A STUDY OF LUBRICATED CONTACTS BY ELECTRICAL METHODS

by

John Alfred Leather

A thesis submitted for the
Degree of Doctor of Philosophy
in the University of London
and also for the
Diploma of Imperial College
May 1977

Department of Mechanical Engineering
Imperial College of Science and Technology
London S.W.7

ABSTRACT

A survey is made of the main developments in the study, by electrical methods, of lubricated contacts with particular emphasis on the asperity contact effect. Preliminary experiments with a hydrostatic disc machine demonstrate some of the difficulties of interpreting electrical resistance measurements.

A mathematical model is evolved relating the mean duration of electrical contacts to the statistics of surface topography and shows reasonable agreement with published data.

A method of studying the duration and separation of electrical resistance changes on a statistical scale is applied to a simplified line contact. Over the range examined no dominant frequencies are apparent. Subsequently a basis for the interpretation of the electrical resistance signal is proposed, assuming that contacts occur at random. This theory is applied to a wide range of published results. In many cases a derived variable shows a simple power law dependence on contact load, speed or oil film thickness. The reinterpretation of results obtained with different electrical circuitry yields essentially parallel curves.

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Professor A. Cameron, for the opportunity to undertake this work, for his interest and for his careful reading of the manuscript.

I am greatly indebted to Professor J.C. Anderson for many stimulating discussions, for reading parts of the manuscript and for placing the considerable facilities of the Electrical Materials Section at my disposal.

It is a pleasure to thank Dr. P.B. Macpherson for his encouragement and his advice on many aspects of this work.

Financial support was provided by Westlands Helicopters Ltd. and Ministry of Defence Procurement Executive and is gratefully acknowledged.

I thank my contempories in both the Lubrication Laboratory and the Electrical Materials Section for many interesting discussions and acts of kindness, particularly Mr. R.T. Smythe for preparing a sputtering target, Dr. R.C. Leveson for instruction in sputtering, Mr. E.J. Weidmann and Mr. P.J. Cranfield for help with and advice on photography, Mr. N. James and Mr. A. Holmes for help with computing, Mr. D.P.H. Smith for checking the mathematics, Mr. C.A. Jones for help with drawings and Mr. R. Dobson and Mr. A. Whymark for much practical guidance and for the manufacture of the sputtering masks.

Finally, my most sincere thanks to my wife, Heather, for moral support and encouragement and for typing the script and preparing some of the figures.

CONTENTS

LIST OF FIGURES

NOMENCLATURE

Infrequently used terms are defined in the text. Common usage has resulted in the duplication of some symbols and these have been included under chapter headings where appropriate. Subscripts 1 and 2 normally refer to the two surfaces in contact.

General.

W	applied load
A_o	nominal area of contact
A_r	real area of contact
ν	Poisson's ratio
E	modulus of elasticity
E'	reduced modulus of elasticity
b	Hertzian half width
η	oil viscosity
U,u	mean surface speed inside contact zone
γ	contact time fraction
T'	limit no-contact time fraction
t	sampling time
C_m	number of observable electrical contacts occurring in time t
C'_m	$\frac{C_m}{t}$ i.e. count rate
C	number of independent electrical contacts occurring in time t
C'	$\frac{C}{t}$
C^*	contact capacitance
m	average number of simultaneous electrical contact points

κ	separation between onset of electrical contacts
τ	duration of independent electrical contact
τ'	a particular value of τ
τ_m	mean value of τ
h_o	separation between ideally smooth lubricated surfaces
$\bar{h}$	separation between mean planes of height distributions
h	$\bar{h}$ and h_o where no distinction is drawn between these two values
h'	separation between mean planes of distributions of asperity summits
σ	combined root mean square roughness for two surfaces in contact (with subscript this refers to one surface only)
σ^*	combined standard deviation of peak height distributions of two contacting surfaces (with subscript this refers to one surface only)
σ_c	centre-line average roughness for profile
σ_m	standard deviation of roughness filtered at contact width
β_1, β_2	radii of curvature of asperity peaks for contacting surfaces (assumed constant)
β	$\frac{1}{2}(\beta_1+\beta_2)$
β^*	autocorrelation length

Chapter 3.

R_1, R_2	resistances in potential divider circuit, R_2 in parallel with lubricated contact

d	R_2/R_1
r	contact resistance in low resistance state
r'	contact resistance in high resistance state
R,S	series resistor
V_o	potential applied to divider circuit
V	potential difference across contact
V'	maximum value of V
V*	discriminating voltage
S'	value of series resistor in equivalent circuit

Chapter 4.

L	length of profile
y(x)	equation of profile
λ	factor relating smooth and rough surface pressure distributions
w	overlap or compression of two opposing asperities

Chapter 7.

I	intensity of incident beam
R, R_1, R_2, R'	coefficients of reflection
μ_1, μ_2	refractive indices

Chapter 8.

L	length of tapered roller
R	mean radius of roller

Chapter 9.

λ'	oil film thickness divided by the sum of the c.l.a. roughness values for the two surfaces in contact

R_c	contact resistance
r	Hertzian radius in point contact
R	"average resistance" of dynamic contact
v	average potential difference across contact

Chapter 1. General introduction.

Electrical measurements in the field of lubrication have a somewhat chequered history. Whilst appearing to detect small changes in experimental conditions these types of measurement have been notoriously difficult to interpret. It is a sad fact that no serious examination of the electrical parameters and processes has been undertaken by any of the researchers interested in applying these measurements. Equally well, the dynamic lubricated contact is hardly the ideal experimental arrangement from the electrical engineer's point of view. Important characteristics, such as the dependence of current on temperature or applied potential, cannot be measured directly and, at best, must be obtained by inference. This is further complicated by the fact that some of these parameters may change significantly during the course of an experiment.

The aim of this programme is to take a fresh look at the use of electrical measurements in studying lubricated contacts, from all relevant disciplines, with a view to a more direct interpretation of electrical behaviour. The work is concentrated on the partial elastohydrodynamic regime and on the "asperity contact" phenomenum.

In view of the many faceted nature of this work it was felt necessary to give a short guide to its layout and content.

The idea of a pure rolling contact is perhaps analagous to that of a completely smooth surface in that both are ideal. There is probably always some sliding present in a rolling contact, particularly when distortion of the surfaces is evident. For instance, if two separately driven one inch discs differ in diameter by 0.001 of an inch then there is .1 per cent sliding present. Thus, although the experimental work is nominally under rolling conditions, the effects of sliding are also considered. Theoretical work indicates that further information, concerning surface topography, may be obtainable when a known amount of sliding is present.

A representative survey of electrical measurements applied to tribology is made in chapter two. The lack of a theoretical basis for interpreting asperity contact effects limits the use of such observations.

The electrical circuitry employed by various workers is discussed in chapter three and the equivalence of the two types of circuits used is demonstrated. The series resistor is only necessary when current through the contact is monitored. The possible conduction mechanisms are discussed and the weight of evidence points to an electronic conduction process. This is important as it gives credence to the correlation of electrical contacts with surface features.

Some important developments in surface topographical analysis and modelling are described in chapter four. A model is then developed to examine the

duration of electrical contacts when sliding is the dominant behaviour. Good agreement is obtained with what little experimental data are available and the possible applications of this approach are considered.

In chapter five a series of experiments on a new design of disc machine is described. The results are generally in agreement with those found in the literature, and a chemical effect was detected.

A method of observing contacts on a statistical basis is evolved in chapter six and the electrical circuitry built to perform this function is also described. The problem of reproducing and maintaining experimental conditions led the author to attempt a new experiment with a simplified contact. This involved much work with surface coatings deposited by the relatively recent technique of radio frequency sputtering. This work is detailed in chapter seven. Chapter eight contains the results of the experiments performed with the above equipment. These results suggested a new theoretical approach to analysing the electrical resistance signal. This idea is developed in chapter nine. When applied to published results, derived parameters show a simple power law dependence on contact variables. The practical and theoretical significance of these findings is discussed in chapter ten, and areas for future research are identified.

A consistent system of notation has been utilised throughout this work. The contact time fraction is denoted by γ, so that the "per cent metallic contact" is 100γ.

The rate of contact interruptions is written as C_m' or $\frac{C_m}{t}$, representing C_m distinguishable contact events occurring in a time interval t. The quantity "m" that appears in chapter nine is the expected number of simultaneous micro-contact points.

Chapter 2. A review of electrical measurements in lubrication.

2.1 Introduction.

When the oil film is sufficiently thin the electrical resistance of a lubricated mechanical contact, in motion, appears to fluctuate rapidly between a high and a low value, with few intermediate values occurring. The frequency of this switching effect is commonly taken as a guide to the state of lubrication between components under examination, usually by displaying the electrical output on an oscilloscope.

Attempts to quantify this electrical information are complicated since it is a function of the actual circuits employed as well as the physical conditions inside the contact. One problem is that any electronic measuring equipment must have a wide frequency response in order not to obscure the very fast resistance fluctuations. An intrinsic limitation is caused by the finite capacitance of the contact. These and other electrical effects are discussed in chapter three.

Several electrical techniques have been applied to lubrication studies and some of the more important methods are described in this chapter. Theoretical developments, assuming electrical contacts correspond to actual contact between the two surfaces, have not been entirely successful.

Next, the "voltage discharge" method has declined in interest following the difficulty of obtaining general

calibration constants.

Contact capacitance measurement, although having certain difficulties of operation, has produced useful results.

An outline of the development and present state of knowledge of these electrical measuring techniques is given next.

2.2 Early work.

Needs (1) designed viscometers to investigate thin films and noticed that there was a measurable resistance between the dropping disc and the fixed disc at a separation of 2.5μm. and a low resistance at 1.5μm.

Beeck, Givens and Smith (2) used resistance measurements to indicate the transition between hydrodynamic and boundary regimes in a four ball machine which was brought up to speed and then allowed to run down under its own friction. From this a critical velocity was found and compared with that corresponding to a friction maximum.

Courtney-Pratt and Tudor (3) took measurements between the cylinder liner and piston ring in an internal combustion engine. They found low resistance regions close to Top and Bottom Dead Centres, where the piston speed was low, and high resistances corresponding to higher piston speeds. They were also able to make qualitative judgements on the effects of engine speed and lubricant viscosity. They noted a negative resistance region as the current through the contact was increased. Tudor (4)

observed the running-in of a journal bearing, and noted that the incidence of electrical contacts decreased with time.

In a letter to The Engineer, Lewicki (5) noted that the simple interpretation of resistance measurements based on bulk oil conductivity led to discrepancies in film thickness of three orders of magnitude or more.

2.3 The "voltage discharge" technique.

Brix (6) reported that when a sufficiently high current was applied across a thrust bearing, a potential drop appeared across the bearing, which was a constant amount greater than the voltage drop in the rest of the circuit. Moreover, this potential drop ("film volts") seemed to be dependent only on film thickness.

The same effect was observed by Cameron (7) in line contact and the term "discharge voltage" was identified with Brix' film volts. Considerable work was done by Cameron et al. (8,9,10 and 11) in making accurate calibration measurements of this effect, and in applying it to the line contact and point contact. Similar calibration constants were found by Bathgate and Yates (12) who applied the method to meshing gears.

An independent investigation was carried out by Dyson (13 and 14) who found that the discharge voltage was very dependent on the slide-roll ratio, and had a non-linear dependence on film thickness. An explanation of these differences is suggested by Twisleton-Wykeham-Fiennes (15)

and Fiennes and Anderson (16), who considered the electrical properties of the system, in particular the effect of cleaning the lubricating oil.

This technique is no longer in common use, although Brix (17) reports that some Russian workers are using it with reasonable accuracy.

2.4 Average resistance.

Another attempt to quantify electrical results was made by Furey (18). He used a simple averaging circuit in parallel with a point contact rig undergoing pure sliding. The average resistance measured is, strictly speaking, an average of the time during which the contact was in a low resistance state. The term "percent metallic contact" was also used by Furey. This is also misleading, since what is measured is the fraction of time during which one or many electrical short circuits exist somewhere in the contact. Notwithstanding, the values obtained seemed to be reproducible.

The effect of increasing applied voltage on metallic contact was found to be negligible until sparking occurred and a constant low resistance observed, (Fig. 2.1).

The effect of varying load, speed and viscosity are shown in Fig. 2.2.

Furey also examined the effect of rough surfaces on metallic contact (19). These results were very interesting as they showed an increase in contact with roughness, up to a point, and then a marked decrease. These

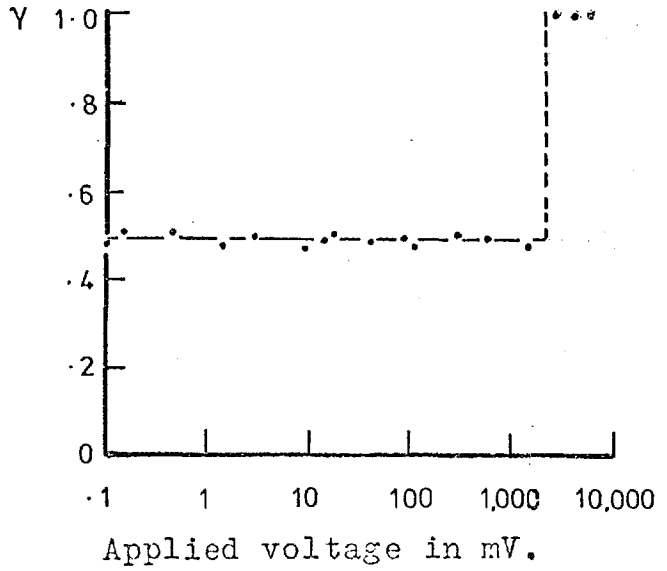
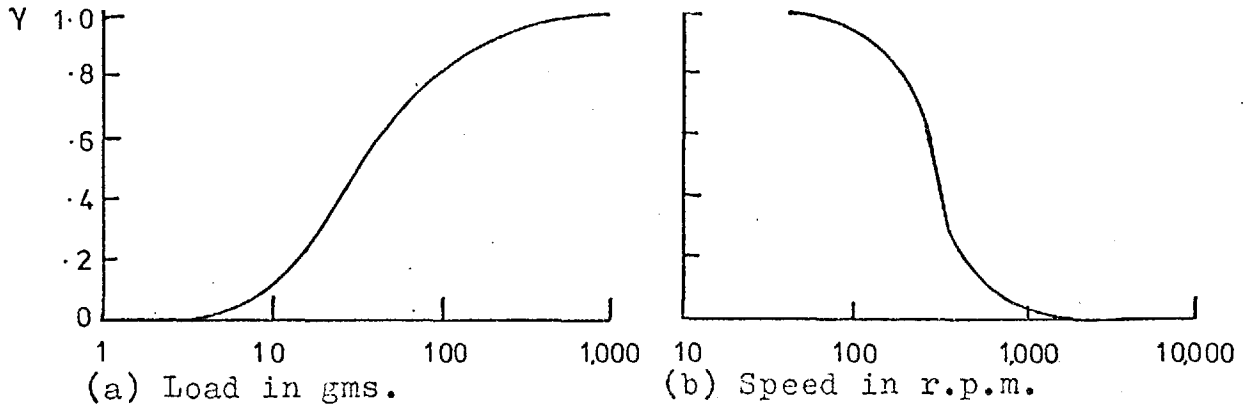


Fig. 2.1 The effect of applied voltage on γ.



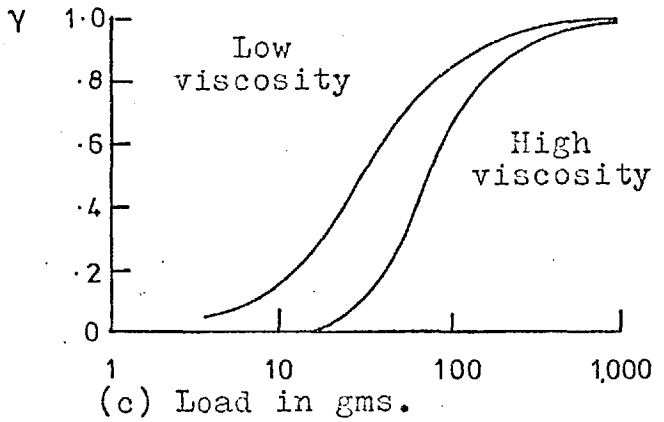


Fig. 2.2 Variation of γ with load, speed and viscosity.

results are shown in Fig. 2.3 and it can be seen that there is a maximum in the region of .25-.5μm. (10-20μins.) c.l.a.

Furey and Appeldoorn (20) also examined the effect of lubricant viscosity on metallic contact. An interesting observation in this paper was that a plot of metallic contact against load on Gaussian probability paper gave a very good straight line, although no explanation for this was proposed.

Christensen (21) looked at metallic contact and friction during the running-in process. The two discs on a disc machine were run for a period of several hours at a fixed load and the load was increased in steps. Metallic contact rose sharply as the load was increased then gradually dropped to a near constant value after about twenty minutes. This led him to suggest that the running-in time for discs could be considerably reduced as most changes in the surfaces of the discs occurred in a fairly short time. He then went on to fatigue tests and found that, in some cases, the discs scuffed as soon as 100% metallic contact occurred, and in others the discs ran on quite happily at 100% contact without giving any indication of failure. Clearly this is a possible short-coming of this method. Once 100% contact is reached no further information is available to indicate changes in the discs by this method.

An interesting observation was made by Millns (22). A particular industrial oil (Shell Tellus 27) built up a partially insulating film when a potential difference was applied between test discs. The potential difference was removed for a length of time and on re-application a

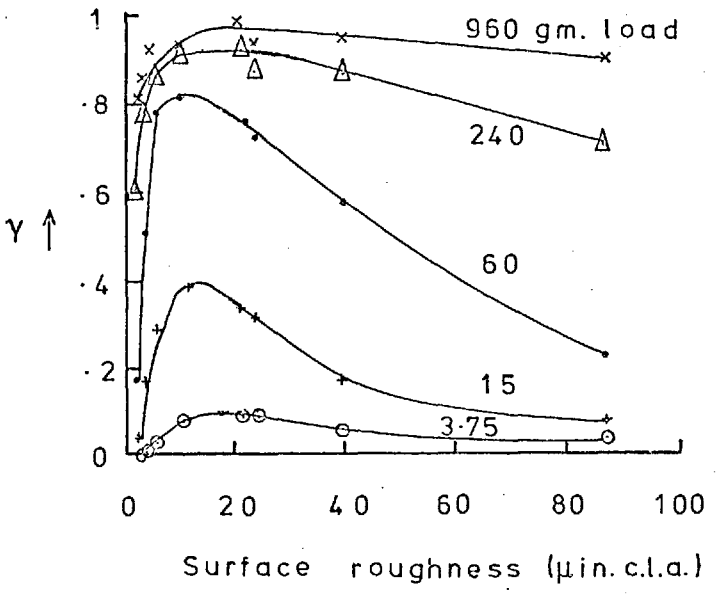


Fig. 2.3 The effect of surface roughness on γ.

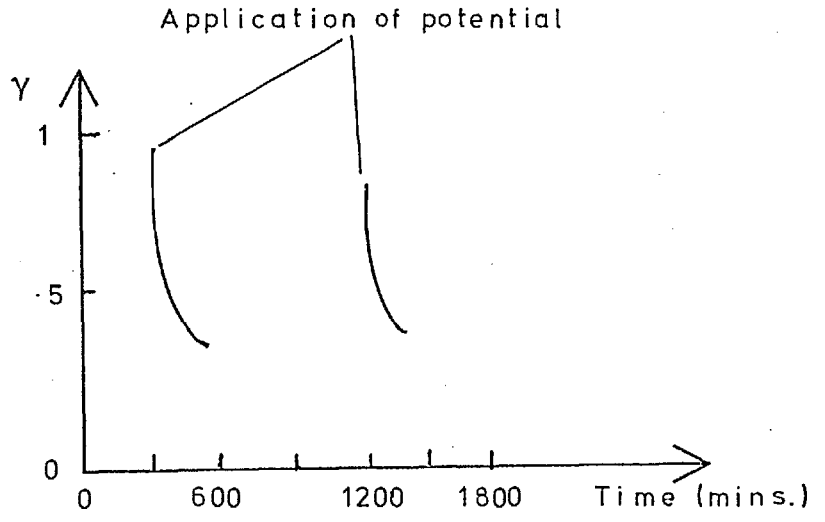


Fig. 2.4 An effect observed by Millns (22).

higher degree of contact was observed. This behaviour is shown in Fig. 2.4. The time taken to regain a steady value was about two hours.

