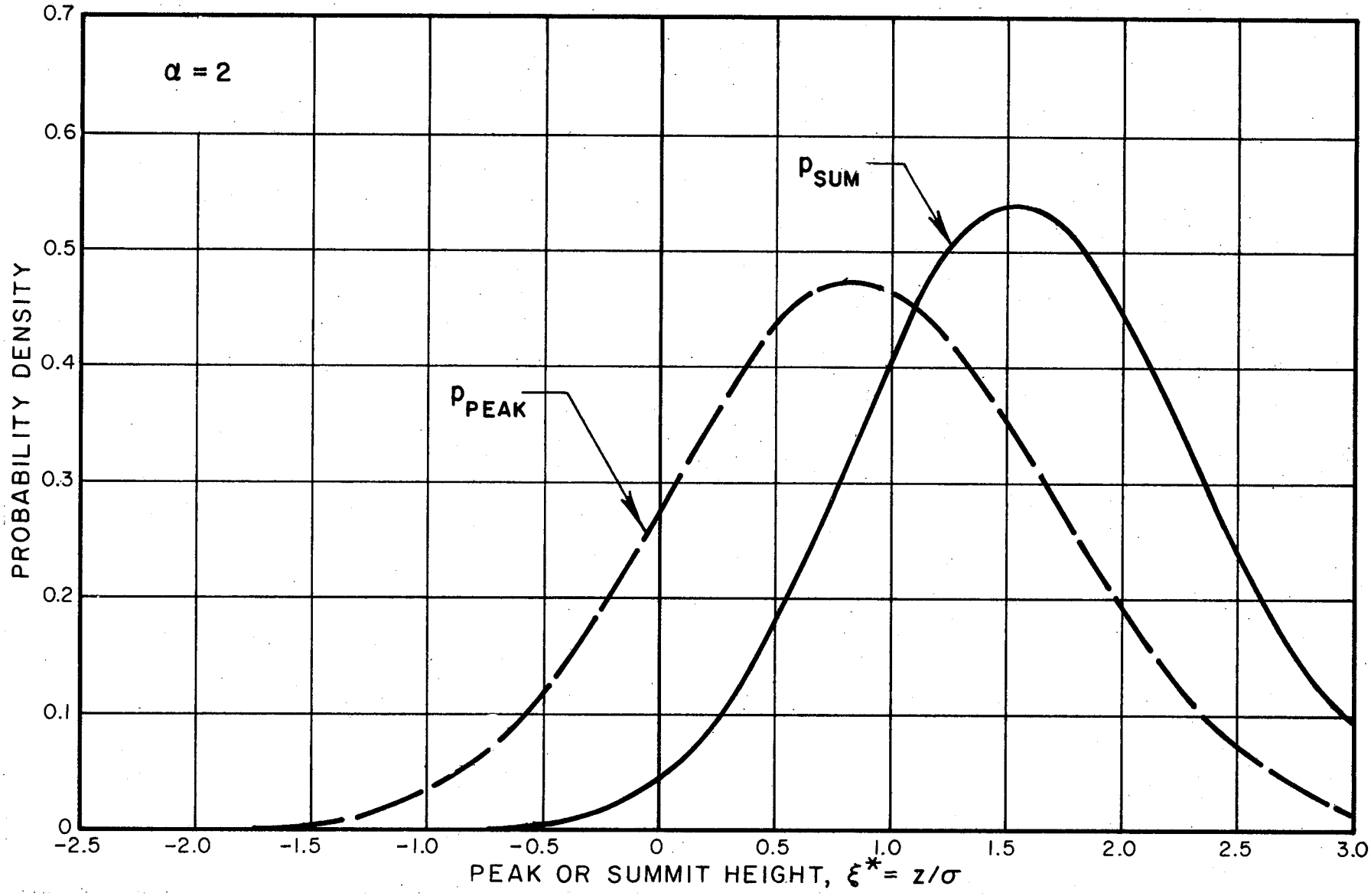


FIG. 5.8 COMPARISON OF PROBABILITY DENSITIES FOR PEAK AND SUMMIT HEIGHTS

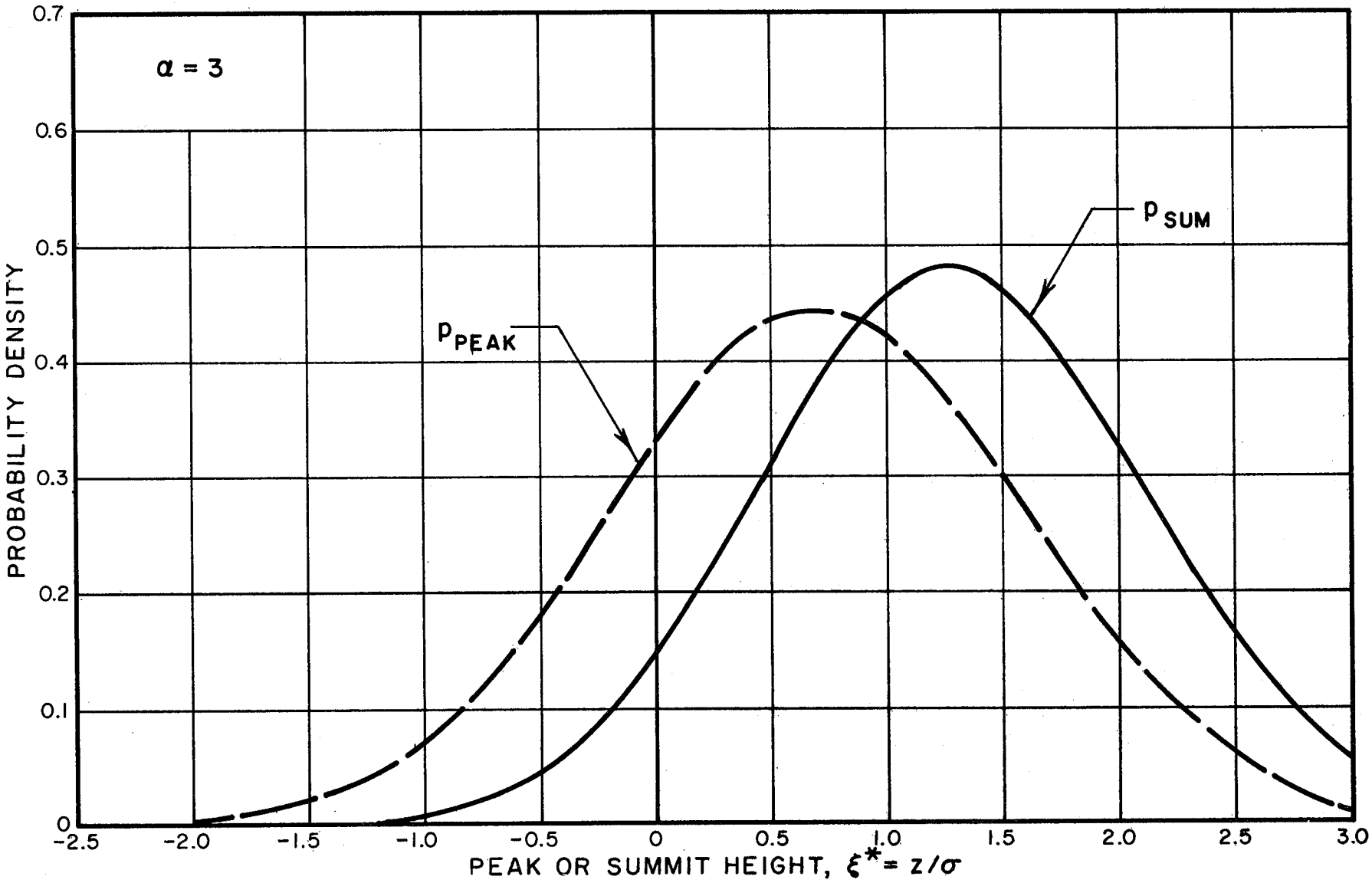


FIG. 5.9 COMPARISON OF PROBABILITY DENSITIES FOR PEAK AND SUMMIT HEIGHTS

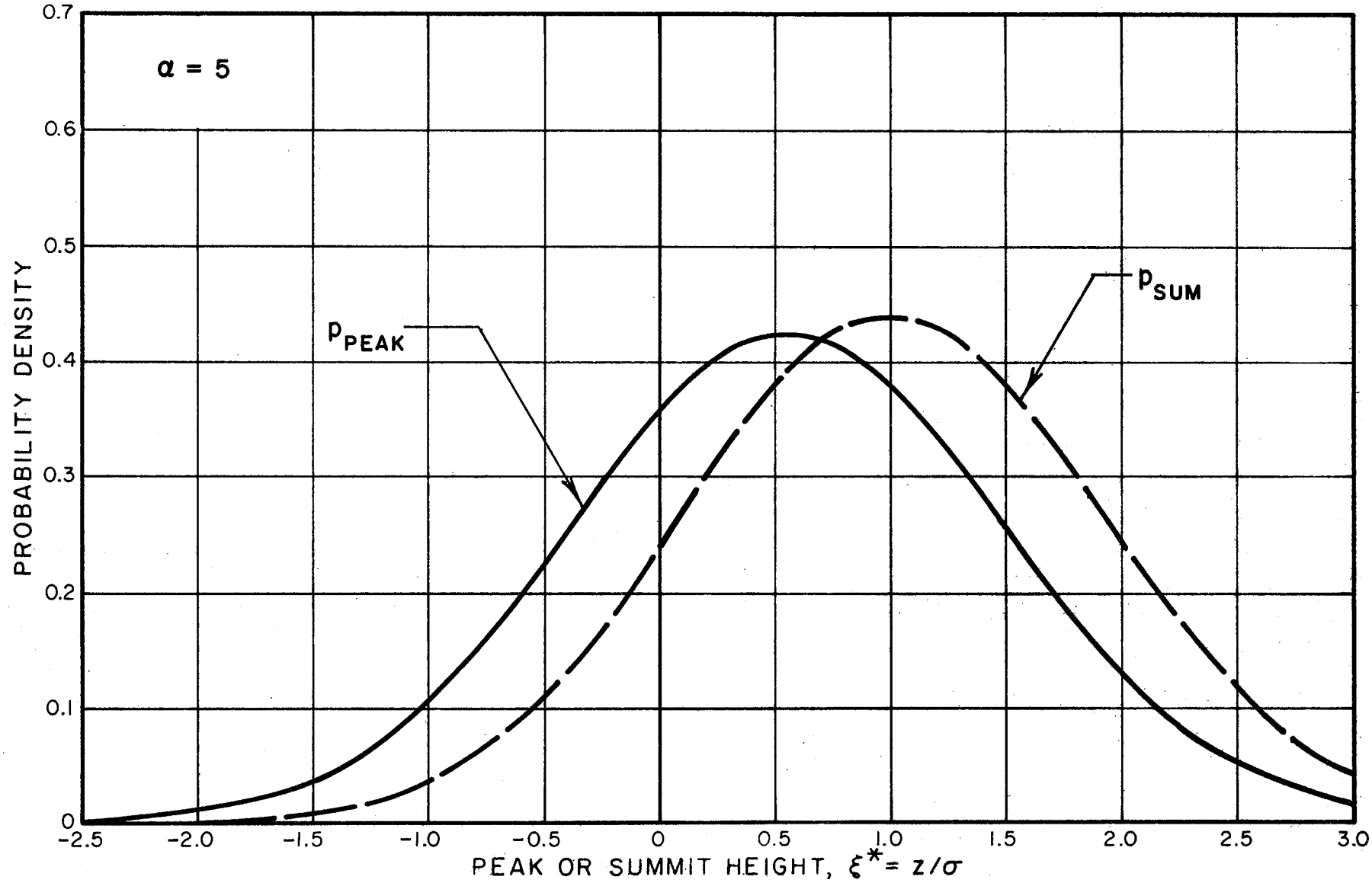


FIG.5.10 COMPARISON OF PROBABILITY DENSITIES FOR PEAK AND SUMMIT HEIGHTS

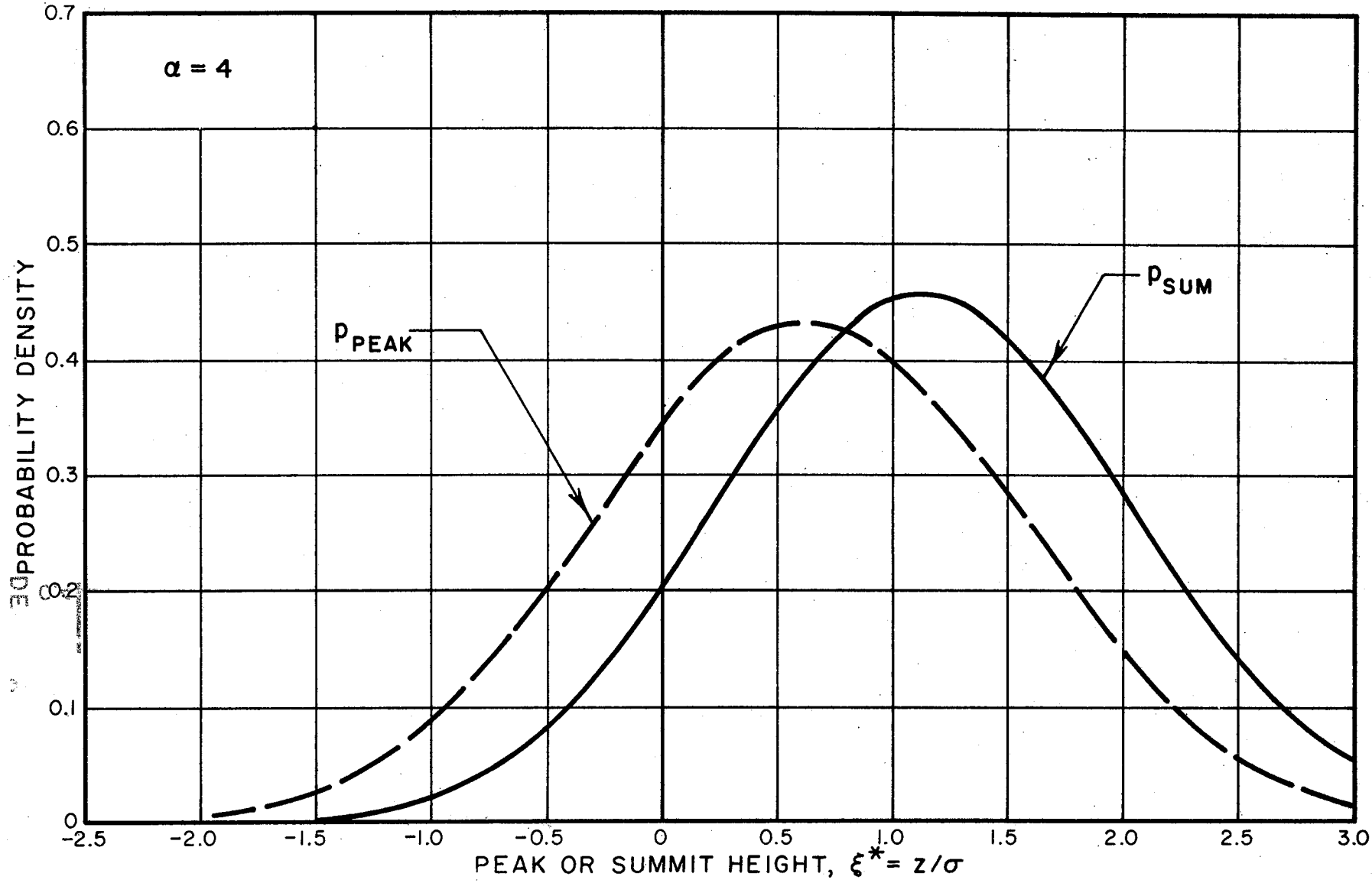


FIG.5.11 COMPARISON OF PROBABILITY DENSITIES FOR PEAK AND SUMMIT HEIGHTS

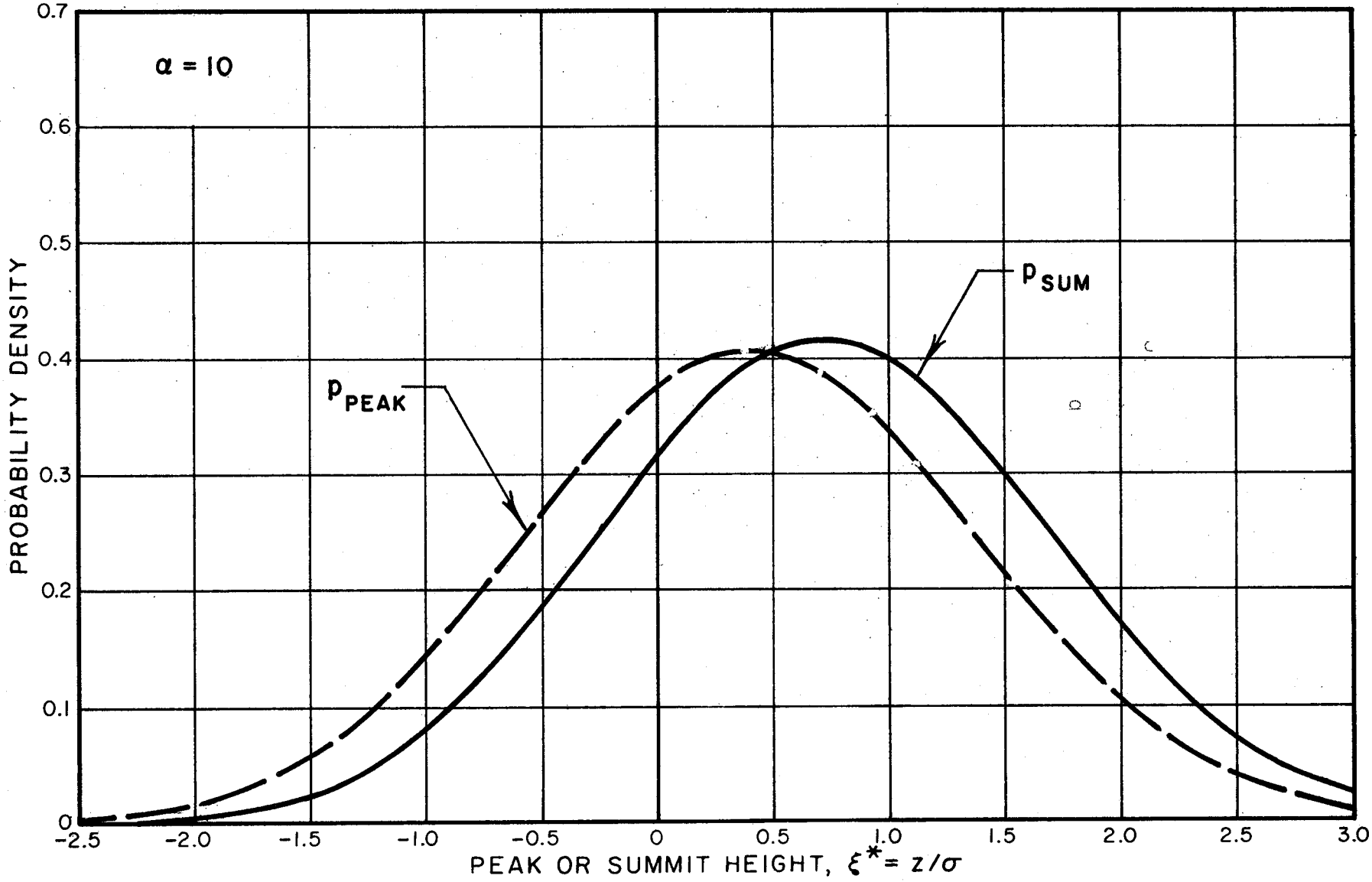


FIG.5.12 COMPARISON OF PROBABILITY DENSITIES FOR PEAK AND SUMMIT HEIGHTS

The cumulative probability density for peak heights,

$$q_{peak}(\xi^*) = \int_{-\infty}^{\xi^*} p_{peak}(\xi^*)d\xi^* \ , \tag{81}$$

is tabulated in Table 5.2. It is clear from a comparison of the data in Tables 5.1 and 5.2 that there are many more summits above the $+3\sigma$ level than would be indicated by the profile, except when $\alpha \to \infty$.

5.4.2 Peak curvature

In a manner similar to that leading to the expected mean summit curvature, $\overline{\kappa}_m(\xi^*)$, Eq. 63, the expected value of the profile peak curvature is found to be

$$\overline{\kappa}_{peak}\ (\xi^*) = m_4^{\frac{1}{2}}\ \delta\sqrt{2}\ \frac{[\chi + \sqrt{\pi}\ e^{\chi^2}(1+\text{erf}\chi)(\chi^2+0.5)]}{[1\ +\ \chi\ e^{\chi^2}\ \sqrt{\pi}\ (1+\text{erf}\chi)]}\ , \tag{82}$$