These experiments were under steady conditions and were performed on run-in discs. During running-in normal behaviour was observed.

Millns attributes these results to a potential assisted adsorbtion process. He suggests that polar fatty acid molecules migrate to the anodic surface, where adsorbtion is enhanced by the potential present.

Clearly this effect is a hazard to estimating film thickness from electrical measurements. It does suggest, however, a practical way of reducing contact in a bearing. More work in this area would be of interest.

The average resistance across the races of a roller bearing is described by Garnell and Higginson (23) and Garnell (24). The electrical situation is very complex. To register a contact there must be simultaneous electrical contact between the outer race and a roller, and the inner race and a roller. Many parallel current paths are available through the estimated five rollers (from the twelve in the bearing) in simultaneous contact. No doubt roller-cage contacts also occur. Despite these complications there appeared to be reliable correlation between average resistance and film thickness for a range of loads. This is illustrated in Fig. 2.5 which is taken from (24).

A load correction was made to the resistance

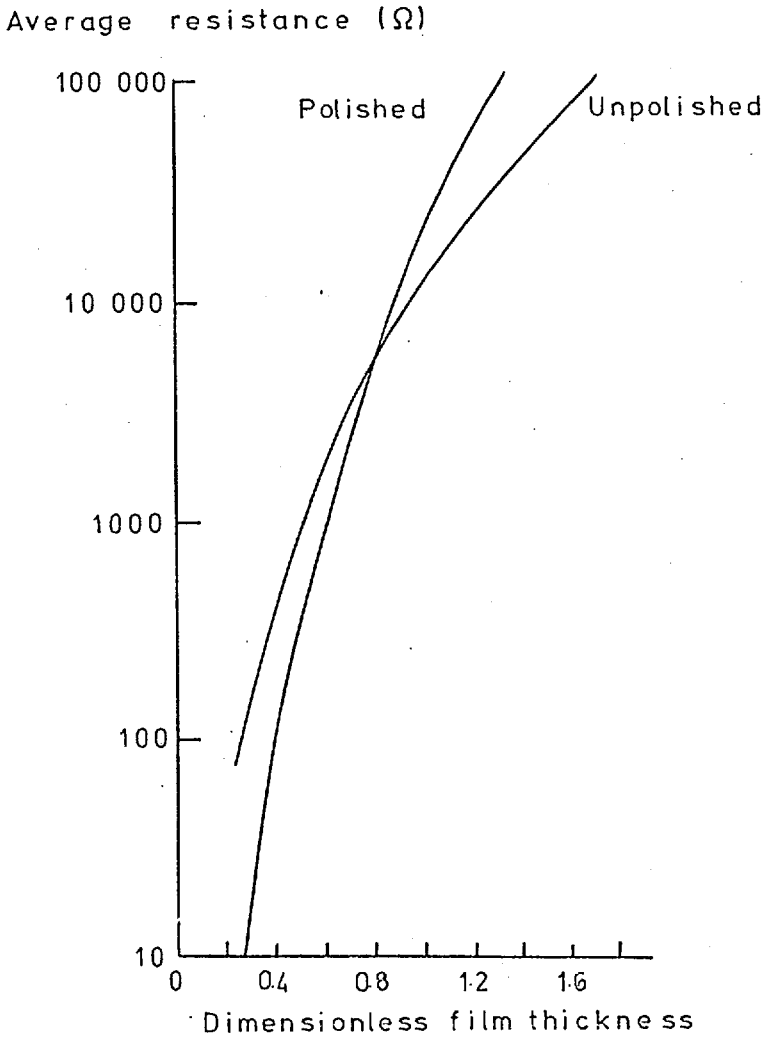


Fig. 2.5 Variation of average resistance with dimensionless film thickness for roller bearing.

values in obtaining the best fitting curves, the measured values being multiplied by $\sqrt{\frac{\text{LOAD in lb}}{9000}}$. The electrical circuit used was not published but the applied potential was 50mV and the large resistance in a potential divider had a value of 200kΩ.

It is interesting to note that when the applied voltage was increased to 200mV, substantially lower resistances were recorded. This was ascribed to arcing in the contact. There was also evidence to show that the resistance values were reduced when the 200k resistor was reduced to 20kΩ.

The technique was similarly applied to a tapered roller thrust bearing by Higginson and Leaver (25) and Leaver et al. (26). A region of unstable running was observed at all loads. This is shown as the hatched area in Fig. 2.6 below on a speed-torque plot.

The variation of metallic contact with speed is shown in Fig. 2.7 below. The broken line purports to show the region of unsteady running. However, reference to Fig. 2.6 shows that this line should run approximately perpendicular to its present course. The conclusion in this paper (26) that "the boundary of steady running is difficult to define ... and occurs over a band of fractional no-contact of 0.3 ... to 0.6 ..." should have the values of 0.3 and 0.6 amended to about 0.4 and 0.5 respectively.

The curves for the two lowest loads were not published in Fig. 2.7 but it is mentioned that they would

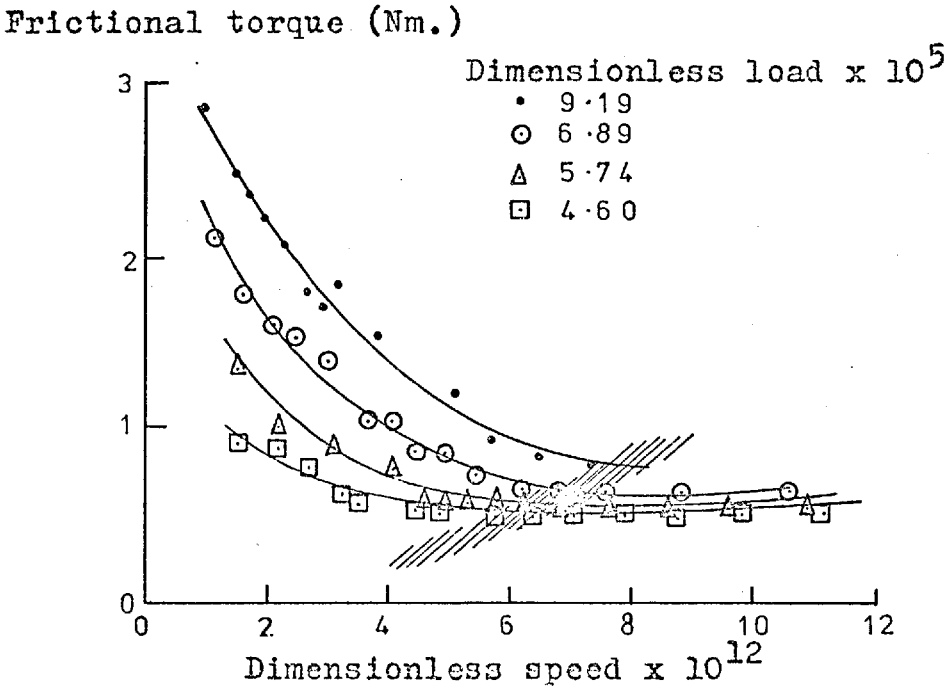


Fig. 2.6 Variation of friction with speed for thrust bearing.

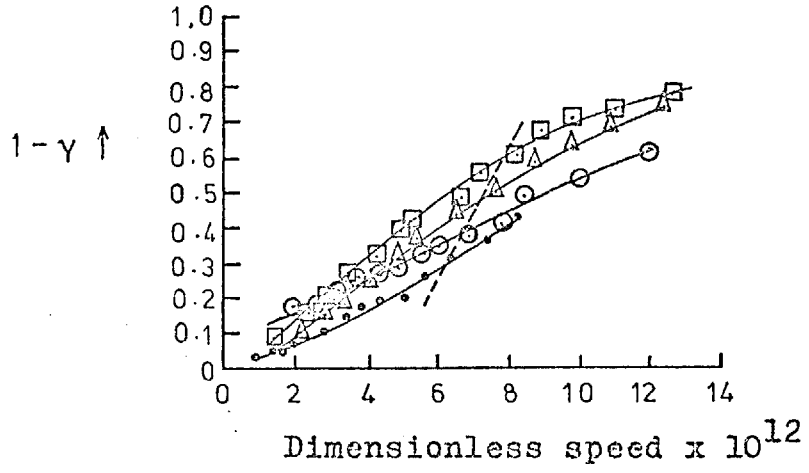


Fig. 2.7 Variation of no-contact time with speed.

not lie in order above the other curves, which are fairly well behaved with load variation. No explanation is given for this but it seems that at the lowest loads there is an unexpected rise in metallic contact. A similar effect occurred in the disc machine experiments detailed in chapter five.

Similar experiments were performed by Poon and Haines (27) on a point contact disc machine. Typical behaviour of film thickness with contact time is illustrated in Fig. 2.8 below. This data was then plotted as a Weibull distribution (28) to give Fig. 2.9 below. This shows the data to fall on two straight lines of different gradients. This is then matched with a discontinuity in the traction coefficient.

The Weibull approach involves rewriting a cumulative distribution F(x). F(x) represents the probability that the random variable is less than x. Weibull rewrote this function as

$$F(x) = 1 - \exp[\phi(x)].$$

The advantage of this is apparent where the failure of one element is sufficient to cause breakdown of a whole system. If, for instance, F(x) is the probability that a metal link should break at or below a particular load, then the probability that a chain of n links should not break is $(1-F)^n$ or $\exp(n\phi)$.

2.5 Capacitance.

Brix (6) reported that capacitance was used to

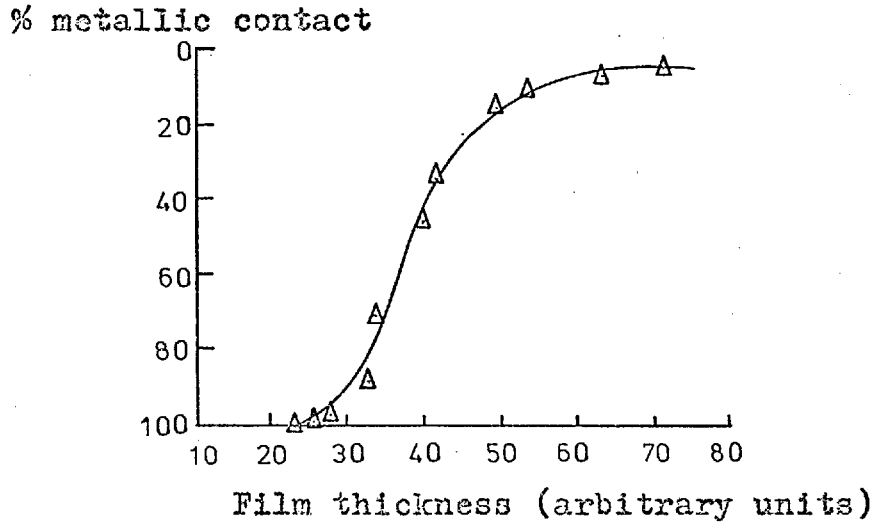


Fig. 2.8 Variation of metallic contact with film thickness.

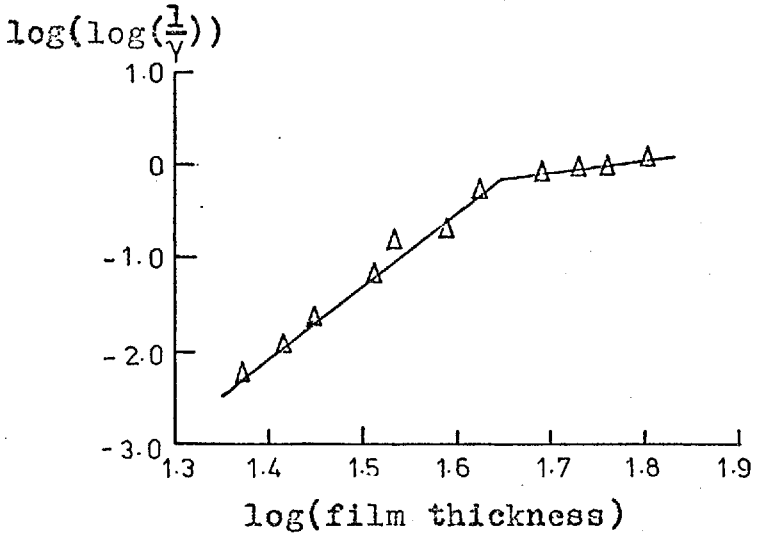


Fig. 2.9 Data of Fig. 2.8 plotted as Weibull distribution.

estimate the oil film thickness in a bearing as long ago as 1926. Since then the technique has been refined. The idea was reintroduced by Lewicki (5) in 1955 and then taken up by Crook (29), who made measurements between test discs and very lightly loaded pads outside the contact region in order to remove the errors due to estimating the actual contact geometry. Archard and Kirk (30) applied the method to point contact on a crossed rollers rig, and found it necessary to introduce a frequency selective amplifier to reduce the effect of spurious contacts.

Crook (31 and 32) was able to make measurements of the film profile by exchanging one of the steel discs for a glass one with a chromium electrode on its surface, and showed the constriction in the exit region by this method. Orcutt (33) used a similar technique and generally confirmed Crook's results. Dyson, Naylor and Wilson (34) performed a large number of tests on a disc machine and obtained excellent agreement with the theory of Dowson and Higginson (35). They found it advantageous to use a higher frequency (19kHz) than other workers, as well as tunable filters.

Hamilton and Moore (36) used temperature and pressure sensing materials in conjunction with a capacitance pad and were able to display these variables simultaneously.

Vichard and Godet (37) applied the technique to cams and tappets and Astridge and Longfield (38) pointed out the deficiencies of the assumptions generally made about the geometry of the contact, when applied to high conformity

situations.

An independent assessment of capacitance measurements was performed by Snidle (39) who compared capacitance with interferometric measurements directly on a crossed cylinders rig. Wymer (40) extended these measurements to line contact. Snidle found quite good agreement between the two methods, except at high speeds. Wymer used frequencies of up to 50kHz and also interposed a dielectric layer between the two surfaces which reduced the effect of surface irregularities and constriction regions on the capacitance readings. He found excellent agreement between the optical and capacitance methods.

When measuring very thin films, the roughness of the surface makes an important contribution towards the capacitance compared with the equivalent smooth surface situation. A modified calibration curve was produced by Napel and Bosma (41) to take some account of this effect. The correction was calculated using a superimposed sinusoidal surface roughness and experimental work confirmed that better agreement to theory was obtained.

2.6 Asperity counting and contact resistance.

The average resistance technique described in section 2.4 was extended by Tallian et al. (42) who examined the electrical contact interruptions in greater detail. In this paper two new terms were introduced, no-contact time fraction, which is self explanatory, and limit no-contact time fraction (T'). This quantity is the no-contact time fraction if the contact were to have no capacitance. A model was set up and a correction function derived which

can be applied to the measured values to obtain the ideal values. In the notation of this thesis the correction is given by

$$1-\gamma = \frac{T'^2}{C_m' RC^* + T'} \qquad 2.1$$

where C_m' is the measured rate of contact interruptions and RC^* is the time constant of the electrical circuit-contact combination.

The effect of changing several electrical factors was explored in this work. These are briefly mentioned here and are discussed in more detail in chapter three. Fig. 2.10 shows the effect on no-contact time of changing the series resistor. The broken curve is a T' curve presumably derived from one of the solid curves. It is very significant that the 48kΩ curve is above the broken one. Inspection of equation 2.1 shows that in all cases T' should be greater than 1-γ. It must be concluded from this that the change of series resistor has also changed the number of contacts detected. As the 48kΩ curve was closest to the T' curve, this value of series resistor was used in subsequent experiments.

Fig. 2.11 shows the difference caused by changing the applied potential. The 100mV potential appears to detect more contacts than the lower applied voltage. This is in direct contrast to Furey's findings (18) displayed in Fig. 2.1, and these differeces have yet to be resolved.

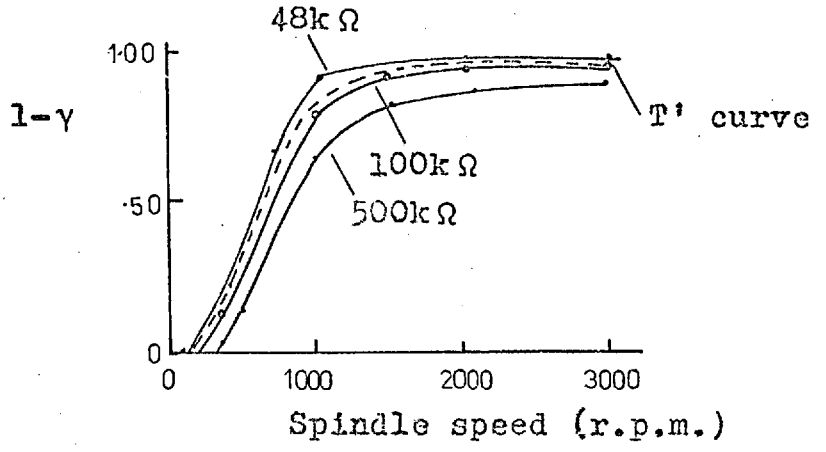


Fig. 2.10 The effect of series resistor on 1-γ.

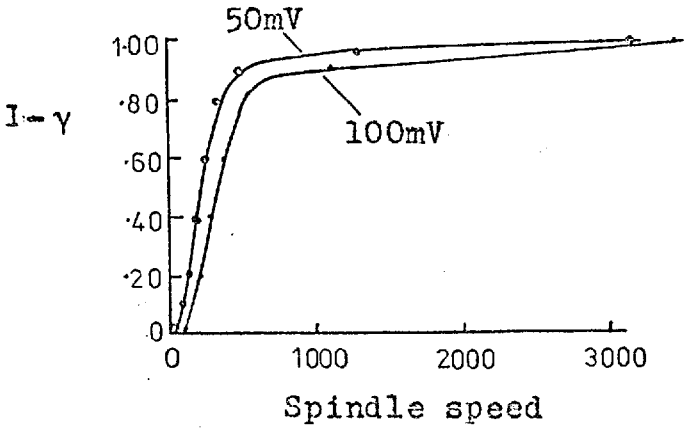


Fig. 2.11 The effect of applied potential on 1-γ.

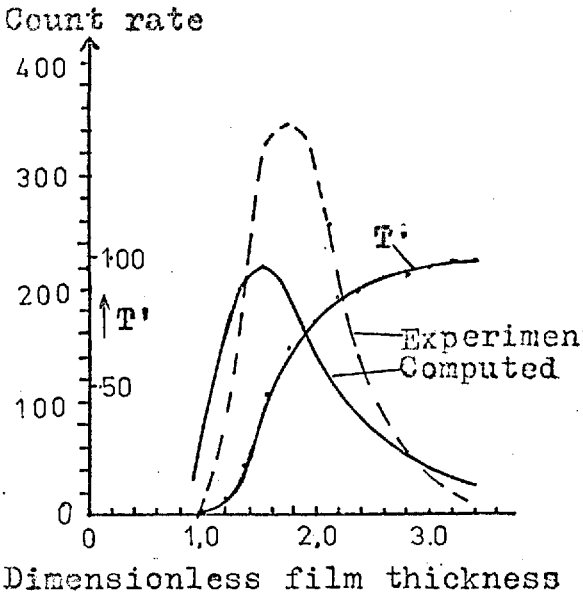


Fig. 2.12 Count rate and T' for 50lb. load.

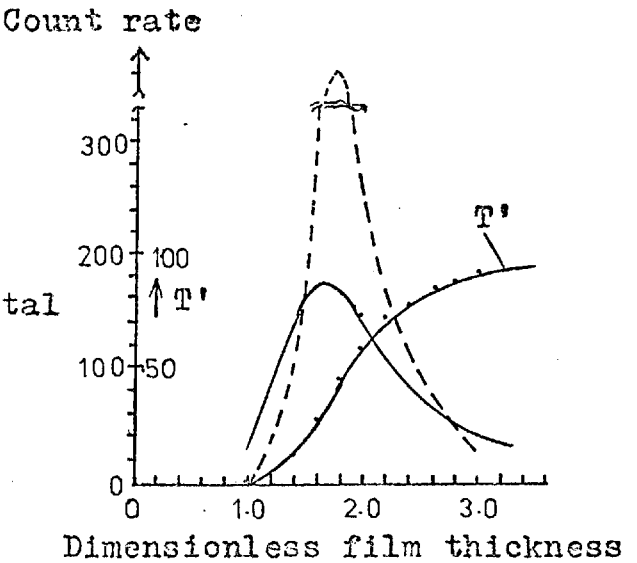


Fig. 2.13 Count rate and T' for 100lb. load.

The change in count rate with film thickness for two loads is demonstrated in Figs. 2.12 and 2.13. The broken curve is the experimental count rate, and the solid curve is that predicted from the theory described shortly. The third curve is the T' curve fitted to the experimental values marked. The experimental and theoretical curves show the same general behaviour although the fit is not exact. The theoretical approach is not described fully in (42) but the main lines of development are as follows. Both surfaces are assumed to have Gaussian distributions of heights. These may be combined to give a single Gaussian distribution with a variance equal to the sum of the variances of the two surface height distributions. Tallian then goes on to consider "level crossings". These are the number of either way crossings of a datum through the profile, parallel to the mean line. Using results from Bendat (43), it is shown that the number of level crossings is normally distributed with height. Finally the intervals between level crossings are examined and a distribution is derived for these. If a profile is examined through a "window" of length d corresponding to the appropriate Hertzian width, then for a contact interruption to occur the interval between level crossings must be greater than d. Thus, with a knowledge of the distributions obtained above, the expected values of no-contact time and count rate may be derived.

This process was extended for a two-dimensional isotropic surface and a circular window corresponding to the Hertzian area. Only the results of this analysis are

presented in (42).

It is noted here that the main physical assumptions of the model are that the duration of an isolated contact is equal to the Hertzian transit time and that the asperities bear no load. The relevant statistical quantities were derived from a profilometric study of the test balls.

Subsequently certain errors in this theory were pointed out by Sidik (44) who developed a general theory for predicting electrical contact behaviour. This unfortunately depended on two-dimensional statistical variables, the behaviour of which is not generally known. It was then necessary to make various assumptions to simplify the mathematics.

In both models the duration of a contact was taken to be the Hertzian transit time. Sidik also found it necessary to ignore multiple contacts. Clearly the good agreement found by Tallian is somewhat suspect if the assumed value for the duration of a contact is wrong.

Tallian, McCool and Sibley (45) extended the technique to examine lubricants and showed that the ratio of film thickness to composite root-mean-square surface roughness was an important factor in determining wear and fatigue life.

In an effort to expand the time scale and examine electrical contacts in more detail, Tallian and McCool (46) used very viscous oils in a slow moving point contact. The

test balls were observed in some instances to slide and spin as well as roll and no correction was applied for this.

Some contacts were noted that were considerably longer than the Hertzian transit time although most were much shorter. These long contacts occurred too frequently to be explained by multiple contacts. Generally, contacts were arranged to be unlikely events so that double or multiple contacts should have been extremely rare.

In the simple model of a parallel circular Hertzian zone, asperity contacts can be infinitesimally short by occurring very close to the side of the zone. A statistical analysis of the experimental results would have been necessary to determine whether they were in keeping with the theory and this was not the authors' intention.

They were able to make actual resistance measurements by comparing simultaneous voltage and current records. Generally resistances were much higher than would be expected from constriction resistances, and of the order of thousands of ohms. This value is much too low to be explained by bulk electrical properties of the oil. These resistances were reduced by about an order of magnitude when the series resistor was reduced by two orders of magnitude. The authors suggested that the most likely mechanism for conduction under their experimental conditions was electronic tunnelling.

A closer examination of individual contact occurrences reveals interesting behaviour close to the no-

contact level. Frequently the contact does not occur instantaneously but has a definite rise time both at the leading and lagging edges of the pulse. This is idealised in Fig. 2.14. It is possible, though unlikely, that the decay at the lagging edge is a capacitance effect. It is not possible to explain the "switch on" curve in this manner. It seems very probable that this curve actually represents the approach of two asperities. This switch on time is of the order of tenths of seconds.

The good correlation between the occurrence of electrical contact and the predictions based on surface profiles led Tallian (47) to believe that the small scale surface features are essentially unchanged as they pass through an elastohydrodynamic contact.

Christensen (48) examined the duration and frequency of electrical contacts on a statistical scale. Fig. 2.15 is a histogram of contact duration in pure rolling. The Hertzian transit time for the conditions in this Fig. is 2 milliseconds. The same experiment was performed with a disc speed ratio of 1.5:1 and these contact durations are very much shorter, as shown in Fig. 2.16.

Christensen also studied the variation of rate of contact interruptions C_m' with load, speed and viscosity. He derived the following empirical formula for this variation:

$$C_m' = k_1 W^{1.28} (\eta u)^{-3.5} \qquad 2.2$$

where k_1 is some unknown constant, and W, η and u have their usual meanings.

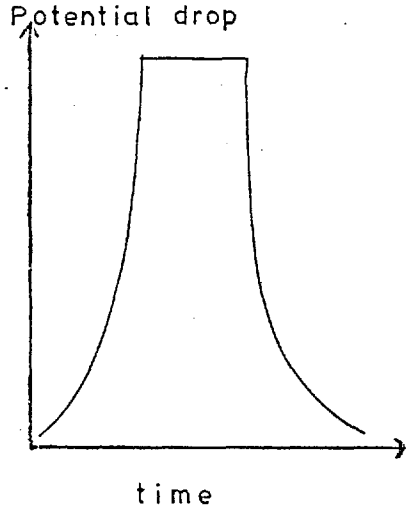


Fig. 2.14 Idealisation of contact.

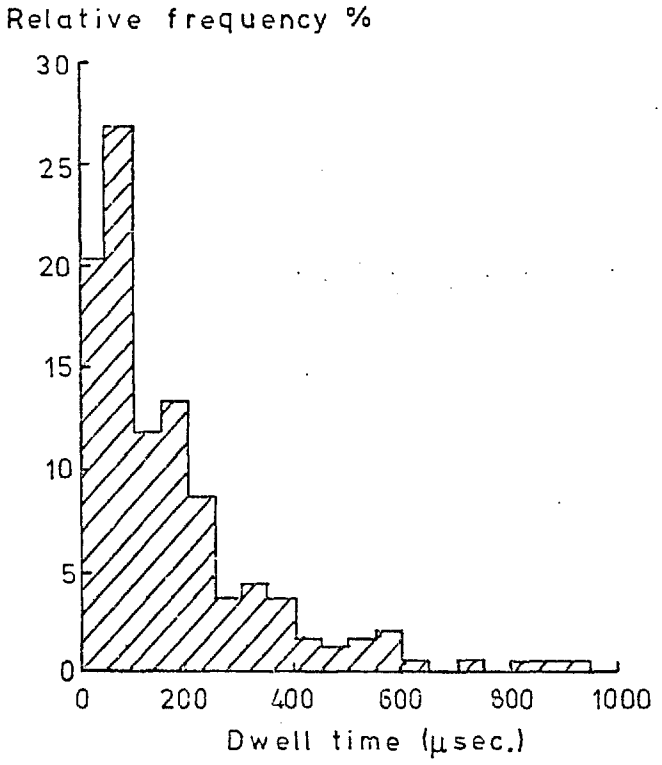


Fig. 2.15 Distribution of dwell times for rolling contact.

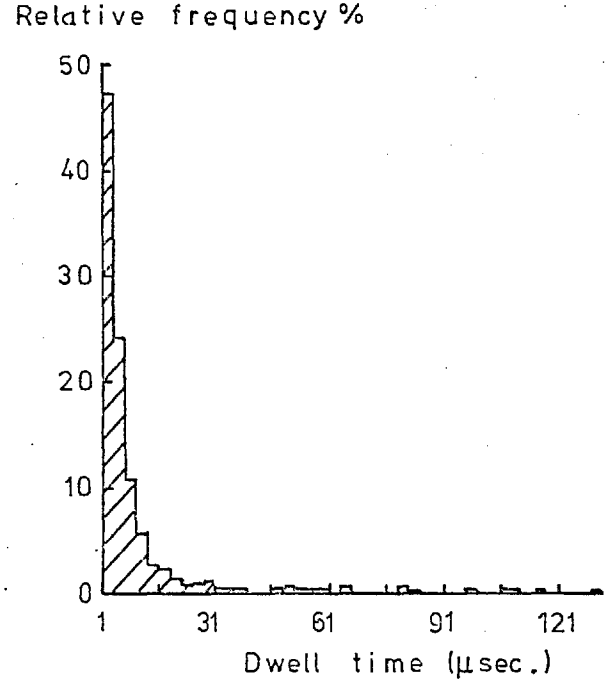


Fig. 2.16 Distribution of dwell times for sliding contact.

A similar equation is proposed for the variation with film thickness, thus:

$$C'_m = \frac{k_2}{h^5} \qquad 2.3$$

However the range of γ for which these experiments were performed was restricted and Christensen warns that an extrapolation of these equations outside the region investigated is highly questionable. Clearly these equations would not be expected to hold for larger values of γ, when multiple contacts predominate.

Fig. 2.17 shows γ, the contact duration τ and the count rate C'_m plotted against viscosity. It immediately strikes the eye that C'_m and τ fall into two distinct groups on either side of a viscosity of about 25cPoise. It would be interesting to know if there was any difference in experimental procedure to account for this. The product of C'_m and τ was plotted against γ and showed a general correlation with γ about twice as large as $\tau C'_m$ for rolling and about ten times greater for rolling with sliding. These results must question the accuracy with which the quantities C'_m and τ were measured as it seems difficult to explain the big discrepancy by statistical variation alone.

All the preceding workers have split the electrical signal fairly arbitrarily into two regimes; contact and no-contact. Czichos (49) has done some work in which three regimes are defined;

metallic contact (resistances less than 10Ω)

quasi-metallic contact (resistances less than

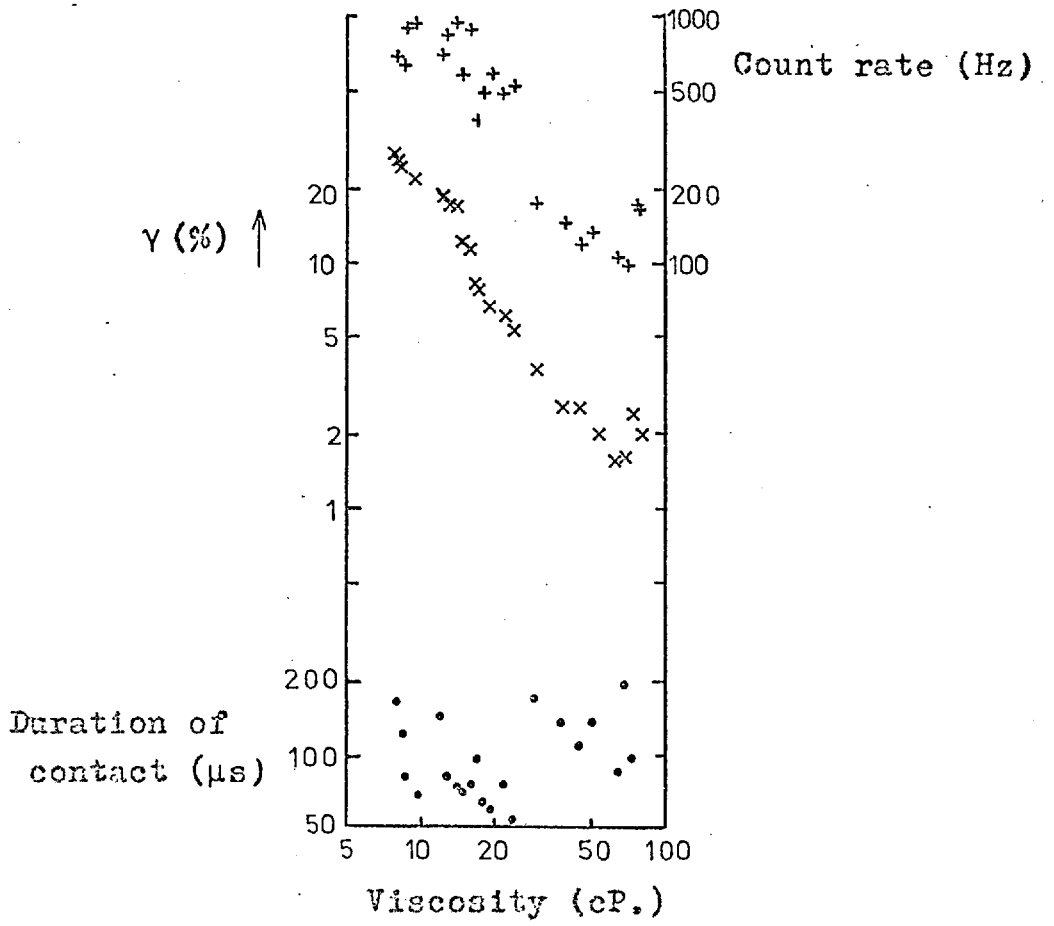


Fig. 2.17 Variation of count rate, contact duration and γ with viscosity.

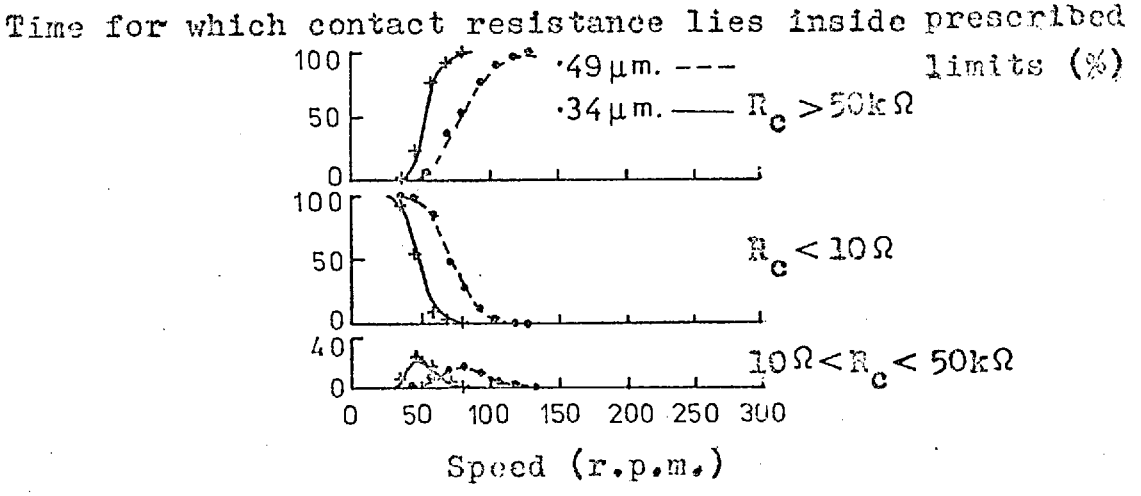


Fig. 2.18 Variation of distributed contact resistances with speed and cylinder roughness.

50kΩ but greater than 10Ω)

no-contact (resistances greater than 50kΩ).

The percentage of time spent in each of these three regions for various speeds is shown below (Fig. 2.18). The solid line represents a roughness of .49μm.(~20 μinches) and the broken line of about .34μm.(~13 μinches). The experimental arrangement was a cylinder rotating on a fixed plate. Similar results were obtained using a plate sliding on three pegs. These plots are compared with the speed-traction characteristics. The minimum value of traction occurs shortly after the peak of quasi-metallic contact time.

In a recent paper Czichos et al. (50) have extended the technique to examine the "time rate" and the fluctuation rate (C'_m) within every decade of contact resistance from 0.1Ω to 100kΩ. This was applied to four lubricants, pure hexadecane, hexadecane plus one of dibenzyldisulphide (DBDS), zinc dioctyldithiophosphate (ZDDP) and oleic acid. The results for 2 minutes, 20 minutes and 1 hour running time are shown in Figs. 2.19, 2.20 and 2.21. It must be assumed in Fig. 2.19 that one of the points on the DBDS graph has been misplotted as the sum of the ordinates is 110%! The effects illustrated were apparently quite reproducible. The only conclusion drawn from this work is that despite having very similar wear and friction characteristics "...the interfacial physico-chemical mechanisms of the ZDDP system and the oleic acid system are entirely different." The oleic acid results indicate a steady increase in the electrical

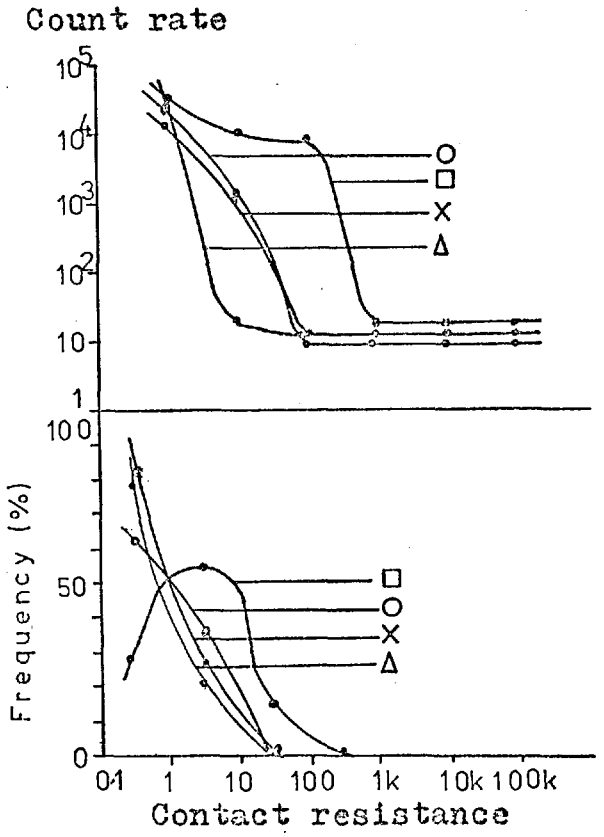


Fig. 2.19 Electrical data at 2 min. test duration.