where δ and χ are as defined in Eq. 78.

The dimensionless peak curvature $\overline{\kappa}_{peak}/\sqrt{m_4}$ is shown in Fig. 5.13. Figures 5.14-5.19 show a comparison of $\overline{\kappa}_{peak}/\sqrt{m_4}$ and $\overline{\kappa}_m/\sqrt{m_4}$. The distortion of the surface by the profile is again evident, but is now quite small for $\alpha > 2$. For $\alpha \lesssim 2.5$, the profile shows a larger peak curvature than the true summit curvature. For $\alpha > 2.5$, the profile peak curvature is less than the summit curvature.

TABLE 5.2 Cumulative Probability Distribution for Peak Heights

ξ^*	$q_{peak} = \int_{-\infty}^{\xi^*} P_{peak}(\xi^*)d\xi^*$							
	$\alpha = 1.0$	1.5	2	3	4	5	10	∞
0.0	0.0	.0956	.1560	.2168	.2558	.2823	.3482	.5
.25	.0308	.1738	.2340	.3143	.3568	.3853	.4541	.5987
.50	.1175	.2662	.3434	.4119	.4545	.4827	.5497	.6915
.75	.2452	.3923	.4506	.5296	.5690	.5948	.6552	.7733
1.00	.3935	.5105	.5760	.6304	.6647	.6870	.7388	.8413
1.25	.5422	.6410	.6789	.7348	.7617	.7792	.8196	.8943
1.50	.6754	.7314	.7804	.8115	.8320	.8452	.8757	.9332
1.75	.7837	.8335	.8510	.8800	.8938	.9027	.9233	.9599
2.00	.8648	.8932	.9104	.9235	.9326	.9386	.9522	.9772
2.25	.9204	.9398	.9458	.9571	.9624	.9659	.9738	.9878
2.50	.9561	.9656	.9714	.9756	.9787	.9807	.9854	.9938
2.75	.9772	.9830	.9846	.9880	.9895	.9905	.9929	.9970
3.00	.9889	.9910	.9929	.9939	.9947	.9952	.9964	.9987

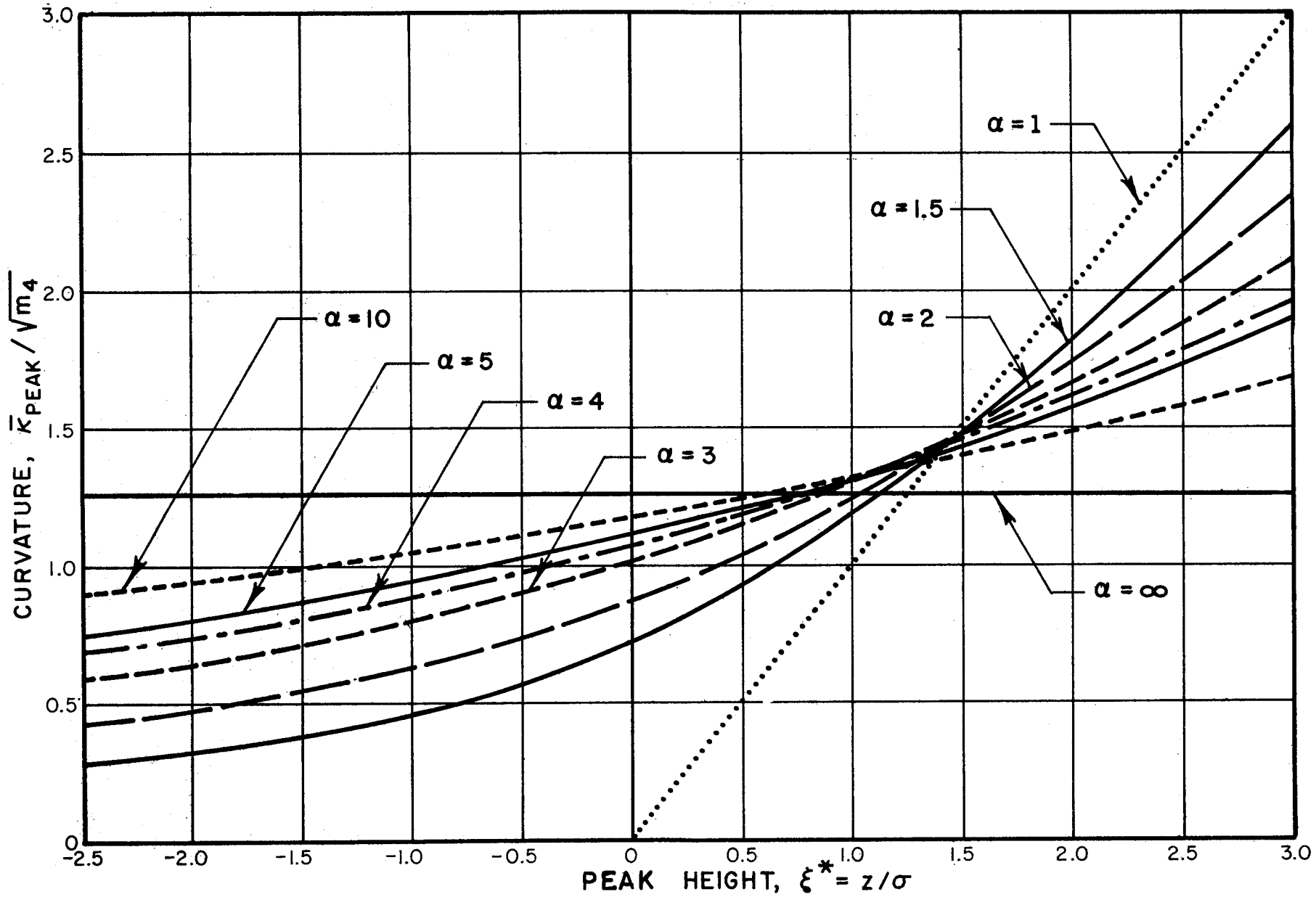


FIG. 5-13 EXPECTED VALUE OF DIMENSIONLESS TIP CURVATURE FOR PEAKS ON A PROFILE

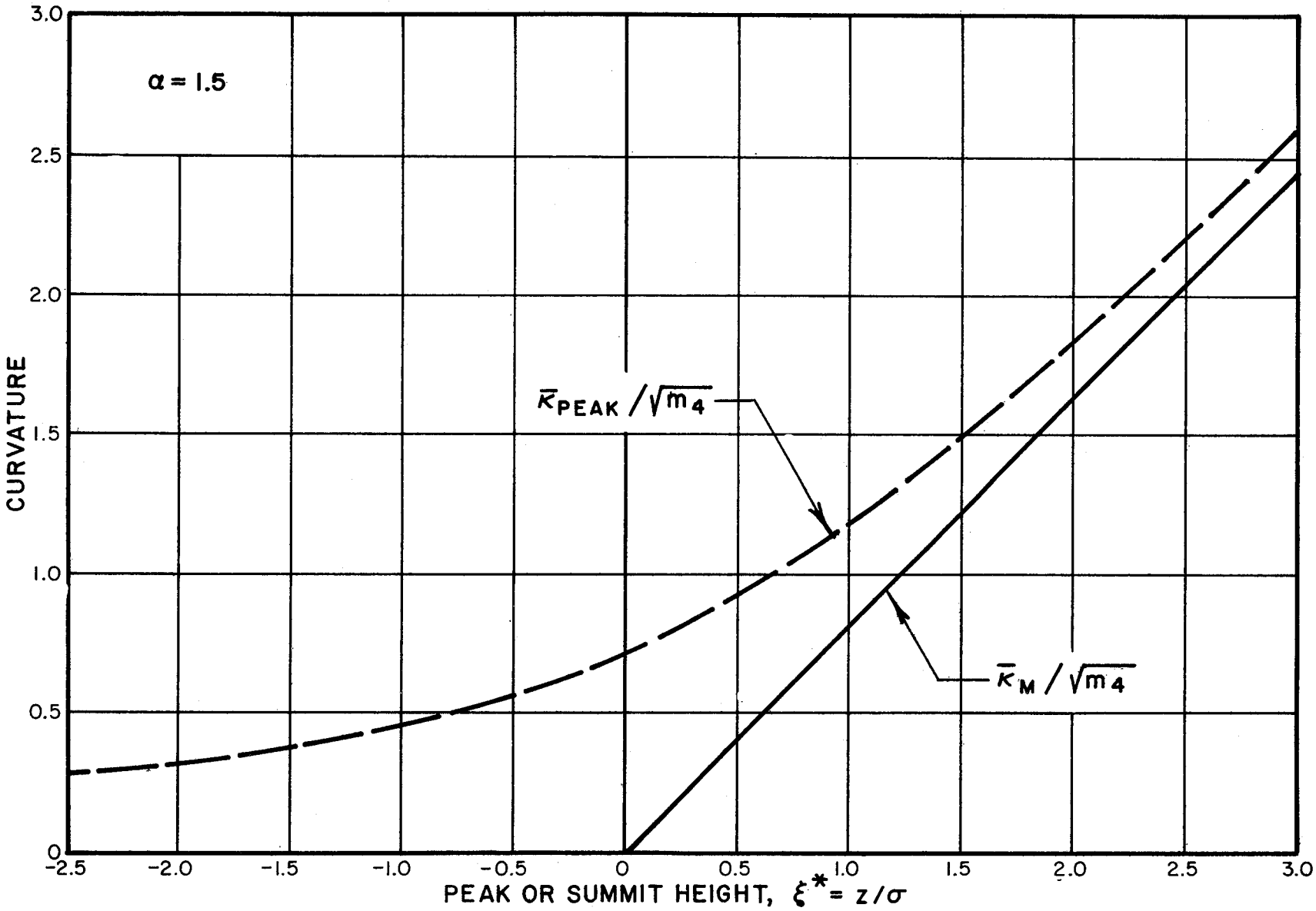


FIG.5.14 COMPARISON OF EXPECTED PEAK AND SUMMIT CURVATURES, AS FUNCTIONS OF PEAK AND SUMMIT HEIGHTS

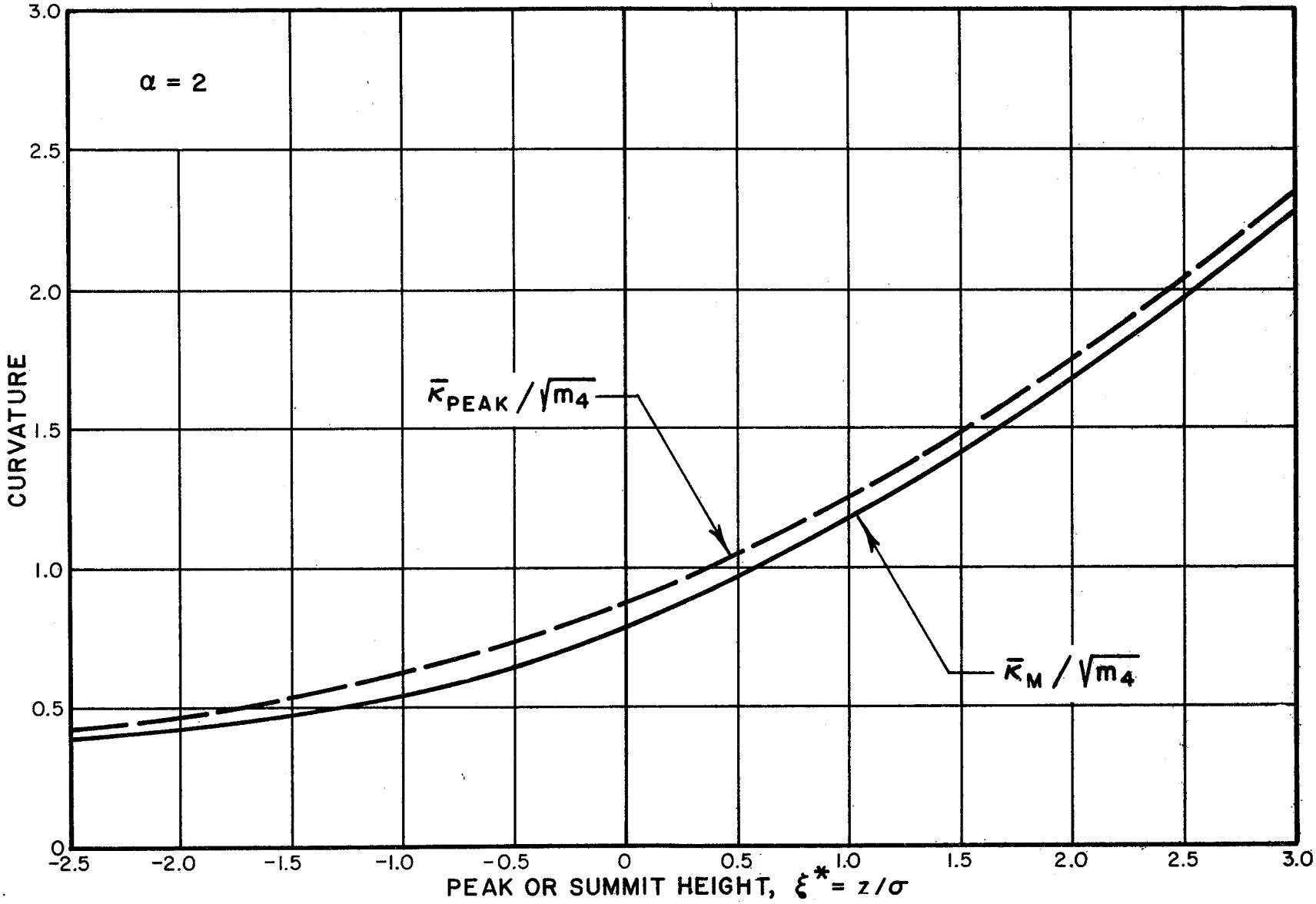


FIG.5.15 COMPARISON OF EXPECTED PEAK AND SUMMIT CURVATURES

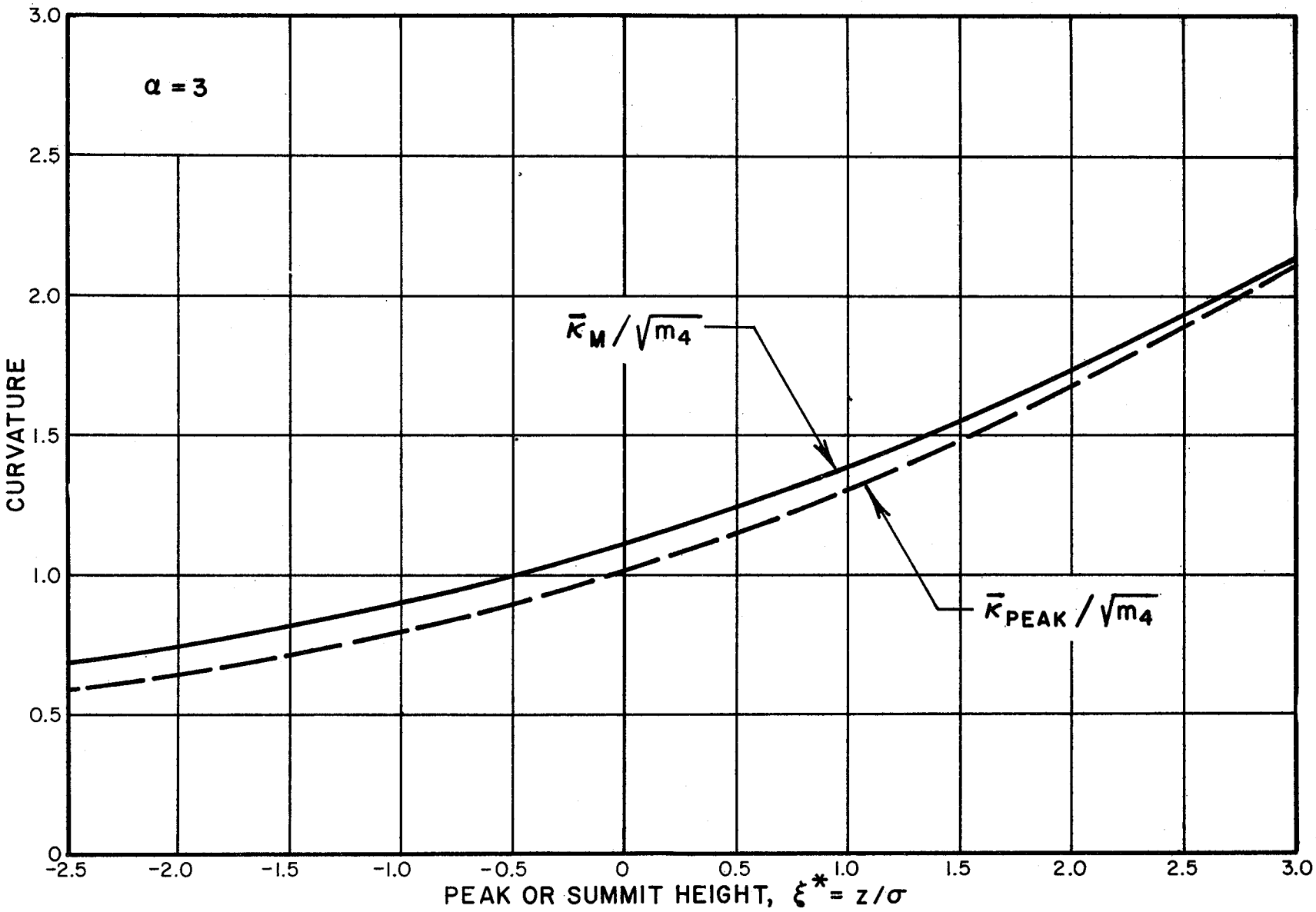


FIG.5.16 COMPARISON OF EXPECTED PEAK AND SUMMIT CURVATURES

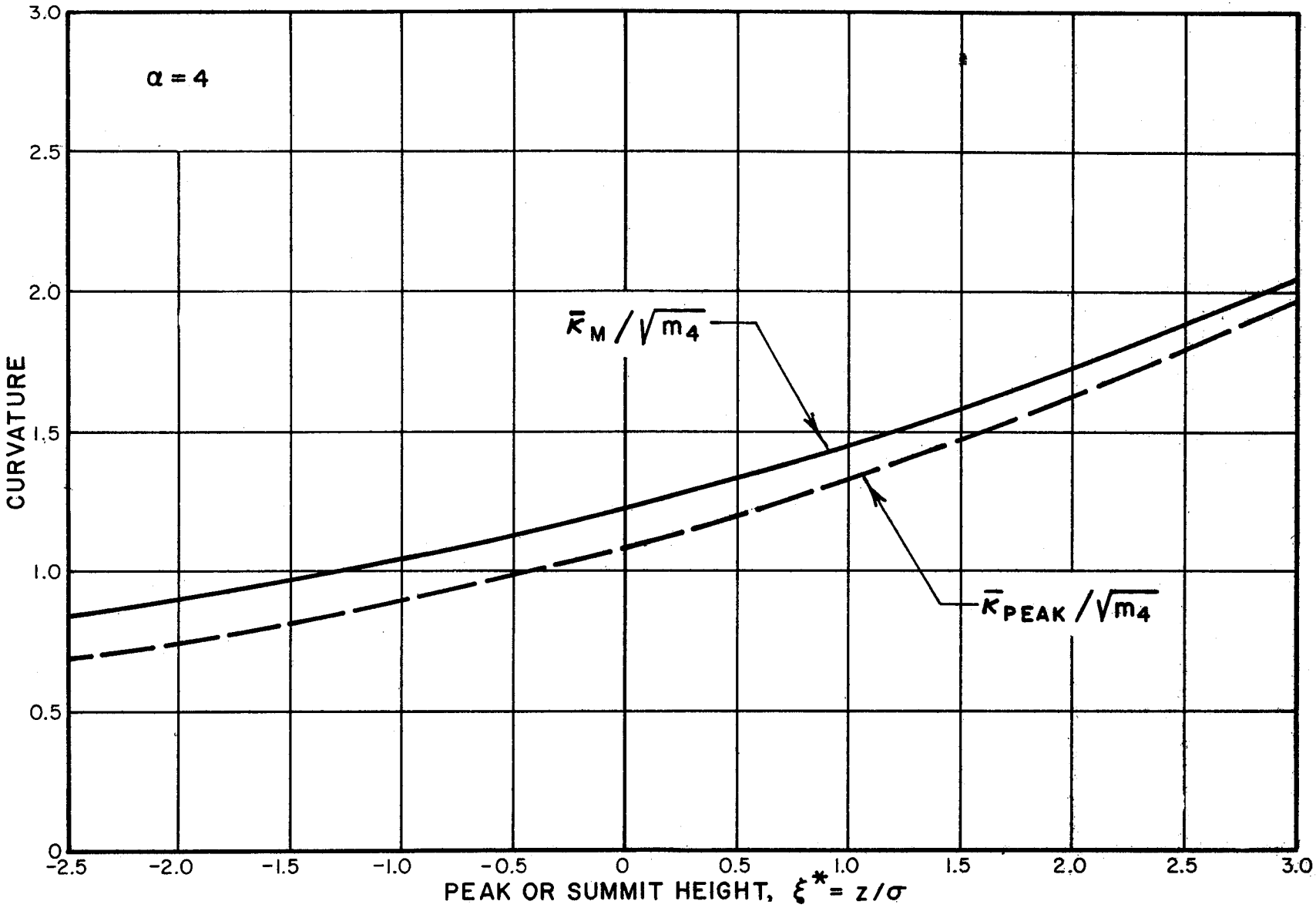


FIG.5.17 COMPARISON OF EXPECTED PEAK AND SUMMIT CURVATURES

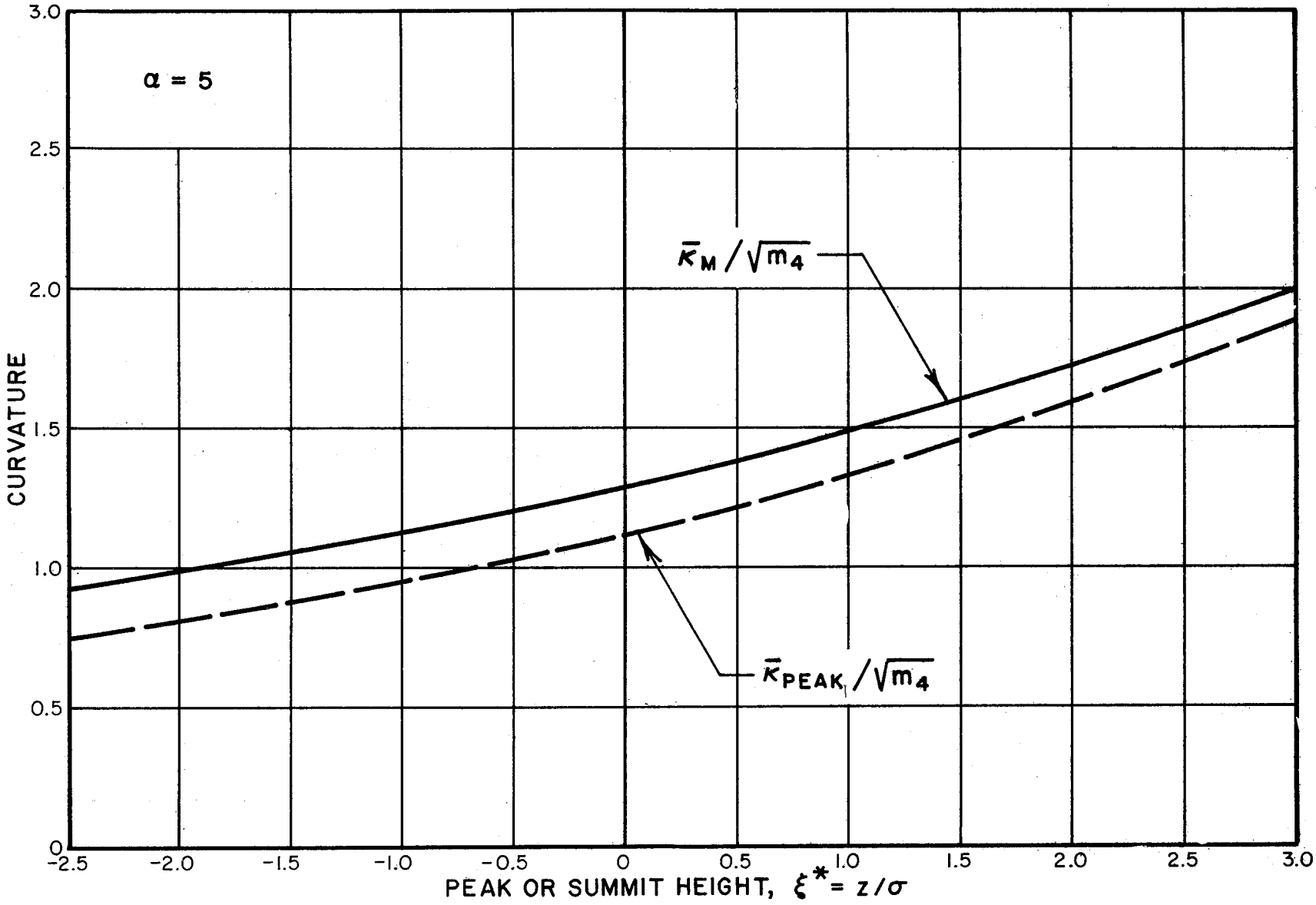


FIG.5.18 COMPARISON OF EXPECTED PEAK AND SUMMIT CURVATURES

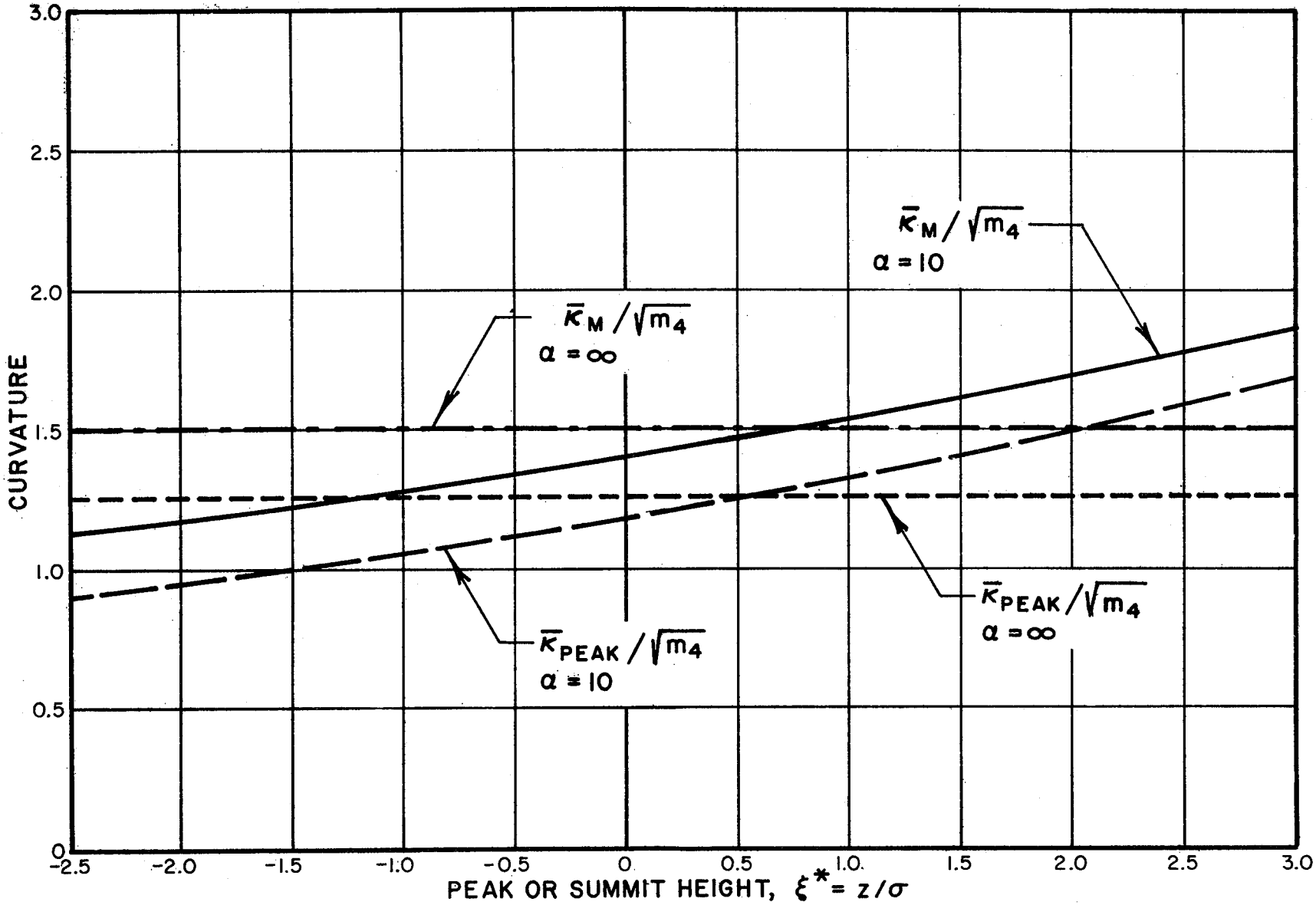


FIG.5.19 COMPARISON OF EXPECTED PEAK AND SUMMIT CURVATURES

5.4.3 The profile slope

Consider a profile taken along the x axis. The height of the profile is $\xi_1 = z$ and the slope is $\xi_2 = \partial z/\partial x$. The joint probability density for ξ_1 and $\xi_2 = \partial z/\partial x$ is found from Eq. 17 to be

$$p(\xi_1,\xi_2) = p(\xi_1)p(\xi_2) \ , \tag{83}$$

since $\overline{\xi_1\xi_2} = 0$. $p(\xi_2)$ is found, by the techniques detailed in Sec. 5.3 to be

$$p(\xi_2) = \frac{1}{\sqrt{2\pi m_2}} \exp[-\xi_2^2/2m_2] \ . \tag{84}$$

Equation 83 indicates that the probability density for surface slopes in a given direction is statistically independent of the surface elevation at which the slope is measured. From Eq. 84, the expected value of the absolute slope $\overline{|\xi_2|}$ at any elevation is

$$\overline{|\xi_2|} = \left(\frac{2}{\pi m_2}\right)^{\frac{1}{2}} \int_0^\infty \xi_2 \exp(-1/2\ \xi_2^2/m_2)d\xi_2 = \left(\frac{2m_2}{\pi}\right)^{\frac{1}{2}} \ . \tag{85}$$

Similar comments apply to ξ_3, the surface slope in the direction of the y axis.

A comparison of Eqs. 69 and 85 indicates that the mean gradient on the surface is always larger than the mean slope on the profile:

$$\overline{\zeta} = \frac{\pi}{2}\ \overline{|\xi_2|} = \frac{\pi}{2}\ \overline{|\xi_3|} \ . \tag{86}$$

5.5 Discussion

Whitehouse and Archard [13] have recently provided a detailed discussion of three-point analyses of surface profiles. We shall briefly examine the connection between their work and ours. Their theory assumes an autocorrelation function of the form

$$R(r) = m_0 \exp(-\beta|r|) , \tag{87}$$

where β is a factor governing the swiftness of decay of the correlation. Using Eq. 8, the profile PSD is found to be

$$\Phi(k') = \frac{m_0}{2\pi} \cdot \frac{2\beta}{\beta^2 + k'^2} . \tag{88}$$

It may be seen that when β is large, the autocorrelation decays switftly, and the PSD is almost flat out to large values of k'. We may now attempt to obtain the parameters m_2 and m_4, using Eq. 13. Upon introducing Eq. 88 into Eq. 13, however, we find a result that is well-known in the theory of Markov processes: for a random process with an exponential autocorrelation function, the mean-square slope (m_2) and the mean-square second derivative (m_4) are undefined. This can be traced directly to the fact that for an exponential autocorrelation function, Eq. 87, the second and fourth derivatives are undefined at $x = 0$. The relation inverse to Eq. 8 is

$$R(r) = \int_{-\infty}^{\infty} \Phi(k') \exp(ik'r)dk' . \tag{89}$$

Differentiating Eq. 89 and combining it with Eq. 13, we obtain

$$\left(\frac{d^2R}{dr^2}\right)_{r=0} = - \int_{-\infty}^{\infty} (k')^2 \Phi(k')dk' = - m_2 \tag{90}$$