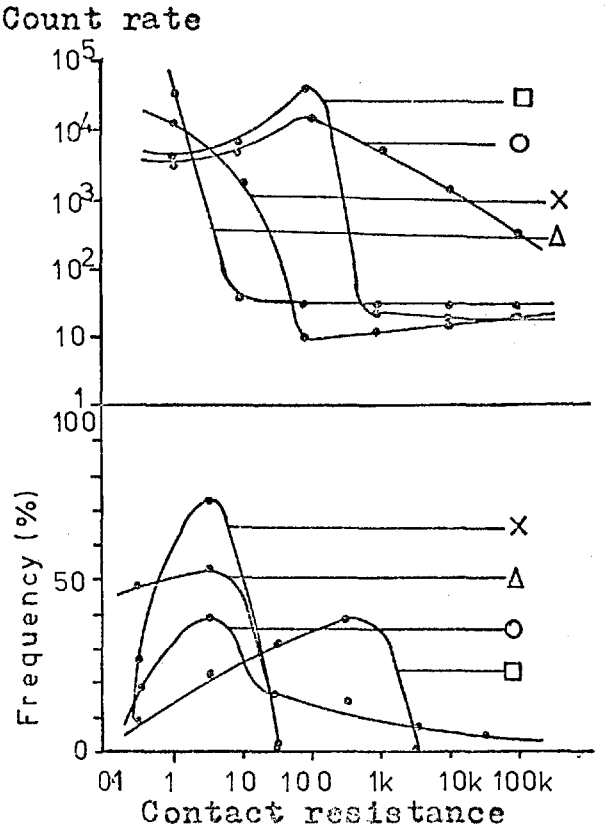
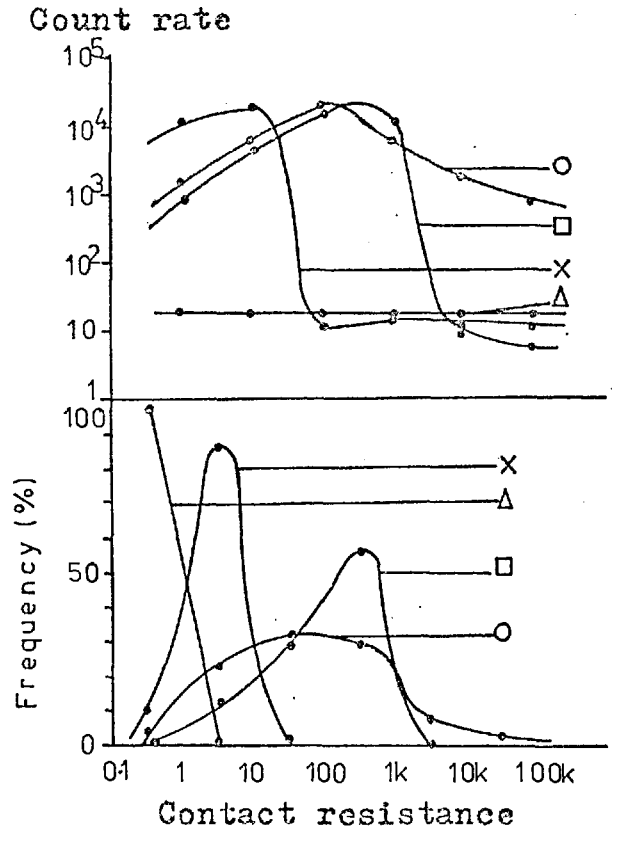


Fig. 2.20 Electrical data at 20 min. test duration.



Key to symbols

- ○ Pure Hexadecane
- □ Hex. + oleic acid
- Δ Hex. + ZDDP
- × Hex. + DEDS

Fig. 2.21 Electrical data at 60 min. test duration.

resistance of contacts with time. Possibly this indicates soap formation. In view of the Millns effect (22), it would be interesting to know whether or not the applied potential was removed after the 2 and 20 minute readings.

The electrical circuitry had a high frequency response and used modern digital techniques. Count rates in excess of 5×10^5 per second were occasionally observed. In an attempt to locate any spurious electrical activity, such as local electromotive forces or semiconducting junctions, the applied potential was removed and no voltage above 10μV was detectable. Reversal of the polarity of the applied potential or reduction from 20mV to 2mV did not alter any of Czichos' findings. Rise times for the contacts were typically of the order of tenths of micro-seconds.

Electrical resistance monitoring of piston and cylinder lubrication was recently reintroduced by Dowd and Barwell (51). Measurement of no-contact time fraction indicated that surfaces of moderate roughness, about 0.13μm. (5.3μin.) c.l.a. gave rise to thicker oil films than very smooth surfaces. This was ascribed to micro-asperity hydrodynamic lift. The results also showed the existence of a range of contact resistance values under any given conditions.

2.7 Summary.

The most useful electrical methods of observing the lubricated contact appear to be capacitance measurement and electrical contact analysis where transitions between

high and low resistance states are regarded as events. To some extent these two techniques are complementary. Capacitance values are most reliable when the oil film is thick compared with the surface roughness, and the exit constric tion may be disregarded. Under these conditions asperity contact is infrequent, so that the information obtainable from studying such contacts is limited. In more severe operating conditions capacitance measurement is more difficult and less reliable, but it will no doubt remain a useful laboratory tool within the confines of roughness and conformity.

Electrical contact analysis is useful in a relative way, but the interpretation of these data is, as yet, only qualitative. The theoretical approach of Tallian gives a reasonable fit with experimental results but does not lead directly to a practical method, for instance, of monitoring bearing performance. So much information has to be supplied to make predictions of electrical contact behaviour that divergence of experimental results from theory does not pin-point particular physical changes in the system. The same general comments must be made of the model developed by Johnson, Greenwood and Poon (52), which is discussed at length in chapter four.

The findings of Christensen that contact durations are, on average, very much shorter than the Hertzian transit time throws doubt on a main assumption of the theoretical models. This may be due to the fact that extremely small amounts of sliding can dominate

electrical contact behaviour, and an alternative premise is developed in chapter four.

Recent work shows that the electrical resistance of a dynamic lubricated contact is not necessarily a two state system of high and low resistance. The fact that γ increases when more current is passed through the contact indicates the existence of a range of values for contact resistance. The observation of essentially high and low resistances may be a subjective error. This is discussed further in chapter three.

It is possible that further information is inherent in the electrical contact signal. The distribution of contact durations has been examined by Christensen, but the intervals between the occurrence of contacts may be worthy of further study. Presumably these intervals must be related in some way to the topography of the contacting surfaces. If it is possible to isolate statistical parameters directly from the electrical contact signal, it would be interesting to study how such parameters vary with current through the contact.

Chapter 3. Electrical considerations.

3.1 Resistance measurement circuits.

One of the problems confronting the researcher is to ensure that any measuring systems do not upset the measured conditions. For instance, when a thermometer is used to measure the temperature of a liquid, it takes a small but finite amount of heat away from the system. Similarly, in the system now considered, the effects of the measuring circuit must be minimised.

The circuit used by Furey (18) is shown in Fig. 3.1. The switch S can be set in any one of four positions, allowing different values of current to flow through the contact, while keeping the applied potential difference V' constant between the cylinder and the ball.

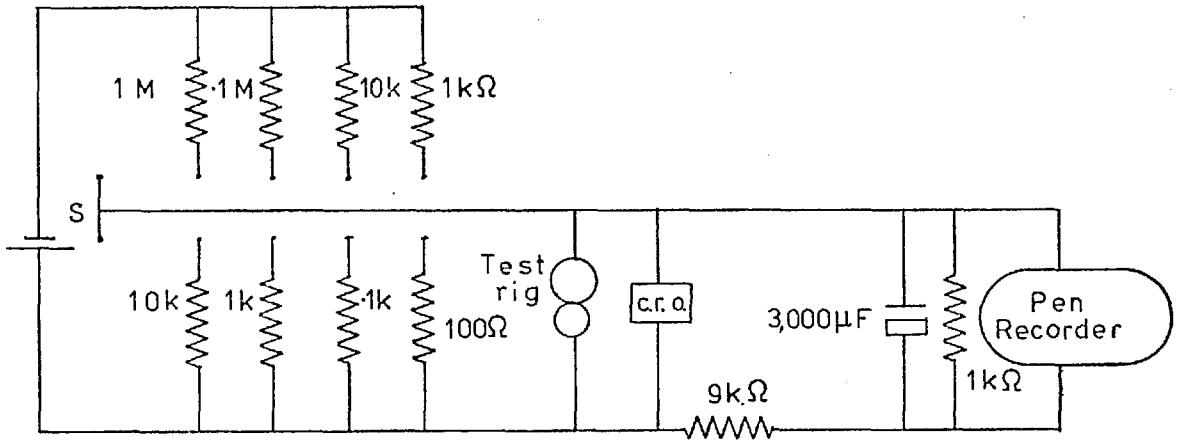


Fig. 3.1 Electrical circuit used by Furey.

This system would not be suitable for asperity counting as the large capacitor would certainly affect the very fast switching process in the contact.

A generalised Furey circuit is shown in Fig. 3.2.

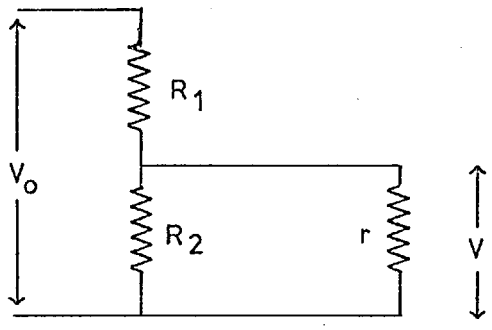


Fig. 3.2 Generalised Furey circuit.

A potential V_o is fed into the potential divider R_1 and R_2, and r, the resistance of the contact, is in parallel with R_2. The potential V across the contact is monitored.

In this circuit, $$V = \frac{rR_2}{r(R_1 + R_2) + R_1R_2}V_o$$

$$= \frac{rd}{r(1 + d) + R_2}V_o \quad \text{where } d = R_2/R_1.$$

Now $d \ll 1$ and $dV_o = V'$, the potential across the contact when $r = \infty$.

Then, $$V = \frac{r}{r + R_2}V' .$$

The variation of V with r is shown in Fig. 3.3. It will be seen from this graph that the value of R_2 determines the excursion from V' that a particular contact resistance r causes. The smaller the value of R_2, the smaller the value of r needed for a particular voltage excursion on an oscilloscope. Conversely, a fixed contact resistance r will cause a bigger potential drop with a smaller value of R_2.

It has been tacitly assumed this far that the

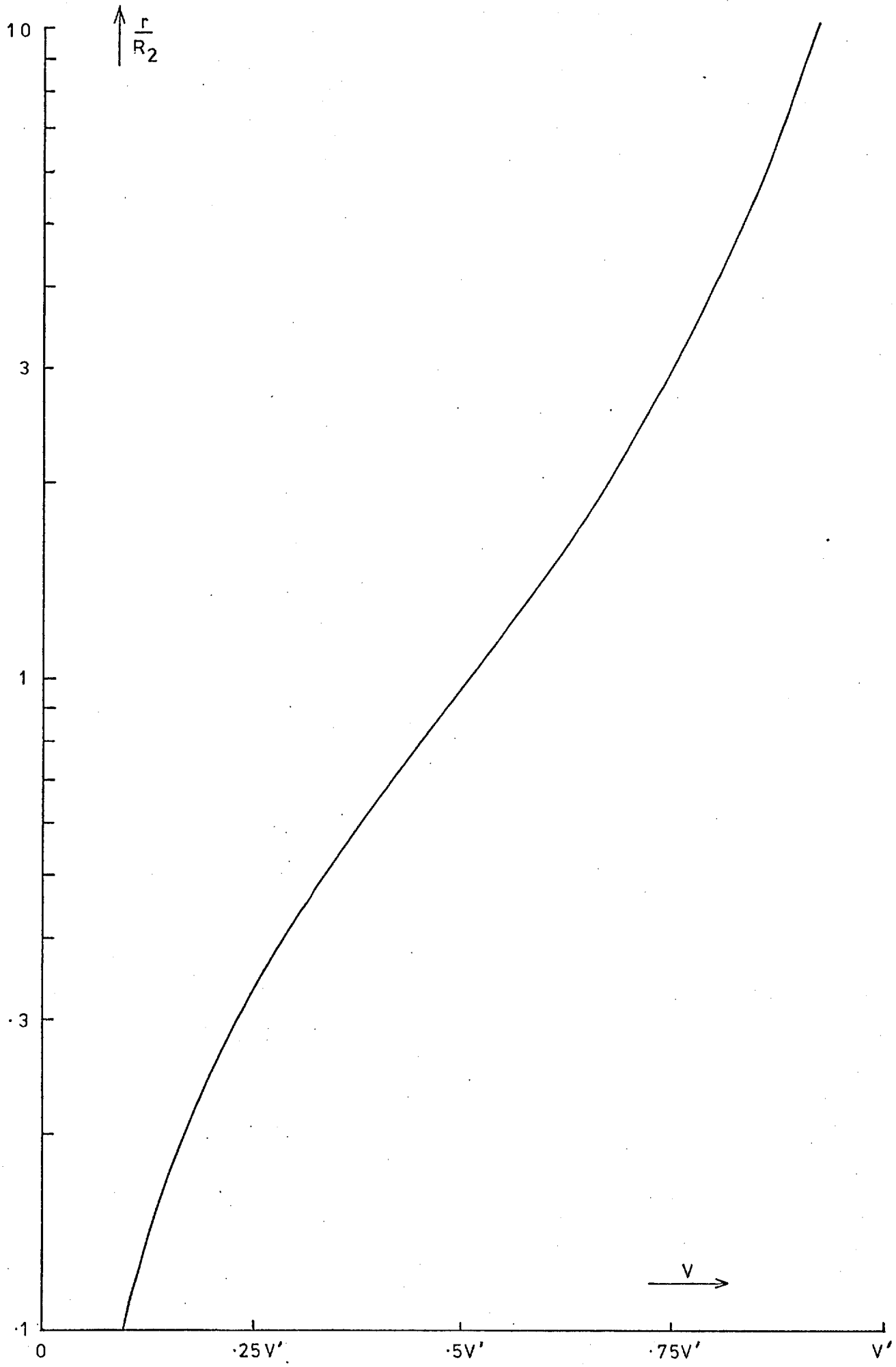


Fig. 3.3 Variation of measured potential with contact resistance.

mechanisms of conduction between the ball and cylinder are purely dependent on the applied potential. However, it may well be that these mechanisms are current dependent. In this case, the increased number of contacts observed on a more "sensitive" scale, obtained by increasing the current flow, may not be due solely to more accurate resistance measurements, but also partly to a different conduction mechanism becoming more viable.

In order to remove this possibility it is necessary to keep the applied potential and the maximum allowable current constant and small.

A series resistor between one side of the contact and the potential divider can be used to limit the current, although this can also introduce problems due to the capacitive effect discussed below.

Tallian used the circuit shown below. (Fig. 3.4)

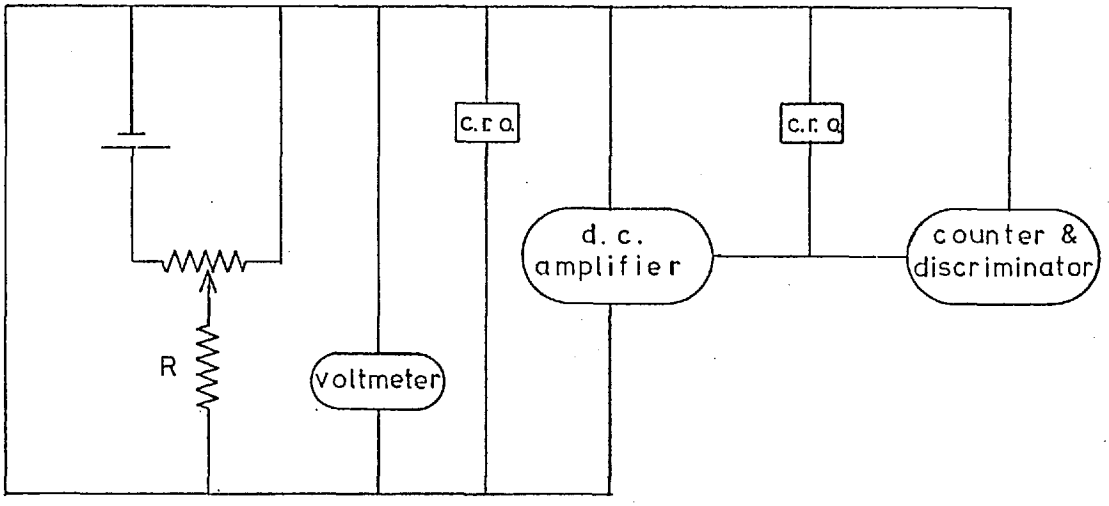


Fig. 3.4 Block diagram of Tallian's circuit.

This has the series resistor R and introduces the electronic counter and discriminator. This enables asperity counts to be recorded. The valve voltmeter, which has a long time constant, allows "per cent metallic contact" or "no-contact time fraction" to be measured directly.

If contacts are assumed to be ohmic in nature, the effect of changing the discriminator level can be reproduced without altering anything except the two resistors in the potential divider circuit. The ratio of these resistors must be kept the same but the actual values change. To illustrate this, consider a discriminator level of $\frac{3}{4}V'$ applied to the circuit of Fig. 3.2. Referring to Fig. 3.3 the discriminator will register a contact if $r < 3R_2$. If the discriminator level is reduced to $\frac{1}{2}V'$ a contact will register when $r < R_2$. The same effect will be observed without change to the original discriminator setting by replacing R_1 and R_2 with resistors one third as large.

If however these contacts are non ohmic there would be a discrepancy between changes in discriminator level and the equivalent resistance alterations. This does suggest a simple method for studying the variation of resistance with current on the timescale usually employed in test machines, although some correction owing to time constant changes may be necessary.

The simple electrical analogue in Fig. 3.5 may be used to illustrate the switching process that is observed in the asperity contact phenomenum.

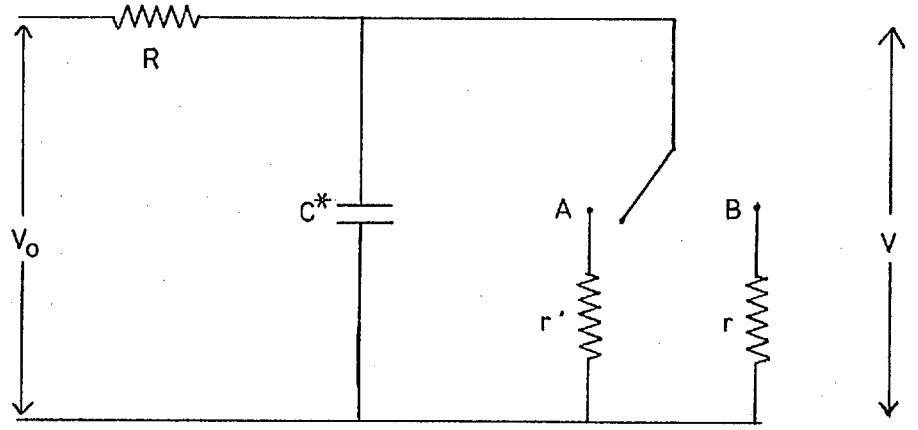


Fig. 3.5 Equivalent circuit of asperity "contact".

The series resistance R (ohms) limits the current to the contact, supplied by the potential difference V_o (volts). The capacitance of the contacting members is C^* (farads) and r and r' (ohms) represent the resistances of the two states of conduction, where $r' \gg R$ and $r \ll R$. The potential across the contact is V (volts).