and

$$\left(\frac{d^4R}{dr^4}\right)_{r=0} = -\int_{-\infty}^{\infty} (k')^4 \Phi(k') dk' = m_4 \ . \qquad (91)$$

Thus, in order for the parameters m_2 and m_4 (and therefore α) to exist, the autocorrelation function must be smooth at the origin, in the sense that its second and fourth derivatives exist. It is entirely likely that R(r) is exponential for large r; however, extrapolation of this behavior to small values of r is manifestly unsafe. There is thus an inherent theoretical contradiction in Whitehouse and Archard's work: their mathematical model does not allow slopes and curvatures to exist, though they proceed to obtain these data from profiles. The reason why this does not amount to a contradiction in practice is that their sampling interval is finite; the effect of a finite sampling interval being to filter out small-wavelength components, and to change the behavior of the autocorrelation function at the origin.

In order to examine the consequences of this filtering out of small wavelengths (i.e., large wave-numbers), consider a profile PSD of the form

$$\Phi(k') = \begin{cases} \dfrac{C}{\beta^2 + k'^2} , & |k'| \leq k_0 \\ 0 , & |k'| > k_0 \end{cases} \qquad (92)$$

Using Eq. 13, we find

$$m_0 = \frac{2C}{\beta} \tan^{-1} (k_0/\beta) ,$$

and $$m_2 = 2C[k_0 - \beta \tan^{-1} (k_0/\beta)] ,$$

$$m_4 = 2C\left[\frac{k_0^3}{3} - \beta^2 k_0 + \beta^3 \tan^{-1} (k_0/\beta)\right] . \qquad (93)$$

The parameter α can now be obtained from Eq. 45, and is found to be

$$\alpha = \tan^{-1}A \left[\frac{1}{3}A^3 - A + \tan^{-1}A\right]/(A - \tan^{-1}A)^2 , \qquad (94)$$

where

$$A = (k_0/\beta) . \qquad (95)$$

It may be seen that $\alpha \to \infty$ as $A \to \infty$. In this limit, the peak-height distribution is found from Fig. 5.6 to be Gaussian, a result that agrees with that of Whitehouse and Archard [13]. For other values of A, α may be obtained from Eq. 95; profile and surface statistics may then be obtained with the techniques described in this paper.

5.6 Conclusions

To reiterate the main theoretical theme of this paper, the extremal statistics of a random surface must be distinguished from the statistics of a profile of the surface. All the information necessary for the analysis of random, isotropic, Gaussian surfaces is contained in the power spectral density of a profile in an

arbitrary direction. The analytical techniques for obtaining the statistics of the surface are given in Sec. 5.3. A simple technique for obtaining the parameters necessary for this analysis is described in Sec. 5.4.

In general, it is found that the profile, if interpreted simplistically, indicates a lower probability for high summits, a smaller summit curvature and a smaller mean gradient than actually exist on the surface. The implications of this distortion for problems involving contact of rough surfaces is obvious.

It may be noted that the theory outlined in this paper is, in principle, easy to extend to nonisotropic Gaussian surfaces. The development of the theory remains valid up to Eq. 27; beyond this point, it is necessary to do without the use of Eq. 16, which imply isotropy. Longuet-Higgins gives the density of summits for nonisotropic Gaussian surfaces [*2*], and discusses in detail surface slopes and gradients [*10*]. However, results are not available for probability densities for summit heights and mean curvatures.

An interesting fact about nonisotropic surfaces is that one needs nine constants to proceed with an analysis analogous to ours: m_{00}, m_{20}, m_{02}, m_{11}, m_{13}, m_{31}, m_{22}, m_{40}, and m_{04}. For a profile along any direction θ_0, three equations containing these nine constants may be written, using Eq. 14:

$$\left.\begin{aligned} m_{0\theta_0} &= m_{00} \\ m_{2\theta_0} &= m_{20}\cos^2\theta_0 + 2m_{11}\cos\theta_0\sin\theta_0 + m_{02}\sin^2\theta_0 \\ m_{4\theta_0} &= m_{40}\cos^4\theta_0 + 4m_{31}\cos^3\theta_0\sin\theta_0 + 6m_{22}\sin^2\theta\cos^2\theta_0 \\ &\quad + 4m_{13}\cos\theta_0\sin^3\theta_0 + m_{04}\sin^4\theta_0 \end{aligned}\right\} \qquad (96)$$

In order to obtain the nine required constants, therefore, it is necessary to have three profiles along three different directions. Nine equations may then be written (three for each direction) and the constants obtained from these equations.

A final observation that may be made is that all the higher-order surface statistics of interest depend only on the parameters m_0, m_2 and m_4, obtained from a single profile for Gaussian, isotropic surfaces, and from three nonparallel profiles for Gaussian, nonisotropic surfaces. Therefore, it seems worthwhile to investigate the possibility of defining the surface finish of solids by these three parameters. The suggestion is particularly attractive because of the ease with which the three parameters may be obtained from a profile, as described in Sec. 5.4.

APPENDIX

The expressions $I_0 \cdots I_5$ appearing in Eq. 63 are as follows:

$$I_0 = \left(\frac{\pi}{2C_1}\right)^{\frac{1}{2}} \exp(-1/2\ \xi^{*2})(1+\mathrm{erf}\beta)\ , \tag{A1}$$

$$I_1 = 1/C_1[\exp(-C_1\xi^{*2}) + \beta\exp(-1/2\ \xi^{*2})\ \sqrt{\pi}(1+\mathrm{erf}\beta)]\ , \tag{A2}$$

$$I_2 = \sqrt{2}/C_1^{3/2}[\beta\exp(-C_1\ \xi^{*2}) + \sqrt{\pi}\ \exp(-1/2\ \xi^{*2}) \times (1+\mathrm{erf}\beta)(\beta^2 + 1/2)]\ , \tag{A3}$$

$$I_3 = 2/C_1^2[(1+\beta^2)\exp(-C_1\xi^{*2}) + \sqrt{\pi}\ \exp(-1/2\ \xi^{*2}) \times (1+\mathrm{erf}\beta)(\beta^3+3\beta/2)]\ , \tag{A4}$$

$$I_4 = \frac{1}{(C_1+1)^{\frac{1}{2}}} \cdot \left(\frac{\pi}{2}\right)^{\frac{1}{2}} \exp\{-[\alpha\xi^{*2}]/[2(\alpha-1)]\}(1+\mathrm{erf}\gamma) \tag{A5}$$

$$I_5 = \frac{1}{C_1+1}\left[\exp(-C_1\xi^{*2}) + \gamma\ \exp\{-[\alpha\xi^{*2}]/[2(\alpha-1)]\}\ \sqrt{\pi}(1+\mathrm{erf}\gamma)\right]\ , \tag{A6}$$

C_1 being defined in Eq. 45, and β and γ being defined in Eq. 51.

REFERENCES

1. J.B.P. Williamson, "The Shape of Solid Surfaces," in *Surface Mechanics*, Proceedings of the ASME Annual Winter Meeting, Los Angeles, California, 16-20 November 1969.

2. M.S. Longuet-Higgins, "The Statistical Analysis of a Random Moving Surface," *Philosophical Transactions of the Royal Society* **249**, Series A, 1957, pp. 321-387.

3. M.S. Longuet-Higgins, "Statistical Properties of an Isotropic Random Surface," *Philosophical Transactions of the Royal Society* **250**, Series A, 1957, pp. 157-174.

4. M.G. Cooper, B.B. Mikic, and M.M. Yovanovich, "Thermal Contact Conductance," *International Journal of Heat and Mass Transfer* **12**, 1969, pp. 279-300.

5. D.J. Whitehouse and J.F. Archard, "The Properties of Random Surfaces in Contact," in *Surface Mechanics*, Proceedings of the ASME Annual Winter Meeting, Los Angeles, California, 16-20 November 1969.

6. T.E. Tallian, Y.P. Chiu, D.F. Huttenlocher, J.A. Kamenshine, L.B. Sibley, and N.E. Sindlinger, "Lubricant Films in Rolling Contact of Rough Surfaces," *Transactions of the American Society of Lubrication Engineers* **7**, 1964, pp. 109-126.

7. J.B.P. Williamson and R.T. Hunt, "Microtopography of Surfaces," *Proceedings of the Institution of Mechanical Engineers* **182** Part 3K, 1967-1968, p. 21.

8. W. Feller, *An Introduction to Probability Theory and Its Applications* **2**, John Wiley & Sons, Inc., New York, 1966, p. 302.

9. I.S. Sokolnikoff, *Tensor Analysis*, John Wiley & Sons, Inc., New York, 1964, p. 190.

10. M.S. Longuet-Higgins, "The Statistical Geometry of Random Surfaces," in *Hydrodynamic Stability*, The Proceedings of the 13th Symposium on Applied Mathematics, American Math Society, 1962.

11. D.E. Cartwright and M.S. Longuet-Higgins, "The Distribution of the Maxima of a Random Function," *Proceedings of the Royal Society* 237, Series A, 1956, pp. 212-232.

12. S.O. Rice, "Mathematical Analysis of Random Noise," in *Selected Papers on Noise and Stochastic Processes*, Wax, N. (Ed.), Dover Publications Inc., New York, 1954, p. 211.

13. D.J. Whitehouse and J.F. Archard, "The Properties of Random Surfaces of Significance in Their Contact," *Proceedings of the Royal Society* 316, Series A, 1970, pp. 97-121.

6. REVIEW OF EXPERIMENTAL WORK DONE AT MIT

Summary

Experimental work done at MIT on creep and traction in rolling contact in the period between September 1969 and August 1970 [*1,2,3*] is reviewed in the light of work done at Bolt Beranek and Newman Inc. and elsewhere [*4,5*]. Our conclusions from this review are as follows:

1. When the surfaces are reasonably clean, surface roughness does not have an effect on the creep-tractive force relationship for the normal loads encountered in railway applications.

2. Under fluctuating normal or tractive load, the steady-state theory of creep is valid at all times if the period of fluctuation is large compared to the time to roll one contact length.

3. Although relationships between a dimensionless creep and a dimensionless tractive load may be valid under a wide variety of conditions, there is an important unknown factor in this relationship, the friction coefficient.

4. The friction coefficient increases as the normal load increases, all other parameters being fixed. There are two mechanisms which might account for this behavior: (a) At high loads, the reduction in friction due to asperity vibrations is smaller than at low loads. (b) At high loads, there is a greater probability of surface asperities breaking through a contaminant film and making metal-to-metal contact, with a resultant increase in adhesion. If the second mechanism dominates, then surface roughness should have a strong effect on the friction coefficient. This is found to be true.

5. At any level of vibration of the contacting surfaces, increasing the surface roughness increases the friction coefficient.

6. Results on the relationship between the velocity of rolling and the friction coefficient are inconclusive. MIT's experiments indicate a reduction in the friction coefficient as the rolling velocity increases. However, Barwell [4] obtained a constant friction coefficient over a fairly wide range of rolling velocities. Since the range of variables used for the two experiments are different, however, further studies are desirable in order to determine how and why rolling velocity influences the friction coefficient.

7. Studies of the effects of a lubricant on dimensionless creep and on friction are also inconclusive. MIT's results indicate a wide scatter in the data, with no correlation given, whereas Halling and Al-Quishtaini [5] found fairly repeatable results, and were able to provide a reasonable correlation. The range of rolling velocities used in the two experiments were widely different, however, and further experimental work is desirable. This work should aim at correlating creep and adhesion with the following variables:

   a) normal pressure

   b) rolling velocity

   c) oil viscosity (including temperature and pressure coefficients)

   d) surface roughness, which, we believe, is an important factor in the behavior of heavily contaminated surfaces

   e) degree of running-in (under lubricated conditions) of the rolling surfaces

f) rate of oil flow into the contact region (or alternatively, the oil film thickness).

8. The effects of high-frequency normal load variation (which should be distinguished from the effects of microvibrations of the surface and its asperities) are not clearly known. With a sinusoidal variation of normal load, there appear to be frequencies at which the friction coefficient varies widely. These frequencies are in the neighborhood of a "contact resonance frequency", a concept which needs to be studied more precisely, and of suspension resonance frequencies. However, the more important case of a fairly broad-band random variation of the normal load is of greater interest, and needs to be studied.

6.1 Introduction

A series of experimental investigations of friction and creep in rolling contact were performed at MIT for the Office of High Speed Ground Transportation under the direction of Professor Igor Paul. The following is a review and evaluation of those experimental data of both a quantitative and qualitative nature that were available to us [*1,2,3*].

The experimental apparatus is described in detail in Refs. 1 and 2. For our purpose, it is adequate to note the following features of the apparatus used in [*1*] and [*2*]. One of the rolling elements had a spherical rolling surface of 1 in. radius. The other was a cylinder of 5 in. radius. Both were made of easy-machining steel. A torque was applied to the cylinder. Longitudinal creep was obtained by applying a braking torque to the spherical element about its rolling axis. Transverse creep was obtained by giving the cylinder a harmonic transverse displacement

$$y(t) = \frac{1}{24} \sin(\pi t/4) \text{ ft}, \tag{1}$$

where t is the time in seconds.

All forces were measured with a three-dimensional dynamometer holding the bracket in which the axle of the spherical element was mounted. Acceleration measurements were made with an accelerometer mounted on this bracket.

In the experiments described in [*1*] and [*2*], acceleration levels were increased only by increasing the level of inherent vibrations of the apparatus. This was accomplished by increasing

the rpm of the cylinder. In some subsequent experiments, not available to us, external vibration was introduced by shaking the bracket holding the dynamometer with an electromagnetic shaker.

Harmonically varying normal loads were obtained by forcing the bracket holding the dynamometer with a rotating cam.

6.2 Definitions and Symbols

To clarify the subsequent discussion, we present a brief (and slightly modified) glossary of the terms and a list of the symbols used in [*1, 2, 3*].

a: semi-axis of contact ellipse in direction of rolling

b: semi-axis of contact ellipse in transverse direction

E: Young's Modulus, = 3×10^7 psi

f: friction coefficient

F_{Tx}: tractive force in direction of rolling

F_{Ty}: tractive force in transverse direction

F_Z: normal force

n: rpm of spherical element

N: rpm of cylinder

r: rolling radius of spherical element (1 in.)

R: rolling radius of cylinder (5 in.)

V: rolling velocity, ft/sec = $(2\pi RN/60)$

V_{Sx}: longitudinal slip velocity

V_{Sy}: transverse slip velocity

μ: rms roughness of a surface

ξ_x: longitudinal creep = (V_{Sx}/V)

ξ_y: transverse creep = (V_{Sy}/V)

ν: Poisson's ratio $\approx$ 0.3

Adhesion: Term describing the phenomenon underlying the fact that a force is required to separate two surfaces (or surface elements) that have been pressed together.

Rolling Friction Coefficient: Ratio of the maximum tractive force that may be transmitted without gross slip between two rolling elements to the normal force.

Sliding Friction Coefficient: Ratio of tractive force to normal force for two sliding elements. Both the sliding and rolling friction coefficients depend on the adhesive characteristics of the contacting surfaces. However, the term "friction" should be distinguished from the term "adhesion".

Unworn Surface: Newly prepared, virgin surface that has not been placed in contact with another surface.

Worn Surface: Surface subjected to rolling contact with a normal load but with insignificant tractive loads. Unworn surfaces used here are fairly symmetric about the mean plane; worn surfaces are highly asymmetric, due to squashing of the surface asperities. The amount of wear is small.

Tracked Surface: Surface obtained by making longitudinal slip persist for rolling surfaces. A significant amount of wear takes place, resulting in a visible track on the surfaces. Due to the wear, the surfaces tend to be symmetric about the mean plane.

Clean Surface: Surface cleaned with lint-free fabric using trichlorethylene.

Contaminated Surface: Surface smeared with water or motor oil.

6.3 Experimental Results

Urciuoli [1] made a study of some factors affecting adhesion in rolling contact. His results are presented in Figs. 6.1 – 6.3.

Figure 6.1 shows the effect on sliding friction of repeated sliding (obtained by moving the cylinder along its axis) for a worn and an unworn surface. Figure 6.2 shows the variation in rolling friction with the rms acceleration level, for a worn surface. Figure 6.3 shows the variation in rolling friction with the rms acceleration level for two tracked surfaces, one rougher than the other. It must be noted that in obtaining the results of both Figs. 6.2 and 6.3, the acceleration level was increased by increasing the rolling velocity. Specific values of rolling velocity corresponding to each acceleration level were not reported.

Prabhu [2] studied (a) the effect of rolling velocity on the dimensionless creep-dimensionless tractive force relationship, (b) the effect of rolling velocity on rolling friction with transverse slip, and (c) low-frequency fluctuating transverse creep. Some of his results are shown in Figs. 6.4 – 6.6. In an additional study of the effects of surface roughness (not reported quantitatively in [2]), Prabhu found that "variation of the surface roughness of the contacting elements has no effect on the creep (both longitudinal and transverse) at the loads dealt with in rail travel".