In the high resistance state, (position A)

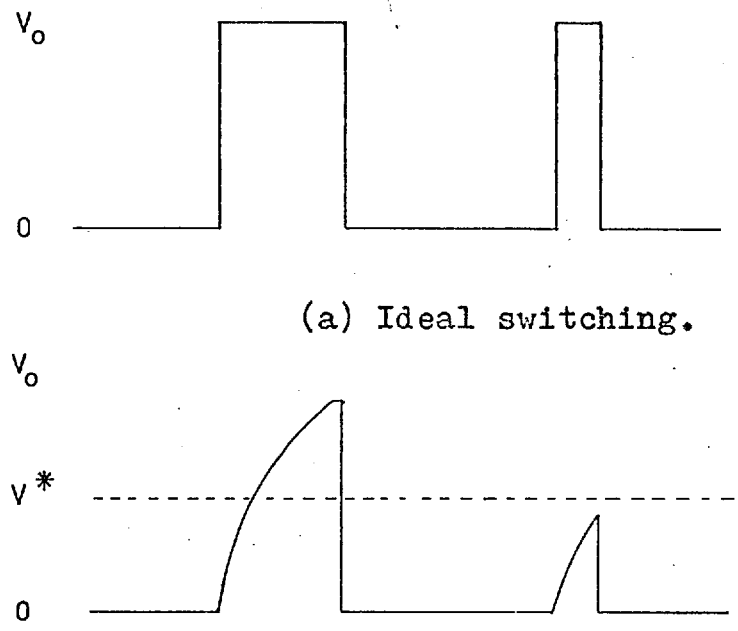
$$V = \frac{r'}{R + r'} V_o \simeq V_o \text{ since } R \ll r'$$

On switching to the low resistance state, (position B) V will fall to $\frac{r}{R + r} V_o$ which is small as $R \ll r$. In addition the capacitor will discharge most of its energy in a time of $5C^*r$ seconds. This energy is $\frac{1}{2}C^*V^2$ watts.

On reverting to the high resistance state, V will rise exponentially towards V_o and will be close to it after $5C^*R$ seconds.

This analogue illustrates two effects. Firstly,

if C*R is comparable to the contact dwell time, instantaneous switching will be masked by the finite rise time of the contact voltage. The typical waveforms in Fig. 3.6 illustrate this.



(a) Ideal switching.

(b) Signal modified by capacitance.

Fig. 3.6 The effect of capacitance on switching.

In (b) the second interruption would not be apparent if the discriminator level was set at V^*.

Secondly, since C*R is comparable to the dwell time, C*r is much shorter than this. Hence all the energy of C^* is dissipated through the contact in a very short time. This could lead to localised heating and surface damage, in conjunction with the mechanical stresses.

Tallian proposed a correction factor to take account of the capacitance effect. Based on an exponential distribution of contacts, he derived equation 2.1 (on page 39).

Clearly, if $C_m'RC^*$ is small enough compared with T', then V/V_o is very close to T'. The only controllable parameter in this equation is R. If R is reduced by a factor of k such that V/V_o is close to T', then it is necessary to reduce the applied voltage by the same factor, in order to maintain the same small current through the contact.

This also has the effect of reducing the energy stored in C^* by a factor of k^2. Since the time in which this energy is dissipated does not change, there must be a drop in local heating.

In practise, an upper limit is imposed on k by the increasing difficulty of distinguishing the signal from electrical noise and a compromise should be reached with k as large as is practicable.

The equivalent circuit of Fig. 3.5 could be modified to replace the capacitor C^* with two capacitors, one representing the capacitance between the test elements and the other being environmental capacitance. Thus C^* is the parallel combination of a fixed and a variable capacitor. Environmental capacitance is frequently the larger of the two. Poon and Haines (27) suggest in their paper that the capacitance of their system is 1pF and is thus negligible. In a lively discussion of this paper Archard points out that this value is unrealistic. The authors concurred with this judgement and subsequently suggested values of between 50pF and 100pF. It is possible to obtain an estimate of the time constant by examination of the rise time of the oscillogram in (27). This yields a value of about 2×10^{-4}

seconds and with a series resistor of 40kΩ, C* is revealed to be about 5000pF.

Digital techniques have been applied to the measurement of γ and C'_m by Dowson and Whomes (53) and later by Czichos (49 and 50). The high resolution of such systems makes them very suitable for these measurements but it must be recognised that a limitation is placed on the frequency response by the capacitance of the test machine and the equivalent resistance of the dividing network. The effect of a long time constant is to increase γ and reduce C'_m.

Following the publication of Tallian's work (42) it became common to add the series resistor S in the idealised circuit shown in Fig. 3.7. The only advantage

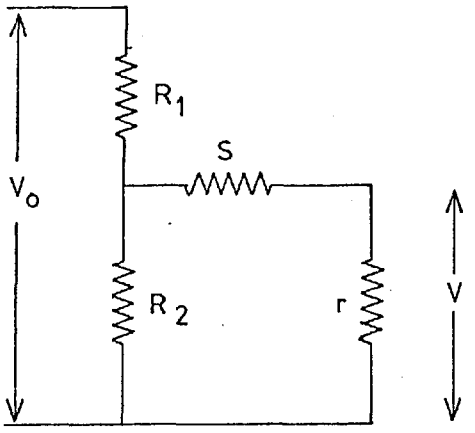


Fig. 3.7 Simple potential divider and series resistor.

of the series resistor is that it allows measurement of the current through the contact. The network in Fig. 3.7 may be replaced by a voltage source and series resistor as shown in Fig. 3.8, where $V' = \frac{R_2 V_o}{R_1 + R_2}$ and $S' = S + \frac{R_1 R_2}{R_1 + R_2}$.

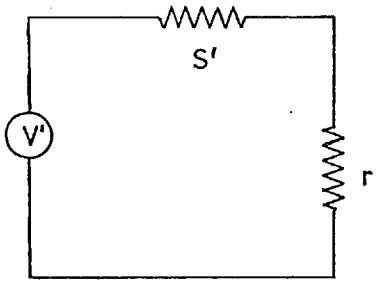


Fig. 3.8 Equivalent circuit of voltage source and resistor.

The idealised Furey circuit of Fig. 3.2 may also be represented by Fig. 3.8 with $V' = \frac{R_2 V}{R_1 + R_2}$ and $S' = \frac{R_1 R_2}{R_1 + R_2}$. It is easily shown that for any circuit of the type shown in Fig. 3.7 there is a unique equivalent circuit of the Furey type.

Attention must be drawn to the "compression" effect that is observed when using simple circuits to measure resistance. Referring to Fig. 3.3 it can be seen that about 86% of the volts scale is representative of two decades change in resistance. If the contact resistances are not essentially constant but evenly spread over, say, ten decades of resistance centred on the parallel resistance, R_2, a subjective interpretation of an oscilloscope trace might diminish the importance of the smaller percentage of resistance peaks occurring near the centre line. Subjectively the oil film resistance would be high or low with few intermediate values. This process is idealised in Fig. 3.9 which shows how ten contacts, each successive one having ten times the resistance of the previous contact, might appear on an oscilloscope screen. The parallel resistance is taken as 10kΩ and the contacts have resistances from 1Ω to 10^9Ω.

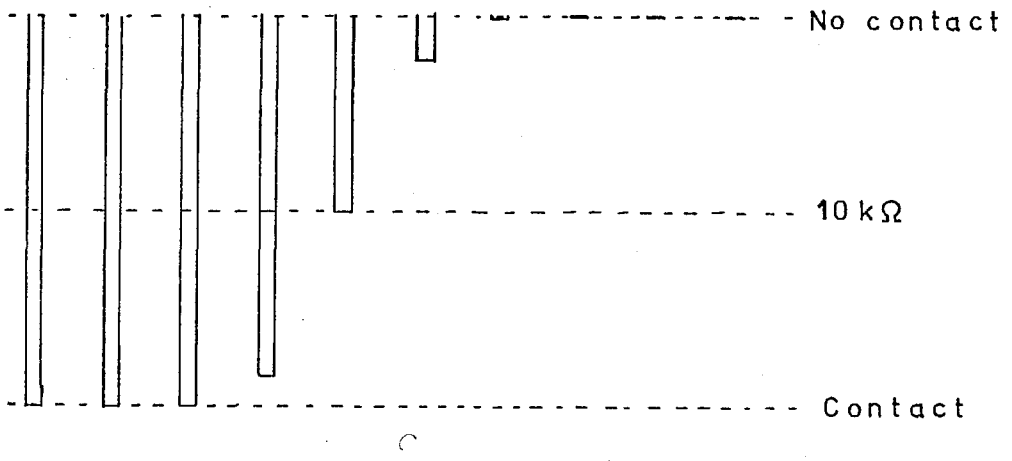


Fig. 3.9 Idealised oscillogram of wide range of contact resistance.

Fig. 3.9 might be subjectively interpreted as a two state system.

3.2 Electrical behaviour of thin films of lubricating oils.

The study of conduction through thin oil films would no doubt be at a considerably more advanced stage than at present if it were possible to set up a test cell with completely smooth surfaces and variable oil film thickness. The experiments described in the literature are far from this ideal case and electrical characteristics must be inferred to a large extent from incidental observations. The slow speed experiments of Tallian et al. (46) probably provide the most constructive approach to this problem and a similar experimental arrangement could be used to study electrical response in more detail. Several aspects of this work are of interest. The finite rise times observed, of the order of tenths of seconds, suggest that at the very slow speeds employed the approach of opposing asperities is reflected in the resistance variation. The actual resistance values observed during contact were typically thousands of

ohms. However when the series resistor was reduced by two orders of magnitude (i.e. the current increased by that amount) the contact resistance dropped by about one order of magnitude. This is a flimsy piece of evidence on which to base a theory. However there are no other data available, so a relationship between current and resistance of the form $I \propto \frac{1}{R^2}$ is tentatively suggested.

The results of Czichos shown in Figs. 2.20, 2.21 and 2.22, and those of Dowd and Barwell (51) generally support the proposition that contact resistance values vary widely although the spread of resistances is different for each system. No doubt the distribution of resistances depends on the electrical properties of both oil and additives as well as chemical factors. An interpretation of this is that the surfaces in nominal contact have many points of close approach and the actual separation varies from contact to contact. For a given system contact resistance is dependent in some way on separation, and probably also on effective conduction area. It is extremely dangerous to generalise in this field, but it is probably safe to state that electrical resistance is critically dependent on small changes of separation. This is most obvious in experiments with extremely smooth surfaces where a 0.025μm.(1μin.) change in oil film thickness causes a large change in count rate and contact time. The effect is masked in rough surface experiments as there is not such a high proportion of potential contact points in a small distance.

Outside the field of lubrication there has been

some interest in the electrical properties of organic thin films. These films can be deposited by evaporation or by the Langmuir process described by Nathoo (54). A few monolayers of, for example, stearic acid can behave in a similar fashion to inorganic insulators and exhibit electronic conduction (55 and 56). Although the controlled conditions required for this type of experiment are very different from those in a typical lubrication study, the application of conventional thin film theory may not be totally unrealistic.

Thin static films of n-heptanol were examined by Rennell and Anderson (57). A one μm. film of very pure fluid showed evidence of ionic conductivity. In order to explain the observed asperity contact resistance values by such a mechanism, calculation shows that the effective conduction area must be several square cms. for a film 100Å thick with an applied potential of 100mV. The actual area of lubricated surfaces separated by 100Å or less is many orders of magnitude less than that calculated by the above method.

Commonly neglected parameters in electrical resistance measurement as applied to tribology are humidity and oxygen effects, as discussed by Holm (58). It is indeed rare to find lubrication experiments performed in a sufficiently clean and dry atmosphere.

3.3 Mechanisms of conduction.

Conduction through thin oil films at higher powers than those employed in asperity contact experiments was studied by Fiennes and Anderson (16 and 59). They suggested that a charge injection mechanism could account for the

observed effects. At applied potentials of 100mV or less it is difficult to imagine that significant charge injection could occur. Generally an ionic mechanism appears unlikely for several reasons. First a mechanism for the generation of ionic species must be established. The asperity contact effect is observed with pure oils as well as with those containing additives. The effect is also observed with dry oil in a nitrogen atmosphere (50). So where do ionic carriers come from? Further, the mobility of large ionic species will be considerably reduced by the high viscosity of the oil in the high pressure region, and one would suspect that this might limit the speed of response to applied potential. Czichos, for example, observed rise times typically of 200ns. Also, polarisation of surfaces is not generally observed.

Thus it seems difficult to explain the observed effects by ionic conduction although it is within the bounds of possibility that an extremely mobile ion such as H^+ or O^+ is the main carrier.

High electrostatic potentials can be generated, for instance, in oil pumps. These are not thought to play an important part in practical contact resistance measurements because of the frequent low resistance states. This assertion can easily be tested by running an experiment without connecting the voltage supply to the test elements (e.g. (50)).

Very little is known about electronic effects in thin oil films. The work on Langmuir films mentioned in section 3.2 shows that very thin organic layers can exhibit

electronic conduction in the same way as inorganic insulators. However, these adsorbed organic films are highly ordered. This may not be the case in a dynamic situation and disorder introduces further complications such as electron trapping sites inside the oil film.

The oxide film on steel surfaces must not be neglected. It is typically ~40Å thick and composed of various oxides with some free metal. It is possible to prepare pure oxides of iron which are good insulators but this would not be typical of a steel surface (60). The presence of free metal causes oxide mixtures to behave in a quasi metallic manner and the effect of oxides on electron tunnelling barriers may be ignored.

Certainly some electron tunnelling must occur between the two surfaces prior to actual contact. The question to be decided is whether or not it is the dominant mechanism. A definitive answer to the conduction processes can only be revealed in future work. A possible experimental arrangement might consist of a lubricated sliding crossed cylinders arrangement where the curved surfaces are highly polished glass or curved mica sheets as shown in Fig. 3.10.

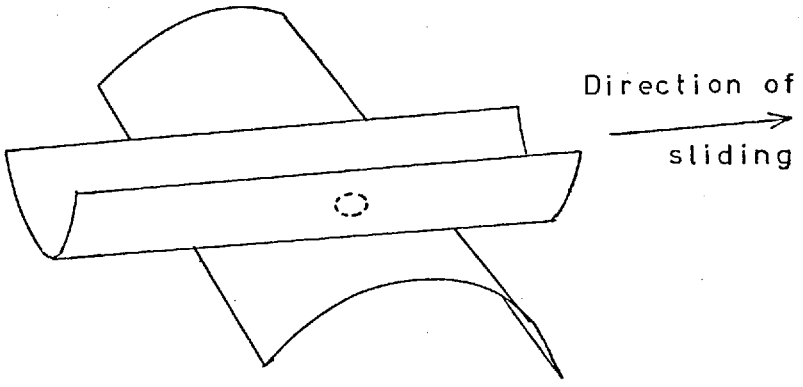


Fig. 3.10 Crossed cylinder type configuration.

The effect on conduction of different metallic coatings forming electrodes on the smooth surfaces might be particularly revealing. Ideally this sort of experiment must be performed in a dry neutral atmosphere.

3.4 Conclusions.

The frequency response of measuring circuits is discussed with relevance to the accurate measurement of contact time and count rate, and it is shown how the two common types of circuit used are equivalent. The observation that resistance values are either high or low could be a subjective error and the effect can be explained by a large range of resistance values.

Physical interpretation of nominal contacts must await better understanding of the electrical properties of oils, and presumably will vary from oil to oil and will also depend on additives and impurity content. The only property that all the systems have in common is that the observed resistance changes are fast and very sensitive to small changes in distance.

It seems that an electronic conduction mechanism such as electron tunnelling is the most likely explanation for the observed contact phenomena. This is encouraging as it suggests that it may be possible to link contact events with surface topography. There is an exponential variation of electron tunnelling current with separation of electrodes under normal circumstances (61, 62 and 63). This implies that it may be possible to establish a critical separation which corresponds to the distance between opposing asperities which is just sufficient to register a contact. The critical

distance will depend on the electrical properties of the oil and any adsorbed boundary films, but for any given system will be constant. It may be that the range of effective conduction areas owing to the multiplicity of asperity shapes and possible alignments reduces the value of the critical distance idea.

At this stage, the most constructive approach appears to be the examination on a statistical scale of electrical contacts regarded as events. It may or may not be possible to link statistical parameters to the variables of engineering interest. However, if the dominant conduction mechanism in this situation is electronic tunnelling, which is sensitive to small changes of separation of the order of 10Å or less, then the contact of dynamic lubricated rough surfaces can be studied on a scale that is not possible by any other method.

Chapter 4. Rough surfaces.

4.1 Introduction.

In the last ten or fifteen years there has been much interest in the mathematical modelling and in the characterisation of rough surfaces. These developments have accompanied technical innovations in the devices used for analysing surfaces, and the increased use of modern computers to handle the large amounts of data generated by the stylus profilometer.

The first part of this chapter presents a review of some of the most important ideas in the history of rough surface theory. It is not intended to be exhaustive, merely to show the changing approaches to the problem.

Subsequently a model is developed to examine the duration of sliding contacts and this shows good agreement with experimental evidence.

4.2 Specification of surface finish.

A homogeneous, isotropic, rough surface is one whose character does not change with position or orientation on the surface. Parameters required to specify such a surface must be invariant with respect to position or direction and are clearly statistical quantities. The most common functions used for this purpose are the root mean square (σ) and the centre-line average (σ_c). These are defined below with reference to Fig. 4.1 which represents a rough surface profile.

$$\sigma = \frac{1}{L}\sqrt{\int_o^L y^2 dx}$$

$$\sigma_c = \frac{1}{L}\int_o^L |y| dx$$

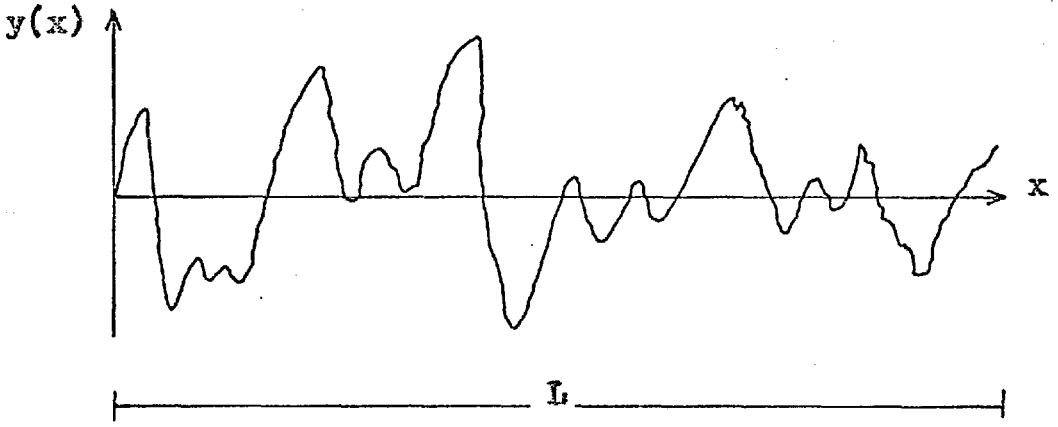


Fig. 4.1 Surface profile.

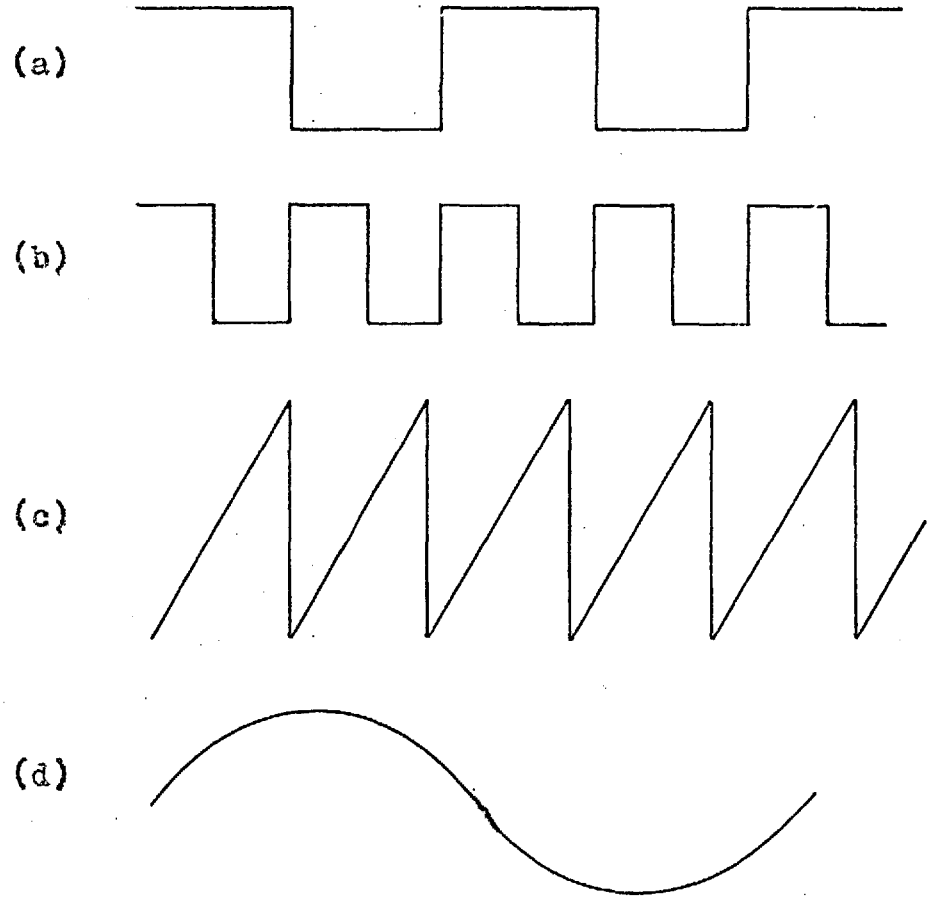


Fig. 4.2 Some profiles having the same c.l.a. value.

The length L must be sufficiently large for the sample to be representative of the surface. The mean line is taken here as the straight line which divides the profile in such a way as to cause equal areas to be cut off above and below, and that these areas are the minimum possible. These averages depend only on the height distribution of the surface and can give no information on, for instance, the shape and density of asperities. Fig. 4.2 shows some curves having the same c.l.a. value.

If the height distribution is known, σ and σ_c will be simply related. The profiles of 4.2(a) and (b) have the property that $\sigma = \sigma_c$. These are unrealistic models of a profile however. The Gaussian distribution is a good representation of many surfaces and it is proved in section 4.6 that $\sigma = 1.25\sigma_c$, for this case.

The most suitable use of these simple averages is for comparison of surfaces prepared by the same method, or the determination of the point at which a particular finishing process may be discontinued. A limitation of this application is demonstrated by considering two ground surfaces of markedly different roughness. Both surfaces will have the same general features of ridges lying parallel to the direction of grinding. A change of grit size will affect not only the roughness, but also the spacing between ridges and material properties modified by local heating.

Fig. 4.3 illustrates how a rough surface profile may be split into three components of roughness, waviness and error of form.

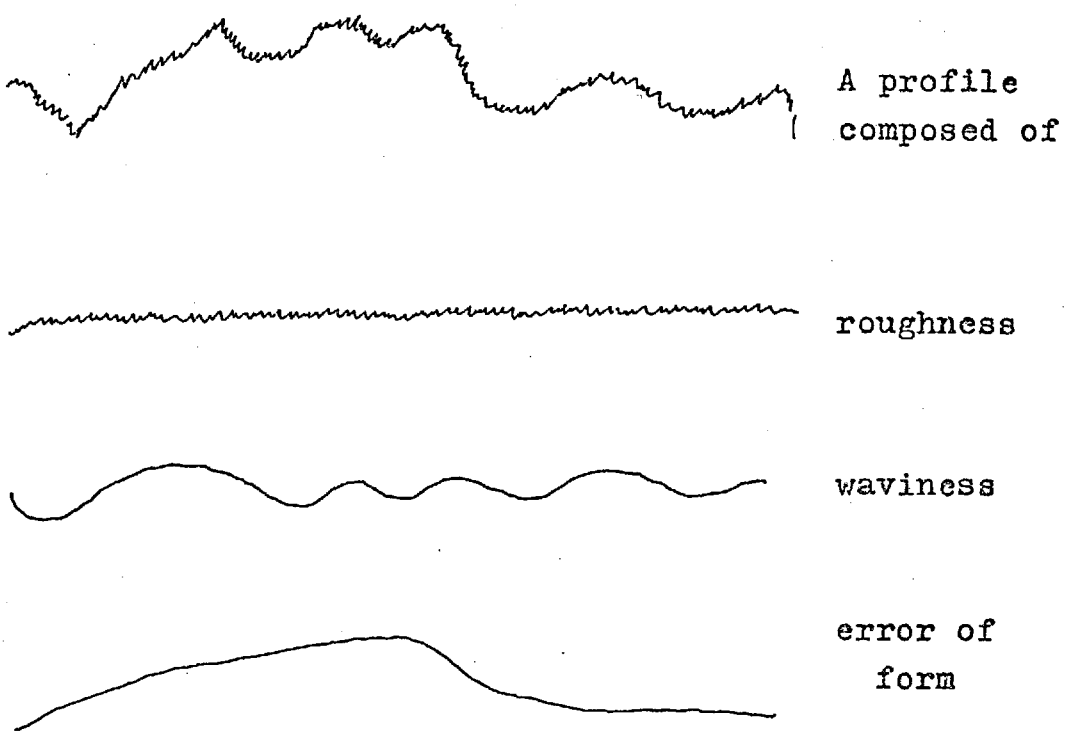


Fig. 4.3 Roughness components of a profile.

This is a somewhat artificial process as the frequency spectrum of a surface is continuous. A good illustration of this may be found in Leaver et al. (26). It does mean that three values can be attached to a specimen rather than one and allows rather better quality control. It can be argued that the roughness component is a function of the process and that the other factors depend on the properties of the particular system in use. Waviness measurement may, for instance, be a method of detecting a grinding wheel out of true or looseness in bearings.

Other statistics obtainable from a profile are the peak density, slope, curvature and the autocorrelation function (Acf.). All these quantities, and derived ones such as mean slope/mean curvature are of interest but the Acf. has

shown particular promise as a useful parameter and is discussed later.

The mathematical models that are used to represent surfaces can be divided generally into two classes, statistical and deterministic. The deterministic approach assumes a particular shape for asperities and where necessary, an alignment of opposing surface features. Calculated results are usually dependent on both shape and alignment and hence limited in application. The statistical approach can be very complicated and simplifying assumptions are frequently needed. It can also be very difficult to interpret some physical effects such as asperity collisions. The most promising approach seems to be a hybrid approach where a shape is assumed for a single asperity and then its dimensions and position are allowed to vary in a random manner.

4.3 The deterministic approach.

A simple theory of friction such as that proposed by Bowden and Tabor (64) requires that the real area of contact A_r is proportional to the applied load W. This is based on the premise that friction arises due to adhesion or welding of micro-contact points. If asperities behave plastically then A_r is proportional to W since the area of contact increases until all points in contact are just below their yield stress.

It is necessary to add a certain amount of bulk elastic behaviour into this simple model to take account of the contact load being reduced without any other motion between the contacting bodies. It is unrealistic to expect all

surfaces to be flowing plastically at each fresh contact, but it is more difficult to explain Amonton's law in terms of A_r when purely elastic models are examined. The real area of contact between smooth elastic spheres was examined by Hertz (65) who calculated that $A_r \propto W^{\frac{2}{3}}$. Lincoln (66) studied the area of contact between a nylon ball and a glass flat using interferometry and was able to verify this relationship. A surface consisting of an array of partial spheres of the same radius shows an $A_r \propto W^{\frac{4}{5}}$ dependence. Other shapes were assumed for asperities, for example Zhuraslev (67), and in general the variation of W with A_r depended on this shape. It was also found to be dependent on the particular alignment of opposing asperities on the contacting surfaces. Archard (68) proposed a model whereby asperities of different magnitude could be combined. This process is illustrated in Fig. 4.4 and it can be seen that as the smaller asperities are included the exponent of W approaches unity. This idea was an important step in the understanding of elastic contact between rough surfaces.

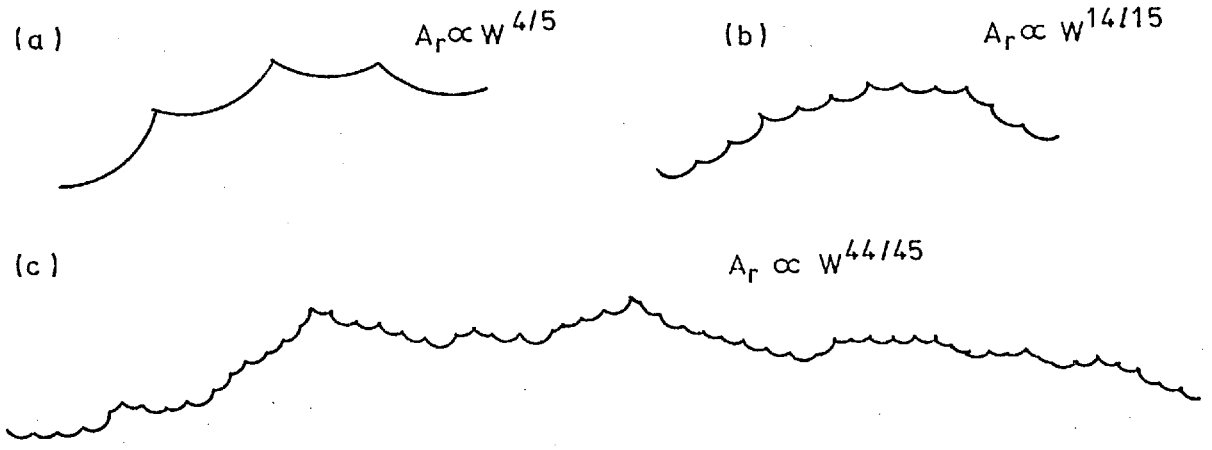


Fig. 4.4 The effect of combining asperities of differing dimensions.

Deterministic models can be very helpful when examining the effect of surface roughness on contact properties that are not easy to describe statistically. Christensen (69) examined the effect of two dimensional sinusoidal ridges on film thickness by means of such a model. Generally the oil film was thinner than the smooth case when the ridges were parallel to the direction of motion. The opposite effect was observed when the ridges were transverse to the direction of motion. A similar model was used by Napel and Bosma (41) who were interested in the effect of roughness on capacitance measurements. Their results showed that the capacitance could be artificially high when the surfaces were close.

These few examples give an indication of the uses of the simple surface models. When the dimensions of the surface features are altered and allowed to have statistical variation the effect can be profound, as indicated by the Archard model described above.

4.4 Statistical models.

The first order statistics contain no information on slope or curvature. Very little can be done with the height distribution only and some other constraint upon the system must be introduced in order to make further progress.

4.4.1 Greenwood and Williamson's approach.

In the theory of Greenwood and Williamson (70) the constraint they use is that the asperities are spherical caps of constant radius β. The peak heights of these caps are distributed normally with standard deviation σ^*. This

notation is used to distinguish σ^* from σ which is the standard deviation of all heights on a profile as opposed to peak heights. The model initially dealt with one smooth and one rough surface (71 and 72) but subsequently Greenwood and Tripp (73) showed that the problem of two rough surfaces in contact can always be represented accurately by one modified rough surface and a smooth flat plane.

4.4.2 Lubricated contact.

A model for load sharing between asperities and the oil film was developed by Thompson and Bocchi (74) in which they combined the Greenwood model with the hydrodynamic lubrication theory of Dowson and Higginson (35). A simple theory of asperity contact in EHL was proposed by Johnson, Greenwood and Poon (52) and this is now briefly outlined. Whitehouse and Archard (75) have shown how the statistics of peak heights are related to those of all ordinates. The separation between the mean plane of the surface and the mean plane of the peak height distribution is approximately 0.8σ, and $\sigma^* \simeq 0.7\sigma$.

A class of functions is now introduced:

$$F_n(u) = \frac{1}{(2\pi)^{\frac{1}{2}}} \int_u^\infty (t - u)^n \exp\left(-\frac{t^2}{2}\right) dt$$

and the values for some of these are given in (73).

The number of asperities in contact per unit area at a separation h' between the smooth surface and the mean plane of the peaks is given by $n = NF_0(\frac{h'}{\sigma^*})$ where N is the density of asperities.

The reduced elastic modulus E' is defined in terms of the elastic moduli and Poisson ratio of the two surfaces in contact by

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

A single asperity compressed by an amount w exerts a force of $\frac{2}{3}E'\pi^{\frac{1}{2}}w^{3/2}$.

The expected total pressure exerted by all asperities in contact is then given by

$$p_a = \frac{2}{3}E'N\beta^{\frac{1}{2}}\sigma^{*3/2}F_{3/2}\left(\frac{h'}{\sigma^*}\right).$$

The effect of surface roughness on the oil film is examined from two view points. They are the way in which the oil flow to the entry region is modified by roughness and the change in bulk deformation. The method used by Christensen (76) is applied for transverse and longitudinal asperities and it is shown that the effect of the roughness on the film formation can be neglected if $\frac{h'}{\sigma^*}$ is large.

The pressure distribution in the contact is assumed to have the same shape as that for a smooth surface although the actual values are changed. Thus $p(x) = \lambda p_f(x)$ where p_f is the film pressure and p is the pressure for a completely smooth surface. It is then shown that Dowson and Higginson's equation (35) can be modified for rough surfaces by replacing W by $\frac{W}{\lambda}$, E' by $\frac{E'}{\lambda}$ and h_o by $\bar{h}$. The $\bar{h}$ and h_o used here refer to the separation of the mean planes of the surfaces

for the rough and the smooth case respectively. Thence it is shown that

$$p_f = \left(\frac{h_o}{\bar{h}}\right)^{6.3}.$$

As the assumptions made so far limit the application of this to $h > \sigma$ it is now proposed that the $\frac{h'}{\sigma^*}$ used in the first part be identified with $\frac{\bar{h}}{\sigma}$ above. This is a reasonable assumption for values greater than about 1.5.

It is then possible to plot p_f and p_a (asperity pressure) on the same axis of $\frac{h}{\sigma}$ ($\frac{\bar{h}}{\sigma}$ for p_f) for various values of $\frac{h_o}{\sigma}$ (smooth case). The ratio of p_f and p_a and the equilibrium film thickness are then found from the point of intersection of the relevant asperity pressure curve with the appropriate $\frac{h_o}{\sigma}$ curve. In all cases where this theory may be applied the equilibrium film thickness is very close to that for the smooth surface.

The effect of the constriction regions in a contact is not included in this theory.

4.4.3 Contact time.

An interesting argument is used by Johnson et al. (52) to explain the variation of γ with $\frac{h}{\sigma}$. This is based on the argument that under a given set of conditions there is an average number of contacts, m, occurring within the nominal Hertzian contact area. Then the following line of reasoning is pursued: "Provided that m is a fairly small number, we might expect that the probability of finding a given number of contacts within the contact area to be given by a Poisson distribution. The probability of finding no contacts is then

$p(0) = \exp(-m)$ i.e. no-contact time fraction

$1-\gamma = \exp(-m)$."

According to the foregoing theory $m = NA_o F_o\left(\frac{h_o}{\sigma}\right)$ where A_o is the Hertzian contact area. The experimental results of Tallian (42) and Poon and Haines (27) are then compared with a theoretical curve obtained by taking $N = 200/mm^2$. This is illustrated in Fig. 4.5.

The effect of a load increase with data obtained from Tallian et al. (42) is to shift the theoretical curve in the right direction by approximately the correct amount. The work by Leaver et al. (26) does not show such good agreement, although based on the foregoing theory.

4.4.4 Conclusions from Johnson model.

1. The oil film is very much "stiffer" than the elastic bodies in contact so that the oil film thickness can be calculated as for the smooth case.

2. The number of contacting asperities is proportional to $F_o\left(\frac{h}{\sigma}\right)$.

3. EHL film thicknesses are virtually independent of load. The total number of asperities in contact is then dependent only on the contact area. This is contrasted with Greenwood and Williamson's model for dry contact whence direct proportionality between load and number of contacts is obtained.

The results obtained by Thompson and Bocchi (74) show a stronger dependence of asperity pressure on $\frac{h}{\sigma}$ than Johnson et al. This probably stems from the assumption of a constant nominal contact area and possibly from the

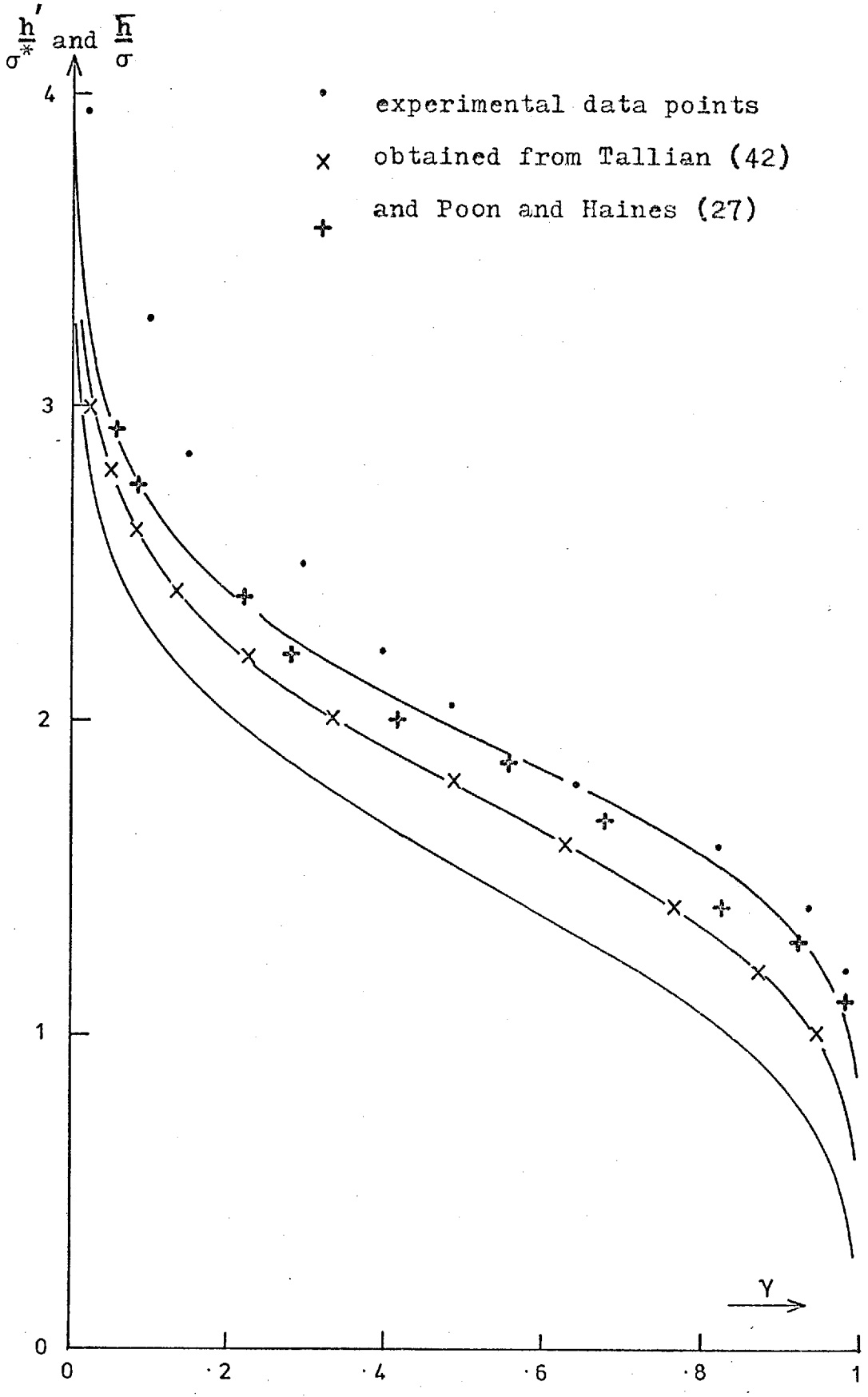


Fig.4.5 Theoretical prediction of no-contact time for three loads, after Johnson (52), compared with experiment.

confusion of $\bar{h}$ with h' and σ with σ^* between which pairs no distinction is drawn.