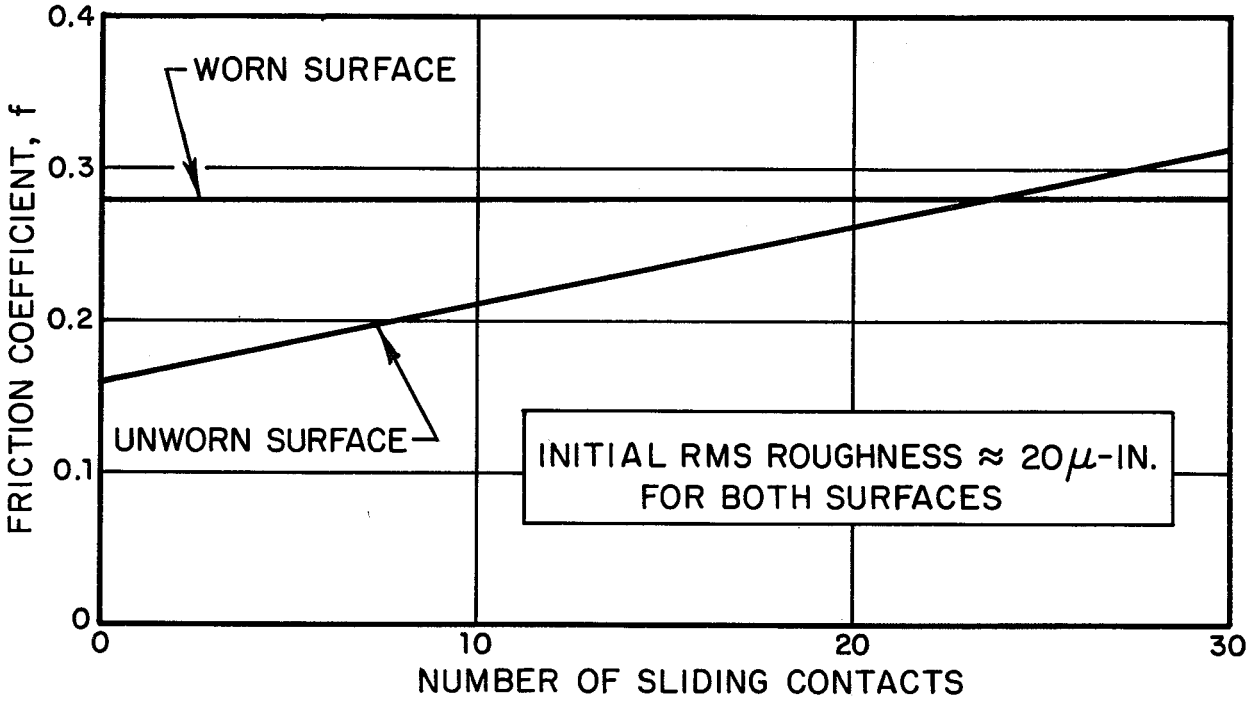


FIG. 6.1 EFFECT OF REPEATED SLIDING ON SLIDING FRICTION OF WORN AND UNWORN ROUGH SURFACES.

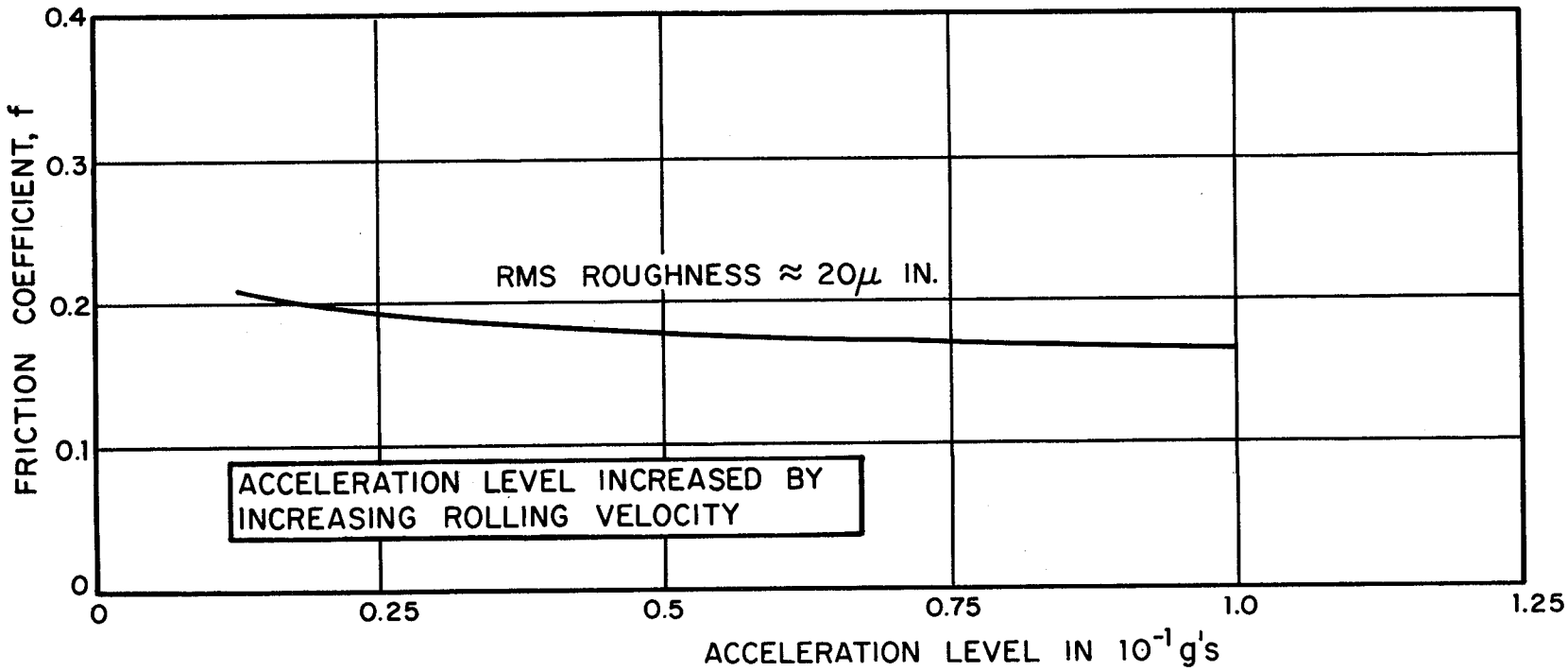


FIG. 6.2 DRY ROLLING FRICTION FOR A "WORN" ROUGH SURFACE AS A FUNCTION OF RMS ACCELERATION LEVEL AT WHEEL AXLE. ROLLING VELOCITY IS NOT CONSTANT.

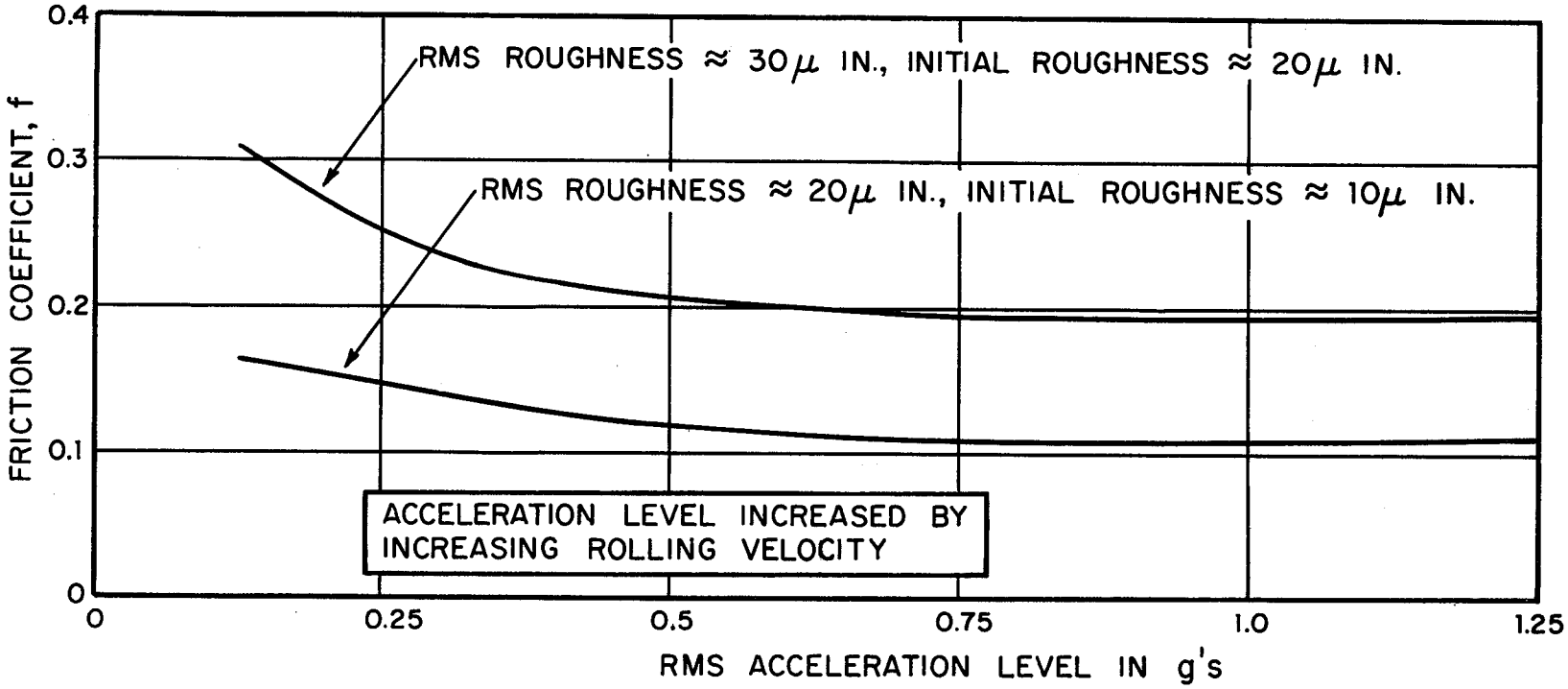


FIG. 6.3 DRY ROLLING FRICTION FOR "TRACKED" ROUGH AND SMOOTH SURFACES AS A FUNCTION OF RMS ACCELERATION LEVEL AT WHEEL AXLE. ROLLING VELOCITY IS NOT CONSTANT.

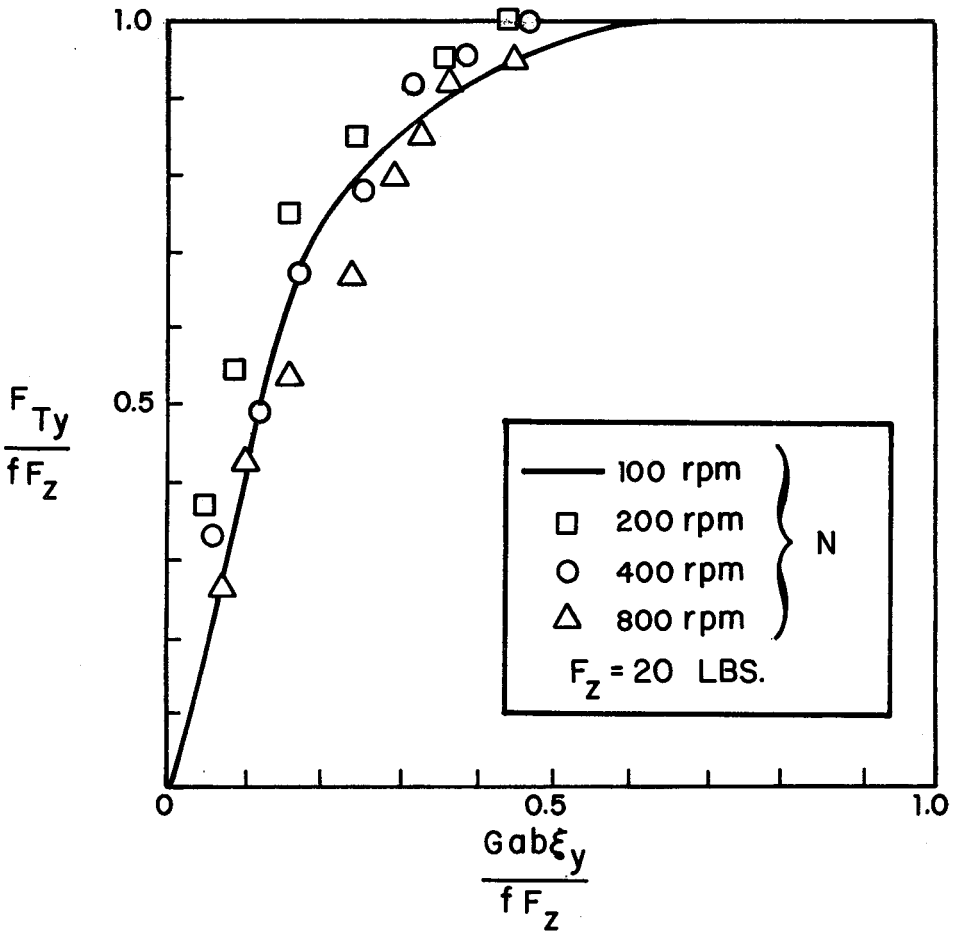


FIG. 6.4 DIMENSIONLESS TRANSVERSE CREEP AT VARIOUS SPEEDS.

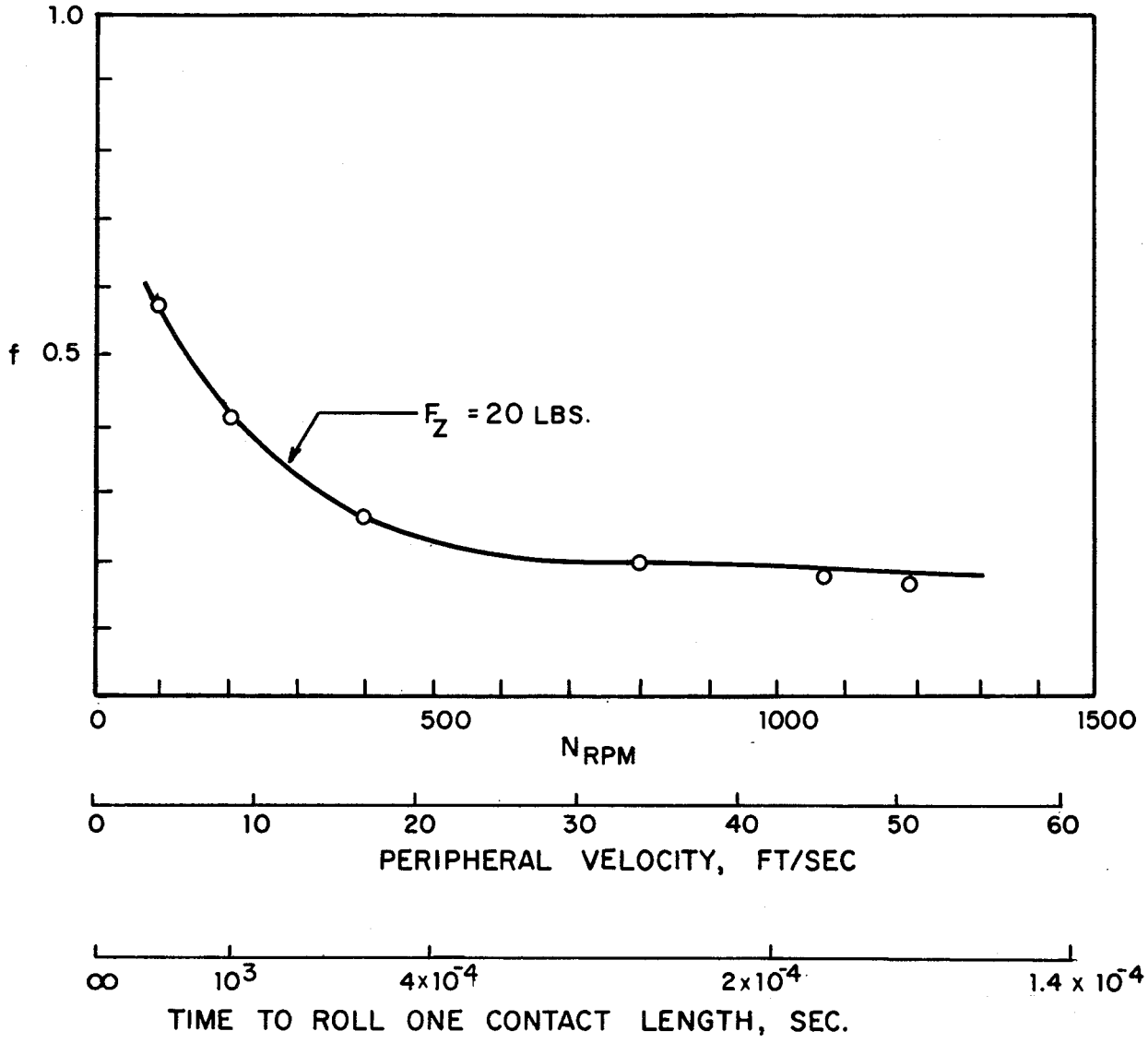


FIG. 6.5 EFFECT OF VELOCITY OF ROLLING ON TRANSVERSE FRICTION COEFFICIENT

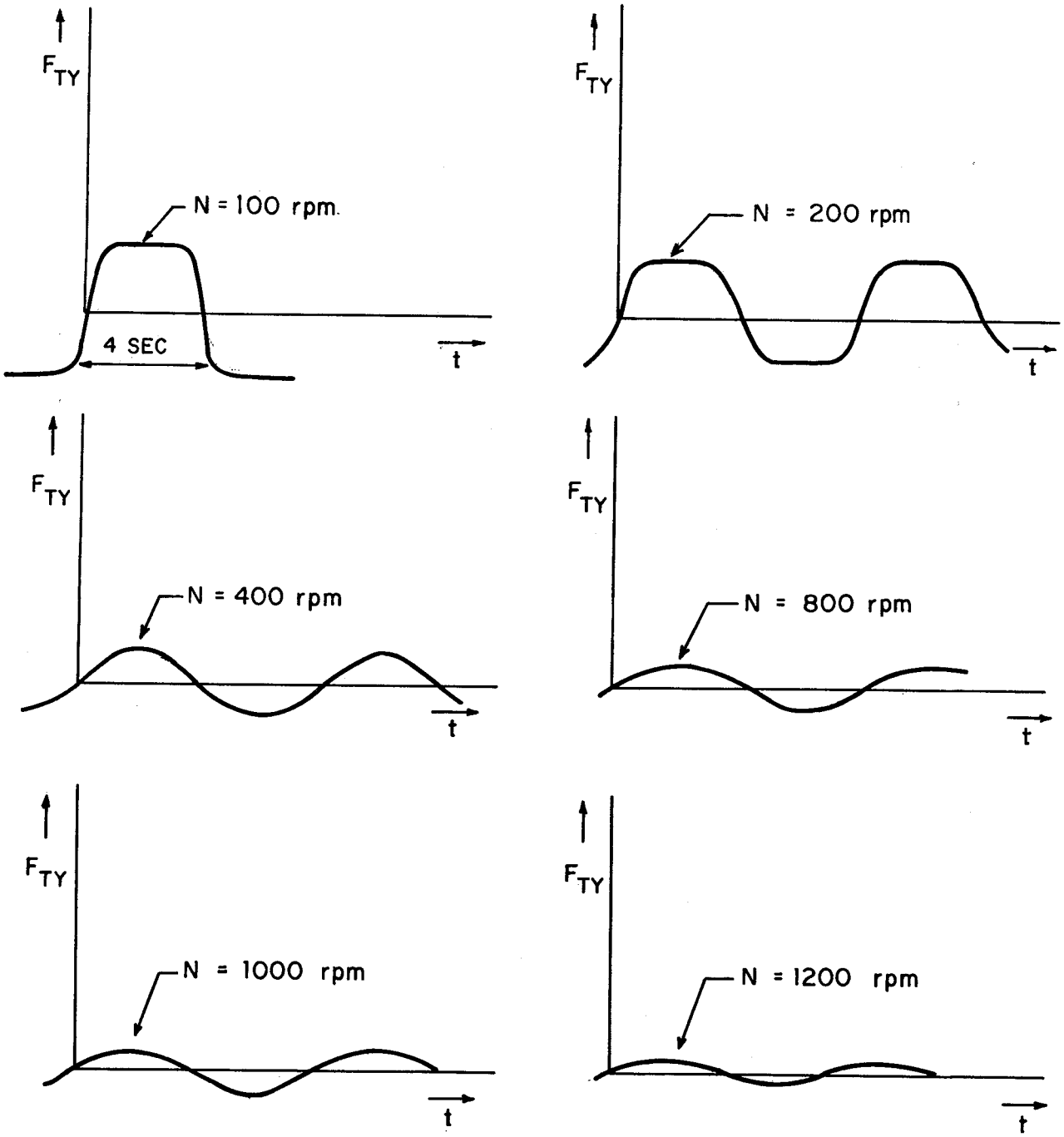


FIG. 6.6 HISTORY OF TRANSVERSE FORCE AT DIFFERENT ROLLING VELOCITIES. TRANSVERSE CREEP DECREASES AS ROLLING VELOCITY INCREASES, CAUSING A DECREASE IN THE TRANSVERSE FORCE.

Figure 6.4 shows experimentally measured dimensionless transverse creep as a function of the dimensionless transverse tractive force, for different cylinder rpm's. Figure 6.5 shows the rolling friction coefficient f as a function of the cylinder rpm. This friction coefficient is used to obtain the dimensionless quantities shown in Fig. 6.4. Also shown in Fig. 6.5 are the peripheral velocity of the cylinder, and the time to roll one contact length, which is the time available to form and break a friction junction. Figure 6.6 shows typical recorded histories of the transverse force, for different cylinder rpm's. With the transverse displacement given by Eq. 1, the transverse creep is

$$\xi_y = \frac{V_{Sy}}{V} = \frac{(\pi/96)\cos(\pi t/4)}{(2\pi \times 5 \times N/60)} = \frac{\cos(\pi t/4)}{16N} \quad . \tag{2}$$

The maximum value of the transverse creep is

$$\xi_{y,max} = (1/16N) \quad . \tag{3}$$

As N is decreased, $\xi_{y,max}$ increases until at some value of N (between 200 and 400 rpm), gross slip occurs, and no more tractive force can be transmitted. This accounts for the flattening of the traces for N=100 and 200 rpm in Fig. 6.6.

The data shown in Fig. 6.4 were obtained by plotting the transverse creep, obtained from Eq. 2 against the measured transverse force F_{Ty} at each time t in the creep cycle.

Dagalakis [3], among other things, studied longitudinal creep with a fluctuating normal load. Some of his data are summarized in Tables 6.1 and 6.2. In both these tables, the first column shows the frequency of the normal load variation. The second column shows the cylinder rpm, N. The third column shows

the measured longitudinal creep, 100 ξ_x. The fourth column shows the acceleration level in terms of the rms volts at the acceleration monitoring meter. The fifth column shows the minimum and maximum values of the fluctuating normal load. Also shown in the fifth column is the percentage overall variation of the normal load, $100\times\left(\frac{\text{Max-Min}}{2}\right)/\text{Mean}$. In all these experiments, the tractive force was held constant (at a value not available to us), and the normal load was F_Z = 20 lbf. The surfaces were clean.

Finally, we present verbatim the conclusions in [*3*], based on the above data and other data not available to us.

"1. At contact stress levels of interest in rail applications (>100,000 psi) the effects of surface parameters (roughness) after an initial wear-in period on creep and adhesion are small. Experiments with a full range of roughness using steel and aluminum wheels showed that at high normal forces the creep varied over a range of - 6% to + 8% compared with performance at a normal worn rail roughness (under identical, slow speed conditions).

"2. The scaling laws used in the literature for steady-state creep are valid for slow speeds (i.e., insignificant contact vibrations) and for normal and lateral force oscillations that are slow compared with the natural frequency of the contact (which is usual). Near natural frequencies of the suspension creep changes drastically and scaling is not possible. Scaling laws were verified to a maximum error of ± 8% using 1 in. to 5 in. diameter steel and aluminum wheels.

"3. Speed *per se* does not affect either creep or *adhesion* but the micro-vibrations which are increased by the increased excitation usually accompanying higher speeds are the critical factors affecting both creep and adhesion. The micro-vibrations occur at the natural frequency of the contact area and are a calculable function of the contact stiffness, nominal load, material properties, surface profile (gross profile not roughness) and mass of the system.

"4. The well-established vibration scaling laws used for model testing of vibration behavior are valid for the contact vibrations occurring at the contact surface. This was verified to be accurate to ± 2.6% using 1.5 and 4 in. diameter wheels.

"5. Contaminants wiped onto the simulated rail wheel had the following effects on steady-state creep (no significant vibrations):

a) Water at low contact pressures (<<100,000 psi) — results varied over a very large range increasing creep + 12 to 210%.

b) Water at high contact pressures (>100,000 psi) — increased the creep by + 12 to + 26%.

c) Oil (automobile oil) at low contact pressures (<<100,000 psi) — results varied considerably increasing creep by + 80 to + 270%.

d) Oil at high contact pressures (>100,000 psi) increased creep by + 14 to + 60%.

"6. Flooding the contact area with the contaminant had the following effects:

a) At low contact pressures the results were meaningless because the creep changed continuously and drastically.

b) At high contact pressures the results were comparable to those in 5b and 5d.

"7. Contaminants wiped onto the simulated rail wheel had the following effects on adhesion (no significant vibrations, conditions corresponding to 5a, b, c and d respectively):

a) Reduced by 24 to 90%.

b) Reduced by 24 to 44%.

c) Reduced by 61 to 94%.

d) Reduced by 30 to 80%.

"8. Under vibration conditions (vibrations not close to the natural frequency of the contact area or the natural frequencies of the suspension) the results for creep were comparable to those in (5) with the *minimum* contact force (normal force oscillations) determining the pressure range (i.e., < or > 100,000 psi).

"9. Under the vibration conditions [as in (8)] the effects on the adhesion were not well reproducible. Under the same vibration conditions the wheel would sometimes slip at a very low value of adhesion coefficient and then frequently regain adhesion and continue to some higher value. Even trying to correlate adhesion based on the minimum normal force (rather than the mean force) gave very inconsistent results. Very small normal force oscillations (± 10% of mean force) gave results similar to those described in (7)

but increasing the oscillations beyond these values was not possible without getting very high scatter in the results.

"10. Oscillations at or near the natural frequencies of the surface contact would produce very significant changes in creep and adhesion. Forced oscillations near this frequency (applied with an electromagnetic shaker) could reduce the creep to zero or increase it by several hundred per cent (i.e., cause gross slippage at times) depending on a combination of magnitude and phase of the applied oscillations with respect to the natural oscillations (the forced oscillations were applied by measuring the contact oscillations and feeding this signal back through a phase shifter and an amplifier to the shaker).

"11. Forced oscillations near the contact natural frequency could reduce adhesion to almost zero or increase it by a factor of *three* compared to values obtained without forced vibrations. The significant factor in each case was to find a combination of applied force amplitude and phase angle such as to reduce the micro-vibrations to as close to zero as possible.

"12. It appears certain from our work that the micro-vibrations are crucial to both creep and adhesion behavior and their control may be feasible for practical application in creep or adhesion control.

"13. An instantaneous creep measuring meter has been developed which accurately and continuously measures the creep between two rolling surfaces at high speeds."

We note that the term "adhesion" as used in [*3*] is equivalent to our "rolling friction coefficient".

6.4 Discussion

In this discussion, we shall mainly concern ourselves with the following:

1. limitations on the steady-state theory of creep,
2. effects of surface roughness on rolling friction,
3. effects of rolling velocity on rolling friction,
4. effects of suspension and other resonances on creep and rolling friction.

6.4.1 Limitations on the validity of the steady-state theory of creep

There are two prime reasons for studying creep in rolling contact: (a) investigations of the lateral instability of vehicles, and (b) investigations of wheel and rail wear. For investigations of vehicle instability, the interest centers on the creep coefficient, which is the slope of the linear part of the creep-tractive force curve.

It may be seen from the smooth-surface, Chap. 2, theory that the creep coefficient is predicted to be independent of the rolling friction coefficient, and dependent only on the geometry and elastic constants of the contacting elements. The question we phrase is, under what conditions is this prediction valid? Examining the data in Fig. 6.4, Tables 6.1 and 6.2, and those reported in [5], we present the following (tentative) answers.

1. The effects of surface roughness in dry contact on the creep coefficient are negligible for usual operating loads.

2. The effects of the velocity of rolling are negligible. This conclusion is also supported by work reported in [4]. In this reference, however, creep coefficients considerably larger than those predicted by theory are reported at all rolling velocities, the discrepancy being ascribed to vibration and plastic deformation of the asperities.