4.4.5 The Whitehouse and Archard approach.

Whitehouse and Archard (75) examined profile traces as though a random signal. Such a signal is completely described by its height distribution and its autocorrelation function. The Acf. is defined below, with reference to Fig. 4.1:

$$C(\beta) = \lim_{L\to\infty} \frac{1}{L} \int_{-L/2}^{L/2} y(x)y(x+\beta)dx.$$

Basically the profile is multiplied by a delayed version of itself. Any periodicity will show up as peaks in the plot of $C(\beta)$ against β. Note that $C(0) = \sigma^2$ (r.m.s. value) and also that $C(\beta) = C(-\beta)$ since the delayed profile may be taken as the reference without changing the value of C in any way.

Peklenik (77) employed the Acf. as a means of characterising real surfaces. In a study of many machining processes several different shapes for the Acf. were observed. The most common however was a sharp initial decay followed by a shallower gradient finishing sometimes with a small amplitude oscillation.

The basis of the Whitehouse and Archard model is the Gaussian distribution of all heights and an Acf. with exponential behaviour.

The value β^*, the autocorrelation distance, is the distance at which the Acf. has fallen to 10% of its initial value.

The practical use of these quantities was admirably demonstrated by Hirst and Hollander (78) in which σ and β^* were clearly related to damage in sliding contacts.

Whitehouse and Archard also developed a plasticity index ψ which is a measure of the likelihood that the surfaces deform plastically or elastically under load. This is defined as:

$$\psi = 0.69\left(\frac{E'}{H}\right)\frac{\sigma}{\beta^*}.$$

This compares with a similar expression derived by Greenwood and Williamson (70).

$$\psi = \left(\frac{E'}{H}\right)\left(\frac{\sigma}{\beta}\right)^{\frac{1}{2}}.$$

A more detailed comparison may be found in Onions and Archard (79).

Some theoretical objections to the exponential Acf. were raised by Nayak (80) and these are discussed below.

4.4.6 The Longuet-Higgins approach.

The statistics of a time dependent random surface were evolved by Longuet-Higgins (81,82 and 83) for oceanographic research. The special case of a metal surface was developed by Nayak (86).

A recurring problem in the area of rough surface modelling is that of inferring three dimensional behaviour from two dimensional cross-sections. In this respect the work of Nayak has been most helpful. Generally the profile suggests lower values for the mean slope, curvature and heights of the asperities than actually occur. For instance,

the mean slope on a profile is a factor of $(\frac{1}{2}\pi)^{\frac{1}{2}}$ lower than that of the corresponding random surface. The density of summits on a surface and peaks on a profile are related by

$$D_{sum} \simeq 1.2(D_{peak})^2.$$

It is further shown how the moments of the distribution are related to extremal and mean line crossings on a profile. These moments can be evaluated from the derivatives of the Acf. of the distribution at the origin. It is in this context that problems arise from the exponential Acf. described in section 4.4.5. Since the Acf. is symmetrical the first derivative at the origin has two values depending on whether the approach is from the positive or negative side. However, in use these problems do not interfere with the practical applications of the Whitehouse and Archard model as the sampled systems by which the Acf. is measured seem to impose smooth behaviour at the origin.

Tallian (84) has used some of the above ideas in a partial EHL theory, which draws together many aspects of interest viz. surface fatigue, wear, smearing and elastic, plastic and oil film load sharing. Tallian also asserts that the effect of roughness on film thickness, as studied by Christensen (69) and Fowles (85), is no more significant than other commonly neglected effects such as non-Newtonian behaviour and thermal considerations. This is not true of traction however. Tallian's experimental evidence (84) shows that traction can be halved with a circumferential lay on a bearing, and doubled or halved with transverse lay. It is also pointed out that if micro-EHL is significant and there

is a pressure gradient over an asperity, then surface roughness modifications must be introduced into partial EHL models.

4.5 Real surfaces and practical measurement.

The approaches to surface characterisation have been outlined above and a closer look is now taken at real surfaces and the validity of certain assumptions occurring in the literature.

A recurring problem is that of scale and what resolution is necessary in order to obtain significant statistical parameters from the surface. The resolution of the stylus profilometer is limited by the tip radius and the construction of its support mechanism. Work with a finer stylus (75) revealed extra surface features and no doubt this process could be continued to reveal surfaces in ever greater complexity, until the stylus load and radius exceeded the plasticity limit. The finite width of the stylus leads to the recorded profile consisting of the envelope of the highest points across the width of the track. There is also a tendency for the stylus tip to be guided along grooves, giving misleading values of roughness and slope.

For applications where the contact of two metal surfaces is being considered, it is perhaps possible to make out a case for ignoring the higher frequencies in the profile spectrum. The asperities on a surface with significant high frequency content must have small values for β, the tip radius of curvature. Referring to the plasticity index defined above, $\frac{\sigma^*}{\beta}$ will then be large. This means that plastic flow is likely until either σ^* is reduced or β increased, indicating

that the high frequency is being modulated. This effect would occur early in the history of the contact and one would expect the lower frequency components to dominate subsequent behaviour.

The most common assumption made in surface modelling is that of a Gaussian distribution of heights or peak heights. In the case of very light contact an exponential distribution may be used to approximate the upper extremes of this distribution with subsequent simplification of the mathematics. Not all surfaces are distributed Gaussianly as reference to Peklenik (77) will confirm. Fig. 4.6 shows the height distribution of a bead blasted aluminium specimen. This is plotted on probability paper which would represent a cumulative Gaussian distribution as a straight line. Fig. 4.7 shows a steel surface finished by abrasion. The distribution of heights seems to be made up of two Gaussian distributions, the lower one having a larger spread than the upper. This is consistent with a surface prepared by one process then finished by another. In the contact of surfaces the upper half of the distribution is the more important. The surface behaves in contact as though it consisted entirely of the upper distribution. This upper part is commonly normal even for otherwise strongly non-Gaussian surfaces.

In his study of the microcartography of rough surfaces, Williamson (86) examines the contact of two equally rough surfaces, the results of which are displayed in Fig. 4.8.

It can be seen that under very light pressure the surfaces are separated by 2.5 standard deviations and that even under fairly heavy loads 90% of the contact takes place

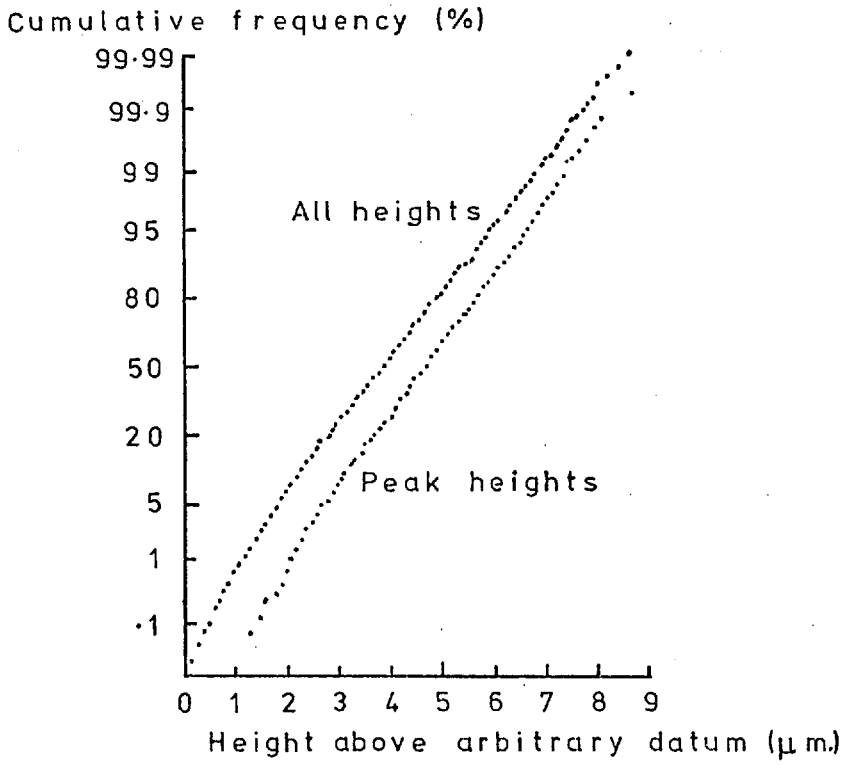


Fig. 4.6 Surface height distributions of bead blasted aluminium.

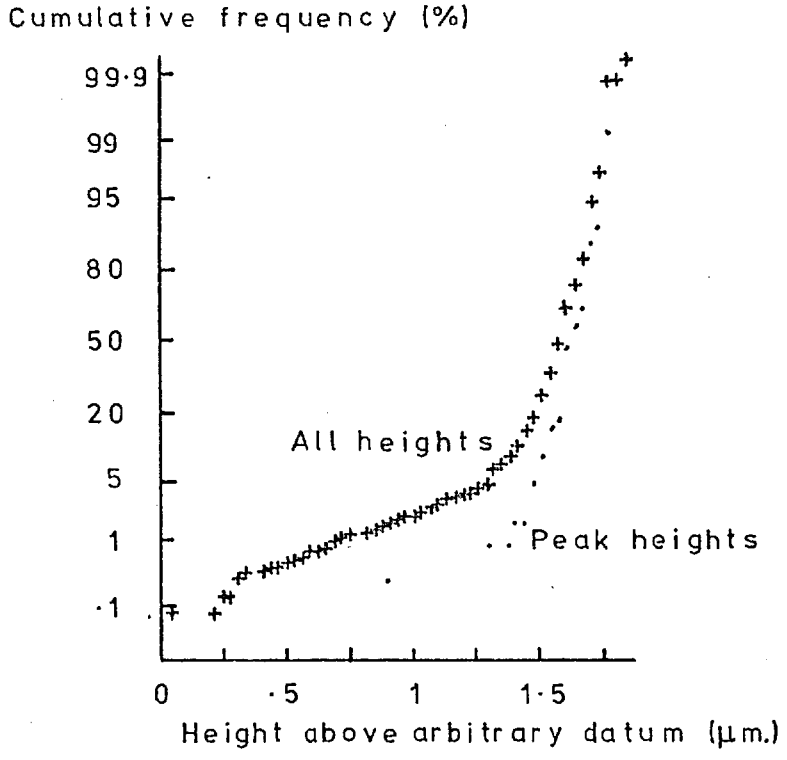


Fig. 4.7 Surface height distributions for abraded steel.

in the upper half of the surfaces. Thus it seems that the Gaussian assumption for rough surfaces is generally acceptable.

Separation between mean planes (standard deviations).	Approximate load on 1 square cm. nominal area (Kg.).	Percentage of nominal area in contact.	90% of contact lies above this percentile.
2.5	0.2	0.4	87
2.0	0.8	2.2	77
1.5	3.0	7.4	60
1.0	8.0	17.0	50

Fig 4.8 Contact between two surfaces with $\sigma = 1.25\mu m$.

4.6 The relationship between averages for a Gaussian surface.

The root mean square value of the surface is σ, the standard deviation of the distribution. The probability of finding an ordinate of height y is then given by

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{y^2}{2\sigma^2}\right].$$

Direct measurement of surface roughness is normally given in terms of the centre line average σ_c. A simple proof is given below to show how these averages are related.

Fig. 4.9 shows a profile of unit length with a Gaussian height distribution. The shaded area represents $\sigma_c = \int_0^1 |y|dx$. This integral must be evaluated in a slightly irregular manner since the equation relating y and x is not

known. Instead of summing small increments in x from the left hand side to the right, the increments are summed with respect to the ordinate y. Thus all the terms in dx that correspond to ordinates lying between y and y+dy are added together and finally integrated for all values of y. The process is illustrated in the figure.

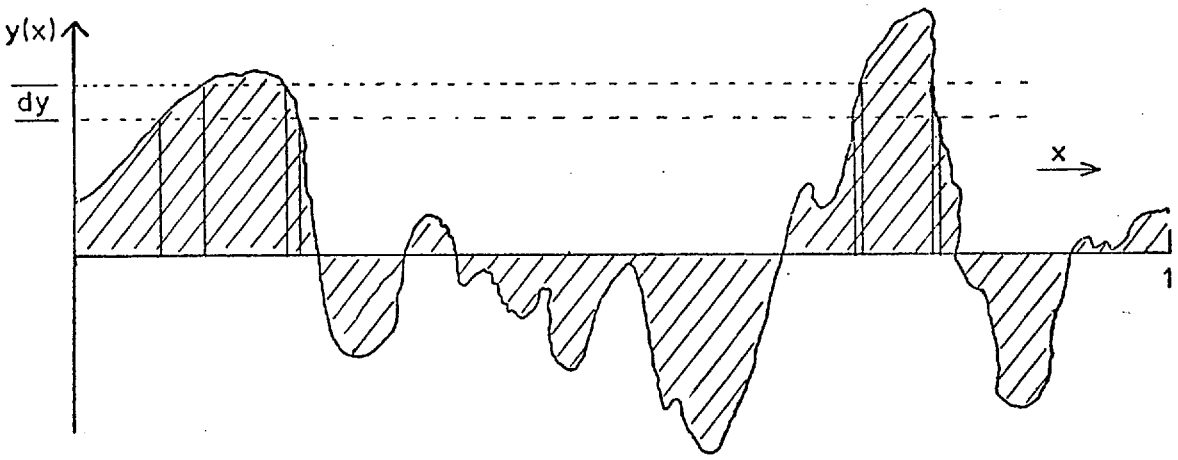


Fig 4.9 Profile of unit length showing method of integration.

Thus $$dx = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{y^2}{2\sigma^2}\right]dy$$

since the distribution of y is known.

Then $$\sigma_c = \int_{-\infty}^{\infty} |y| \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{y^2}{2\sigma^2}\right]dy.$$

$$\sigma_c = \frac{2}{\sigma\sqrt{2\pi}} \int_0^{\infty} y \exp\left[-\frac{y^2}{2\sigma^2}\right]dy$$

$$= -\sigma\sqrt{\frac{2}{\pi}} \exp\left[-\frac{y^2}{2\sigma^2}\right]\Bigg|_0^{\infty}$$

$$= \sigma\sqrt{\frac{2}{\pi}}.$$

Thus $$\sigma = \sigma_c\sqrt{\frac{\pi}{2}} = 1.25\sigma_c.$$

4.7 The duration of sliding contacts.

In pure rolling the electrical contact is normally assumed to persist for a time equivalent to passing through the Hertzian zone. For a point contact this assumption results in a distribution of contact durations, as the two opposing asperities may meet near the edge of the circular Hertzian area. For a line contact, however, all electrical occurrences should have the same duration. If a small amount of sliding is introduced some contacts will be shortened although most will remain unaltered. If the sliding is increased sufficiently the duration of conducting occurrences is no longer governed by the Hertzian transit time but by the time taken by asperities to slip past each other.

In order to appreciate the effect of surface features on contact time a simple mathematical model is described below. It is important to make realistic assumptions when setting up such a model. For instance, the conical asperity model is at its most inaccurate where its behaviour is most critical i.e. at the tip. The Greenwood and Williamson approach is amenable to this sort of analysis and their assumptions have a sound physical basis. Thus Gaussian distributions of asperity peaks with standard deviation σ_1^* and σ_2^* and tip radii of curvature β_1 and β_2 are assumed for the surfaces. The symbols y_1 and y_2 represent the peak height distributions for the two contacting surfaces having standard deviations σ_1^* and σ_2^* respectively. The mean of these functions is zero and the planes in which $y_1 = 0$ and

$y_2 = 0$ are the reference planes. An "overlap function" w is defined for positive values only as

$$w = y_1 + y_2 - h'$$

where h' is the distance between the reference planes. The function $y_1 + y_2$ may be replaced by y which is also normally distributed with zero mean and has standard deviation σ^* where $\sigma^{*2} = \sigma_1^{*2} + \sigma_2^{*2}$.

Now $$p(y) = \frac{1}{\sigma^*\sqrt{2\pi}} \exp\left[-\frac{y^2}{2\sigma^{*2}}\right]$$

Thus $$p(w+h') = \frac{1}{\sigma^*\sqrt{2\pi}} \exp\left(-\frac{(w+h')^2}{2\sigma^{*2}}\right).$$

Then the distribution of w, given that $w > 0$ is given by:

$$q(w) = \frac{\frac{1}{\sigma^*\sqrt{2\pi}} \exp\left[-\frac{(w+h')^2}{2\sigma^{*2}}\right]}{\frac{1}{\sigma^*\sqrt{2\pi}} \int_o^\infty \exp\left[-\frac{(w+h')^2}{2\sigma^{*2}}\right] dw}$$

$$= \frac{\exp\left[-\frac{(w+h')^2}{2\sigma^{*2}}\right]}{\sigma^* \varphi\left(\frac{h}{\sigma^*}\right)} \qquad 4.1$$

where $\varphi(x) = \int_x^\infty \exp(-\frac{1}{2}t^2)dt$.

This distribution is illustrated in Fig. 4.10 and q(w) is proportional to $p(w+h')$.

The contact between colliding asperities is taken to be elastic. No particular mechanism of deformation is suggested; the only condition imposed is that the asperities regain their initial shape as soon as is geometrically

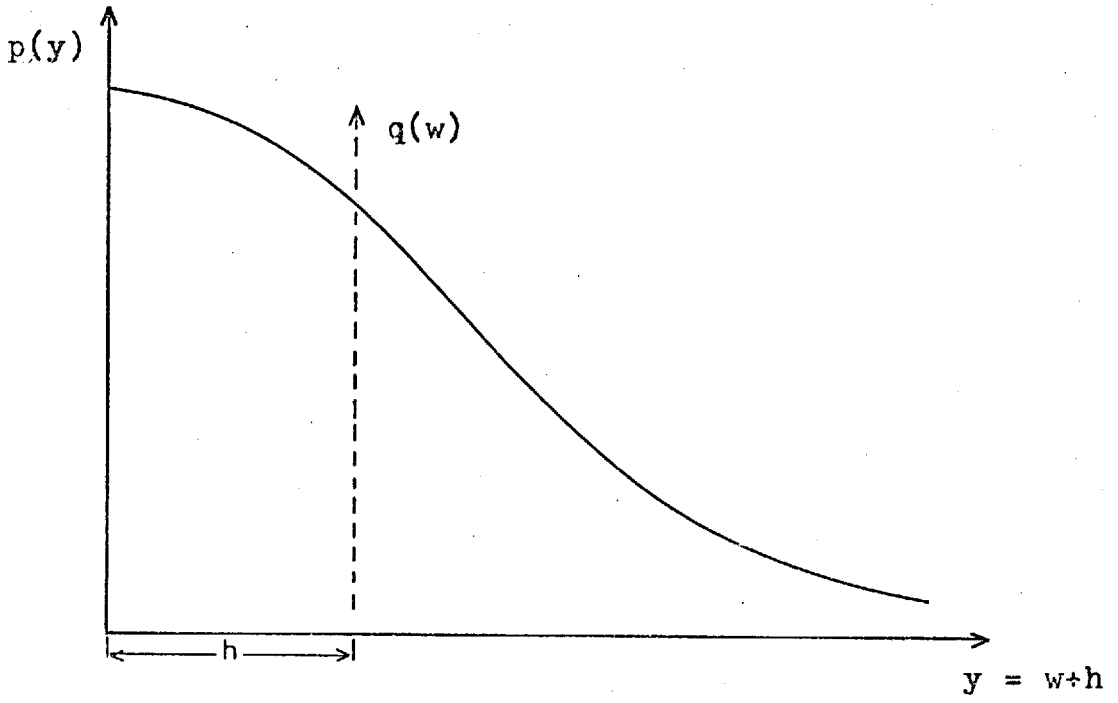


Fig. 4.10 Distribution of overlap function q(w).