3. The effects of vibration on the creep coefficient are not clearly understood. In Fig. 6.4, different rolling velocities correspond to different vibration levels; the creep coefficient is the same for all rolling velocities. In [4], changes in the rolling velocity, which were presumably accompanied by changes in the vibration level, were also found to have no effect on the creep coefficient. These results contradict Conclusion 8 of [3], quoted above in Sec. 6.3. It would seem that the main effect of vibrations would be to reduce the rolling friction coefficient, which would not affect the creep coefficient. However, it is conceivable, for example, that vibrations could act (in a manner that is admittedly obscure at present) in such a way as to increase the apparent surface roughness (perhaps by increasing the time-averaged separation of the mean surfaces of the rolling elements) and thus reintroduce surface roughness effects such as those described in Chap. 2. These possibilities deserve further investigation. For example, it would be extremely worthwhile to try to obtain an empirical correlation between the vibration spectrum and an effective increase in surface roughness for a given situation, and to then predict creep with the rough-surface theory of Chap. 2 in other situations.

4. The steady-state creep coefficient is valid for dry contact with fluctuating transverse loads, when the period of fluctuation is large compared with time to roll one contact length, as evidenced by the data of Fig. 6.4, which were obtained from recordings such as those shown in Fig. 6.6.

5. For fluctuating normal loads in dry contact, with the period of fluctuation again large compared with the time to roll one contact length, Halling and Al-Quishtaini [5] found the steady-state creep coefficient to be valid. This agrees with Conclusion 3 of [3], quoted in Sec. 6.3.

6. For contaminated surfaces, the steady-state creep coefficient for dry contact is not valid. The results summarized in Conclusions 5, 6 and 7 of [3] appear reasonable, but do not provide a means of correlating the creep with contaminant characteristics. Halling and Al-Quishtaini [5] performed a similar experiment, using an oil with a viscosity of 13 centi-stokes at 86°F, at rolling speeds ranging from 7.57 in./min to 141 in./min. They give the following correlation for the creep, valid only for their test, but indicative of the form of correlation to be expected:

$$\xi_x = 7.7\times10^{-3}\left(\frac{V}{F_Z}\right)^{0.23}\left(\frac{F_{Tx}}{F_Z}\right)\exp\left\{25.6\left(\frac{V}{F_Z}\right)^{0.23}\left(\frac{F_{Tx}}{F_Z}\right)\right\}. \qquad (4)$$

Halling and Al-Quishtaini found, moreover, that for lubricated contacts, there is an initial running-in period (whose length they do not indicate), during which wide scatter is obtained in the data. After this running-in period, fairly repeatable data are obtained. Moreover, they postulate the existence of quasi-elastohydrodynamic conditions in the interface, with

some asperity contacts occurring. If this is so, we expect surface topography to play an important part in the creep of lubricated contacts. There is thus a need for a thorough study of the rolling contact of contaminated surfaces, aimed at correlating creep and friction with the following variables:

a) rolling radii,

b) normal pressure,

c) oil viscosity (including temperature and pressure coefficients),

d) oil-metal chemical reactivity,

e) surface roughness,

f) degree of running-in of the rolling surfaces,

g) rolling velocity,

f) oil film thickness.

It must be noted that all the mechanisms discussed above tend to reduce the rolling friction coefficient. Though this does not change the creep coefficient, it decreases the maximum value of creep for which the linear relationship between creep and traction holds.

6.4.2 Effects of surface roughness on rolling friction

The effects of surface roughness on sliding friction are well documented [*6*]. It would be surprising to find no such effects on rolling friction. Based on some of the data presented in Chap. 2, we would expect surface roughness to be most important in the rolling contact of either very clean or very contaminated surfaces, and least important for boundary-lubricated surfaces.

Figure 6.3 exhibits a definite influence of surface roughness on the rolling friction coefficient. On the other hand, the data of Halling and Brothers [7] show no correlation between the surface roughness centre-line-average and the dry rolling friction coefficient. Moreover, using steel on steel, they obtained coefficients ranging from 0.07 to 0.08 (all at very low speeds) compared with values ranging from 0.11 to 0.31 in Fig. 6.3. It is likely that such factors as the ambient humidity, the solvent used to clean the surfaces (trichlorethylene induces brittleness in steel) and the metallurgical composition of the steel strongly influence the rolling friction coefficient. Some data in support of this statement are presented by Barwell [4].

6.4.3 Effects of rolling velocity on rolling friction

As discussed in detail in Chaps. 2, 3 and 4, increases in rolling velocity are likely to reduce the dry rolling friction coefficient due to two causes: (a) due to the decreased strength of the friction junctions, resulting from the smaller time spent in the contact region, and (b) due to increased asperity vibrations. Unfortunately, no experiments aimed at singling out these effects have been performed to date. Figure 6.5 shows a marked decrease in the rolling friction coefficient as the rolling velocity increases; however, the acceleration level also increased with the rolling velocity. The data of Barwell [4] indicate no change in the rolling friction coefficient, even though significant vibration was said to be present (levels were not indicated). Thus, Conclusion 3 of [3] (see Sec. 6.3) does not at present seem justified.

One experiment aimed at isolating the two effects mentioned above would be along the following lines. Make friction measurements at high speed, along with detailed measurements of the inherent vibration spectrum. Next, make measurements at low speed, with the contact excited by a shaker so that its vibrations have approximately the high-speed spectrum. If a significantly lower friction coefficient is obtained in the first experiment than in the second, one would conclude that rate effects in the formation of junctions are important. If substantially equal friction coefficients are obtained in both experiments, asperity or surface vibrations would be shown to be the cause of the falling friction-rolling velocity curve.

Halling and Al-Quishtaini [5] report some data concerning the effect of rolling velocity on friction in a lubricated contact. At V = 7.57 in./min, f = 0.085; at 49.6 in./min, f = 0.068; at 141 in./min, f = 0.065. Thus there is a decrease in f as V increases; it also appears that f reaches some asymptotic value as V becomes large. However, these friction coefficients are to be used with care, as the slip is quite large by the time maximum traction is obtained, particularly at high speeds. A more interesting variable is the ratio of the tractive force to the normal force at some fixed value of slip, say 0.5%. At 7.57 in./min, this ratio is approximately 0.085; at 49.6 in./min, 0.064; and at 141 in./min, 0.052.

As we remarked above in discussing creep of lubricated contacts, surface topography is likely to be an important factor governing the behavior of such contacts.

6.4.4 Effects of suspension and other resonances on creep and rolling friction

The concept of contact resonance has been investigated in some detail in Chap. 2. It is shown there that at the contact resonance frequency, large displacements of the rolling surfaces occur, leading even to loss of contact. Nevertheless, from elementary statical considerations, the mean contact load must remain constant. If one accepts two widely substantiated hypotheses: (a) that the real area of contact is proportional to the normal load, and (b) the friction force is proportional to the real area of contact; then it follows that the mean friction force also does not change. There is no decrease in the friction coefficient. The only explanation for an observed decrease in friction when the contact resonance frequency is excited must be that the constant of proportionality between the friction force and the real area of contact, i.e., the shear strength of the junctions, decreases as a result of the surface vibrations.

This reasoning suggests that surface vibrations in general are important. The most important frequencies are those at which large dynamic contact loads are likely to be generated, namely various resonance frequencies, such as those of the suspension systems. In order to judge the importance of asperity vibrations in general, more detailed investigations of model asperities, such as those described in Chap. 2 are called for.

So far, we have discussed the effects of resonances on the rolling friction coefficient. Turning to creep behavior, it is our feeling that resonant vibrations should not change the creep coefficient drastically unless (a) the amplitude of the vibrations at the contact surface is very large or (b) the period of the vibrations is of the order of the time to roll one contact length.

In order to determine the magnitude of the contact oscillations, we suggest electrical contact resistance measurements, from which contact loads may be deduced. By displaying the voltage drop across the contact region on an oscilloscope, it should be possible to follow the history of the contact load. Conclusions 8 and 10 of [3] (see Sec. 6.3) seem to indicate that surface vibration amplitudes at resonance are fairly large.

TABLE 6.1 Longitudinal Creep with Fluctuating Normal Load

Frequency of Normal Load, Hz	Cylinder rpm	Creep % $100\xi_x$	Vibration Level rms Volts	Load Variation Min/Max, % lbf
0.5	108	0.15	?	9/18
	400	0.156	0.015	
	800	0.167	0.030	33%
	1000	0.211	0.030	
	1150	?	?	
	1400	0.26	0.07	
	1600	0.312	0.08	
	2020	0.22	0.09	
1.0	98	0.159	?	11/17
	410	0.157	0.01	
	820	0.1685	0.02	21.4%
	1000	0.200	0.025	
	1200	0.155	0.045	
	1500	0.267	0.09	
	1980	0.252	0.075	
1.5	112	0.1595	?	12/16
	890	0.132	0.017	
	1000	0.1905	0.025	14.3%
	1200	0.139	0.04	
	1500	0.267	0.075	
	2000	0.250	0.07	
2	200	0.139	0.008	13/16
	550	0.152	0.015	
	940	0.170	0.025	10.3%
	1190	0.140	0.04	
	1520	0.131	0.07	
	2000	0.235	0.07	

TABLE 6.2 Longitudinal Creep with Fluctuating Normal Load

Frequency of Normal Load, Hz	Cylinder rpm	Creep % $100\xi_x$	Vibration Level rms Volts	Load Variation Min/Max, % lbf
0.5	146	0.146	0.01	20/27
	410	0.143	0.012	
	790	0.149	0.020	15%
	990	0.130	0.030	
	1200	0.144	0.04	
	1450	0.153	0.08	
	1710	0.212	0.10	
	2110	0.172	0.14	
1.0	176	0.154	0.01	21/26
	506	0.1235	0.015	
	800	0.156	0.02	
	1010	0.1585	0.03	10.6%
	1190	0.153	0.04	
	1520	0.1595	0.045	
	1775	0.238	0.110	
	2070	0.1755	0.110	
1.5	132	0.158	0.00	21/25
	360	0.159	0.01	
	810	0.1545	0.025	8.7%
	1100	0.1455	0.025	
	1350	0.156	0.03	
	1520	0.263	0.1	
	1740	0.209	0.095	
	2020	0.165	0.075	
2	140	0.149	0.00	22/24
	530	0.151	0.013	
	800	0.156	0.02	4.25%
	1050	0.152	0.025	
	1340	0.157	0.025	
	1520	0.251	0.1	
	1780	0.214	0.095	
	2060	0.162	0.080	

REFERENCES

1. J.C. Urciuoli, "Limits of Adhesion in Rolling Contact," SM Thesis, Department of Mechanical Engineering, MIT (Feb. 1970).

2. R. Prabhu, "Experimental Investigation of Creep in Rolling Contact," SM Thesis, Department of Mechanical Engineering, MIT (March 1970).

3. Professor I. Paul, Private Communication. Also, "Results of Last Year's Work," Technical Memorandum to Mr. Kenneth Lawson, OHS&T (Sept. 1970).

4. F.T. Barwell and R.G. Woolacott, "The N.E.L. Contribution to Adhesion Studies," in *Adhesion*, Proceedings of a Convention arranged by the Railway Engineering Group, The Institution of Mechanical Engineers (November 1963).

5. J. Halling and M.A. Al-Quishtaini, "An Experimental Study of Some of the Factors Affecting the Contact Conditions and the Slip Between a Rolling Ball and its Tract," *Proc. Instn. Mech. Engrs.* 182, Part I, No. 37, p. 757 (1967-1968).

6. E. Rabinowicz, *Friction and Wear of Materials*, John Wiley and Sons Inc., p. 62 (1965).

7. J. Halling and B.G. Brothers, "The Effect of Surface Finish on the Creep and Wear of a Rolling Ball Subjected to Normal and Tangential Surface Tractions," *Wear* 9, p. 199 (1966).

7. RECOMMENDATIONS FOR FUTURE RESEARCH

Research on the frictional behavior of wheel-rail interfaces should concentrate, to our way of thinking, on providing answers to the following questions.

1. Under what conditions are the creep coefficients given by the smooth surface theory of rolling contact no longer valid? How do the creep coefficients change under these conditions?
2. What factors influence the rolling friction coefficient, and how?

Partial answers to these questions are presented in the preceding chapters, and summarized in the abstract. However, a great deal more information is needed. Various problems that need to be investigated are described in some detail in the text of each chapter; here we present a brief list of those that we believe are most worthy of further study.

7.1 Junction Deformation

1. Effects of sliding speed (or contact time) on shear strength of
   a) lubricated junctions
   b) weak, dry junctions
   c) strong, "welded" junctions.
2. Effects of high surface vibration levels at high loads on the behavior of junctions. Various junction models, representative of different surface characteristics, need to be studied.

7.2 Rolling Friction

1. Effects of (a) slip velocity (beyond the creep regime) and (b) rolling velocity on the rolling friction coefficient of
   a) lubricated surfaces
   b) dry surfaces
   c) clean, adhesive surfaces, such as are obtained with a plasma torch.
2. Effects of surface vibration level on the rolling friction coefficient. A possible experiment is to obtain the rolling friction coefficient
   a) at the same rolling speed, with two different vibration spectra
   b) at two different rolling speeds, with the same vibration spectrum.

   In both these experiments, it is important to differentiate between surface vibration and vibration at points distant from the contact region.
3. The importance of contact resonance and suspension resonances in determining frictional behavior.
4. Effects of surface roughness on the rolling friction coefficient for
   a) lubricated surfaces
   b) clean, adhesive surfaces.

7.3 Creep

1. Creep of clean, adhesive surfaces, such as are obtained with a plasma torch. The behavior of such surfaces could be simulated by using a highly adhesive metal such as Indium (which also oxidizes to a negligible extent) with appropriately scaled loads.

2. The effects of high surface vibration levels on the creep coefficients, as distinct from the effects of dynamic loading.

3. The effects of contact resonance on the creep coefficients. As a preliminary study, the importance of the contact resonance phenomenon with broad-band random excitation, such as is obtained when a wheel rolls on irregular track, should be investigated.

4. The creep of lubricated surfaces. The goal should be to obtain a correlation between the creep coefficients and the following variables:

   a) normal load
   b) rolling velocity
   c) oil viscosity (including temperature and pressure coefficients)
   d) degree of running-in
   e) oil film thickness
   f) surface roughness.

GPO 901-675