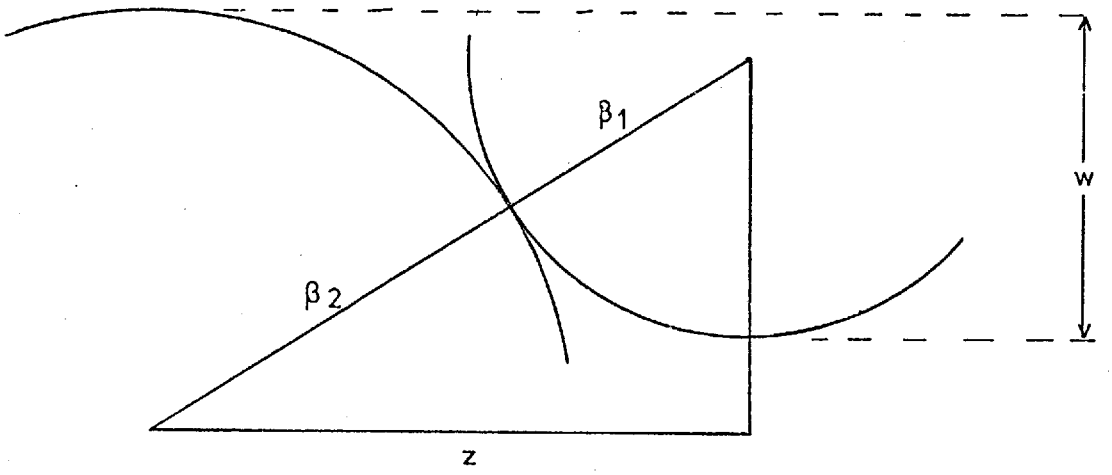


Fig. 4.11 Cross-section through asperities at initial contact.

possible.

The initial contact of two asperities of radii β_1 and β_2 is shown in cross section in Fig. 4.11. The projection of the line joining the centres of the two asperities onto a reference plane is of length z where

$$z^2 = 4\beta^2 - (2\beta - w)^2$$

$$= w(4\beta - w) \text{ where } \beta = \tfrac{1}{2}(\beta_1 + \beta_2).$$

In general $\beta \gg w$ so that $z \simeq 2\sqrt{w\beta}$ 4.2

This line of length z can further be resolved into two components, y perpendicular to and x parallel to the relative direction of motion of the two surfaces. The duration of the contact will be 2x in terms of distance and for any given w y is equally likely to take any value between 0 and z. The mean value of x is now calculated for any value of z.

The distribution of y is constant, thus

$$p(y) = \frac{1}{z}$$

since $$\int_0^z p(y)dy = 1$$

The mean value of x or $\sqrt{z^2 - y^2}$ is then given by

$$\bar{x} = \int_0^z \left[\frac{1}{z}\sqrt{z^2 - y^2}\right] dy$$

$$= \left[\frac{y}{2z}\sqrt{z^2 - y^2} + \frac{z}{2}\sin^{-1}\left(\frac{y}{z}\right)\right]\Bigg|_0^z$$

$$= \frac{\pi}{4}z = \frac{\pi}{2}\sqrt{w\beta} \text{ from equation 4.2.}$$

Therefore the mean duration of contact is given by

$$2\bar{x} = \pi\sqrt{w\beta} = K(w) \text{ (say)} \qquad 4.3$$

This is the expected value of duration for any

given w, but for a separation h′ there is a range of values of w given by equation 4.1

$$\text{i.e.} \quad q(w) = \frac{\exp\left[-\frac{(w+h')^2}{2\sigma^{*2}}\right]}{\sigma^* \varphi\left(\frac{h'}{\sigma^*}\right)}$$

Thus the expected value of K(w) for any fixed h is given by

$$\int_o^\infty K(w)q(w)dw.$$

$$\text{Thus} \qquad \bar{K} = \frac{\pi\sqrt{\beta\sigma^*}}{\varphi(\lambda)} \int_o^\infty \sqrt{\mu} \exp\left(-\tfrac{1}{2}(\mu+\lambda)^2\right) d\mu$$

$$\text{where} \quad \mu = \frac{w}{\sigma^*}, \quad \lambda = \frac{h'}{\sigma^*}.$$

$$\text{We may write} \quad f(\lambda) = \frac{1}{\varphi(\lambda)} \int_o^\infty \sqrt{\mu} \exp\left(-\tfrac{1}{2}(\mu+\lambda)^2\right) d\mu \qquad 4.4$$

since f is a function of λ only and then $\bar{K} = \pi\sqrt{\beta\sigma^*}\, f(\lambda)$ i.e. the mean duration depends only on the surface parameters and the dimensionless film thickness.

λ	$f(\lambda)$
0.0	0.822
0.5	0.730
1.0	0.656
1.5	0.599
2.0	0.549
2.5	0.509
3.0	0.474
3.5	0.435
4.0	0.333

Fig. 4.12 The function $f(\lambda)$ defined in equation 4.4.

The function $f(\lambda)$ tabulated in Fig. 4.12 is obtained from a set of integrals published by Greenwood and Tripp (73) which are simply related to the functions developed above. It can be seen that $f(\lambda)$ varies by about 60% from $\lambda = 0$ to 2.5 which is the practical working range. This function is plotted in Fig. 4.13.

Whitehouse and Archard (75) have shown that the standard deviation, σ^*, of peak heights on a surface is related to the standard deviation, σ, of all heights by $\sigma^* \simeq 0.7\sigma$. Thus $\bar{K} = \pi(0.7\sigma\beta)^{\frac{1}{2}} f(\lambda)$. If the surface speeds are U_1 and U_2, then the mean contact dwell time will be

$$\tau_m = \frac{\bar{K}}{U_1 - U_2} \simeq \frac{2.63\sqrt{\sigma\beta}}{U_1 - U_2} f(\lambda) \qquad 4.5$$

The maximum error caused by replacing $f(\lambda)$ with a constant will be about 19% in the range considered. Thus the mean duration is approximately described by

$$\tau_m = \frac{1.7\sqrt{\sigma\beta}}{U_1 - U_2} \qquad 4.6$$

The only direct measurement of contact duration was made by Christensen (48) and his measured values are compared with a calculated figure. Christensen's experimental conditions were as follows:

$$U_1 - U_2 = 0.27 \text{ metres/sec.}$$
$$\sigma = \sqrt{2} \times 0.035\mu\text{m.}$$

No value for β was published and this is taken as 40μm. which is reasonable for a mean value although much larger values can occur.

Then $\tau_m = \frac{1.7}{0.27}\sqrt{(.035 \times \sqrt{2} \times 40)}$ μsecs.

$\simeq 8.6$μsecs.

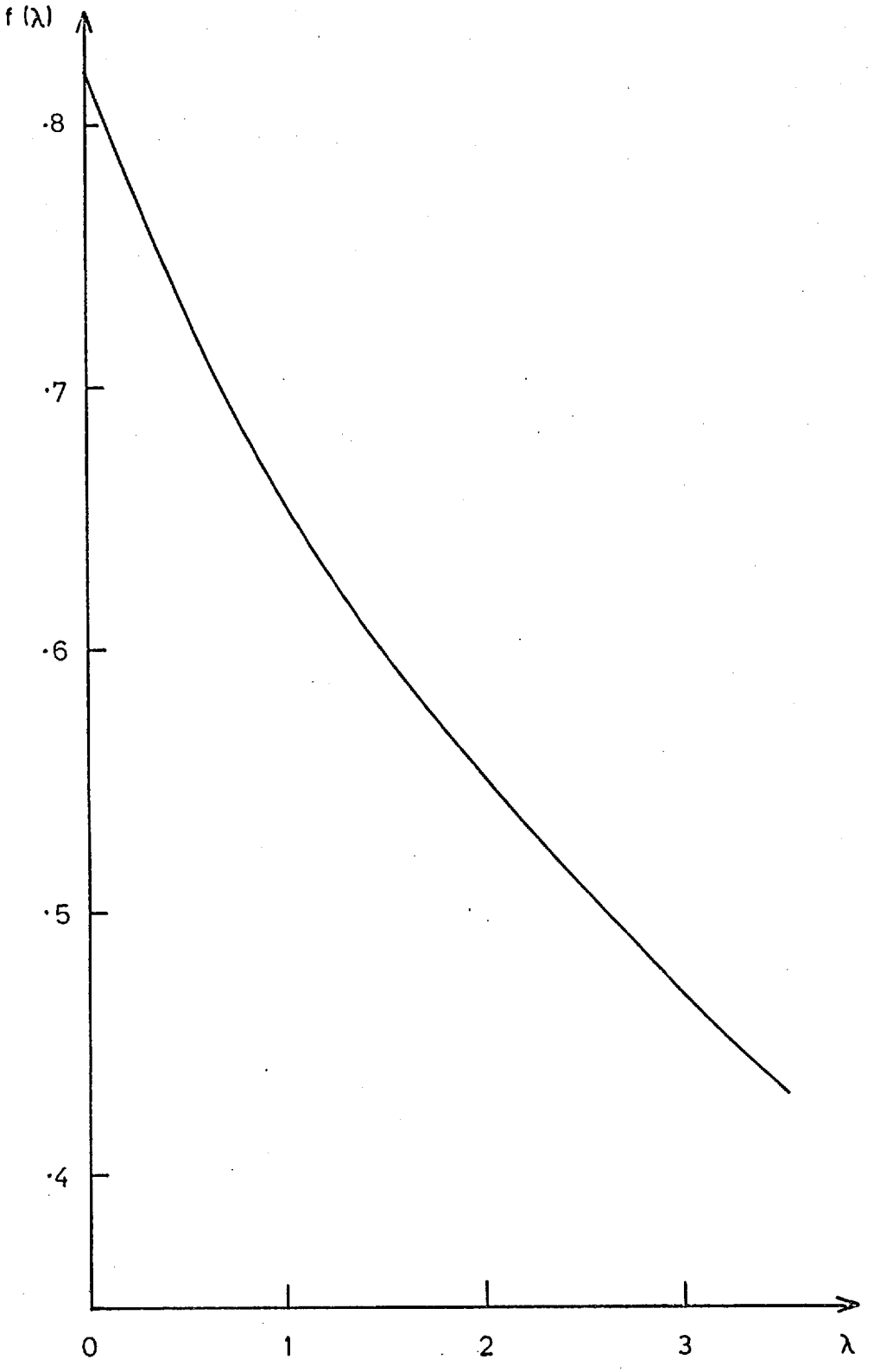


Fig 4.13 The variation of $f(\lambda)$ with λ.

The observed durations shown in Fig. 2.16 show peak values occurring between 1 and 4μsecs.

A similar model based on conical asperities along the lines proposed by Hisakado (87,88 and 89) yielded a result two orders of magnitude in excess of the above calculated value.

The distribution of dwell time is also of interest and is dominated by the overlap function. This function shown in Fig. 4.10 can be approximated by an exponential curve when h' is large compared with w. Christensen's results have again been employed to test this supposition and the results are shown in Fig. 9.12.

This model may be simply extended to asperities of more complicated shape. For elliptical asperities, the value of β used must be that measured along the surfaces in the direction of sliding. In the case of the transverse radius becoming infinite i.e. cylindrical asperities, the difference in τ_m is less than 22% of that for the spherical case.

Thus if τ_m can be measured accurately by electrical means the basis exists for an on-line measurement of the $\sigma\beta$ product . A method for obtaining τ_m experimentally is discussed later.

4.8 Rolling or sliding dominance.

If τ_m is much greater than the rolling contact transit time (say one tenth of the nominal Hertzian transit time) observed dwell times will be typical of the rolling

contact although the small amount of sliding present will increase the scatter. A simple criterion for determining how much sliding is necessary to change the electrical behaviour from rolling to predominantly sliding would then appear to be, for line contact,

$$\tau_m < \frac{2b}{10U_1}$$

where b is the Hertzian half width. The value taken for contact duration under rolling is one tenth of the Hertzian transit time.

Thus $$\frac{1.5\sqrt{\sigma\beta}}{U_1-U_2} < \frac{2b}{10U_1}$$

or $$\frac{U_1-U_2}{U_1} > \frac{7.5\sqrt{\sigma\beta}}{b}$$

Again using Christensen's results where $b \simeq 310\mu m$, this value is about 0.7%

To extend this argument to point contact it would be necessary to estimate a typical value for the duration of rolling contact.

This value derived above is that at which sliding behaviour dominates. It is not unreasonable to suppose that the effect of sliding will be marked at somewhat lower values of sliding.

Chapter 5. Disc machine experiments.

5.1 Introduction.

Several experiments were conducted on a new design of disc machine. The discs were supported on hydrostatic bearings and loaded by oil pressure (Fig. 5.1).

Both the shafts connected to the discs were driven by thyristor controlled motors, arranged so that one motor provided the main drive whereas the other applied a controlled braking force. A fuller description of this machine may be found in (90).

The driven disc was electrically isolated from the shaft assembly and a thin layer of epoxy cement ensured that the side of the disc did not contact the earthed case. The elctrical take off was by means of several silver and graphite slip rings connected in parallel. The driving disc was earthed to the body of the machine.

The spigot used to drive the insulated disc in these electrical tests was somewhat weaker than its conducting counterpart and it was necessary to restrict the operating conditions accordingly.

Temperature was measured by means of a trailing thermocouple at the inlet region of the contact. It was found that the system heated up by the action of the high pressure oil pump. Some control over the oil temperature was possible by means of a heating element and water cooled coils located in the oil sump.

The test discs were run under flooded conditions,

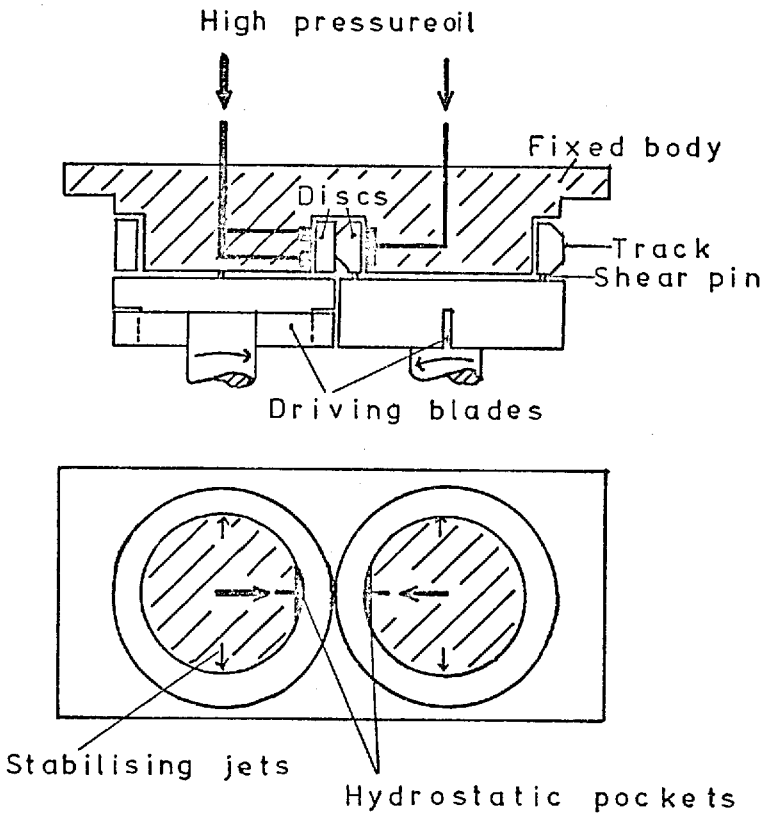


Fig. 5.1 Support and drive arrangement of disc machine.

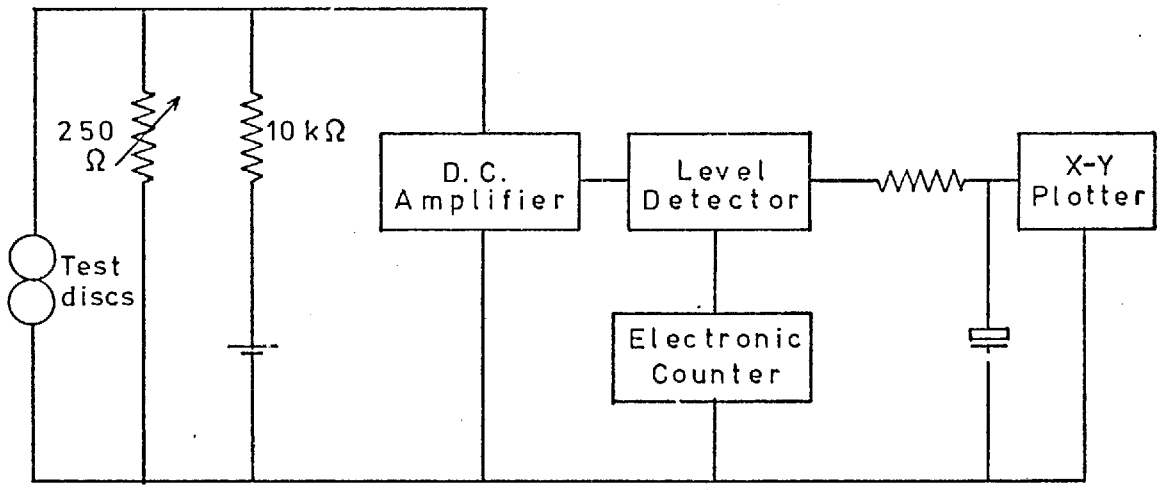


Fig. 5.2 Block diagram of electrical circuitry.

the oil being recirculated and returned via an oil filter. The oil used for all the tests was a formulated industrial oil conforming to D. Eng. R.D. 2487 with a diacid ester base and containing antioxidant and corrosion inhibiting additives.

The mating discs ran on a track 6mm. wide and one disc was chamfered at an angle of 4° on both sides of the running track to reduce plastic flow. The discs were of EN39B steel and were finished by grinding circumferentially.

5.2 Electrical circuitry.

The electrical arrangement is shown schematically in Fig. 5.2. The potential applied to the discs was supplied by a 1.5volt dry battery through a dividing network and it was set at 15mV. The output from the discs was amplified by the d.c. amplifier and this drove a simple switching circuit. This circuit had two outputs, the first of which went into an electronic frequency counter, and the second drove an averaging circuit and x-y plotter. The time constant of this circuit was measured from its decay curve and was about 2.7 seconds. It is emphasised that there was no interference between these two outputs. This was tested by applying various known frequencies to the input and examining the counter display both with and without the averaging circuit connected. This simple circuit performed surprisingly well when tested for frequency response and linearity. A pulse generator was connected across the input and known mark-space ratios fed into the circuit at frequencies from 10Hz to 50 kHz (the limit of the generator). Both frequency response and linearity were within 2% of full scale and lay within the

errors of estimating mark to space ratio on an oscilloscope.

Electrical noise produced by the three large motors on the test machine initially caused considerable obstruction to the proper functioning of the circuit. This could be demonstrated by disconnecting one side of the battery while running and observing counts on the frequency meter. This problem was eventually overcome by careful earthing and the use of screened leads throughout. The circuit was triggered by a disc resistance of about 30Ω.

5.3 Experimental results.

5.3.1 Running-in phenomena.

The discs were run in at fixed speed and the load was increased in 100lb. steps. The specimens were run for twenty minutes at each load for the lower loads and at forty minutes each at the higher loads.

At the higher loads, the application of the next increment caused the contact time fraction to rise to 100%. Furthermore, the count rate dropped to zero for a short time indicating that any excursions that occurred from the contact voltage were faster than the frequency limit of the system. Typically this state continued for about twenty or thirty seconds after which low count rates appeared, increasing rapidly with time. The contact time fraction gradually dropped from the 100% level over the running-in period.

Removal and re-application of the applied potential for short periods of time did not appear to change the measured values appreciably.

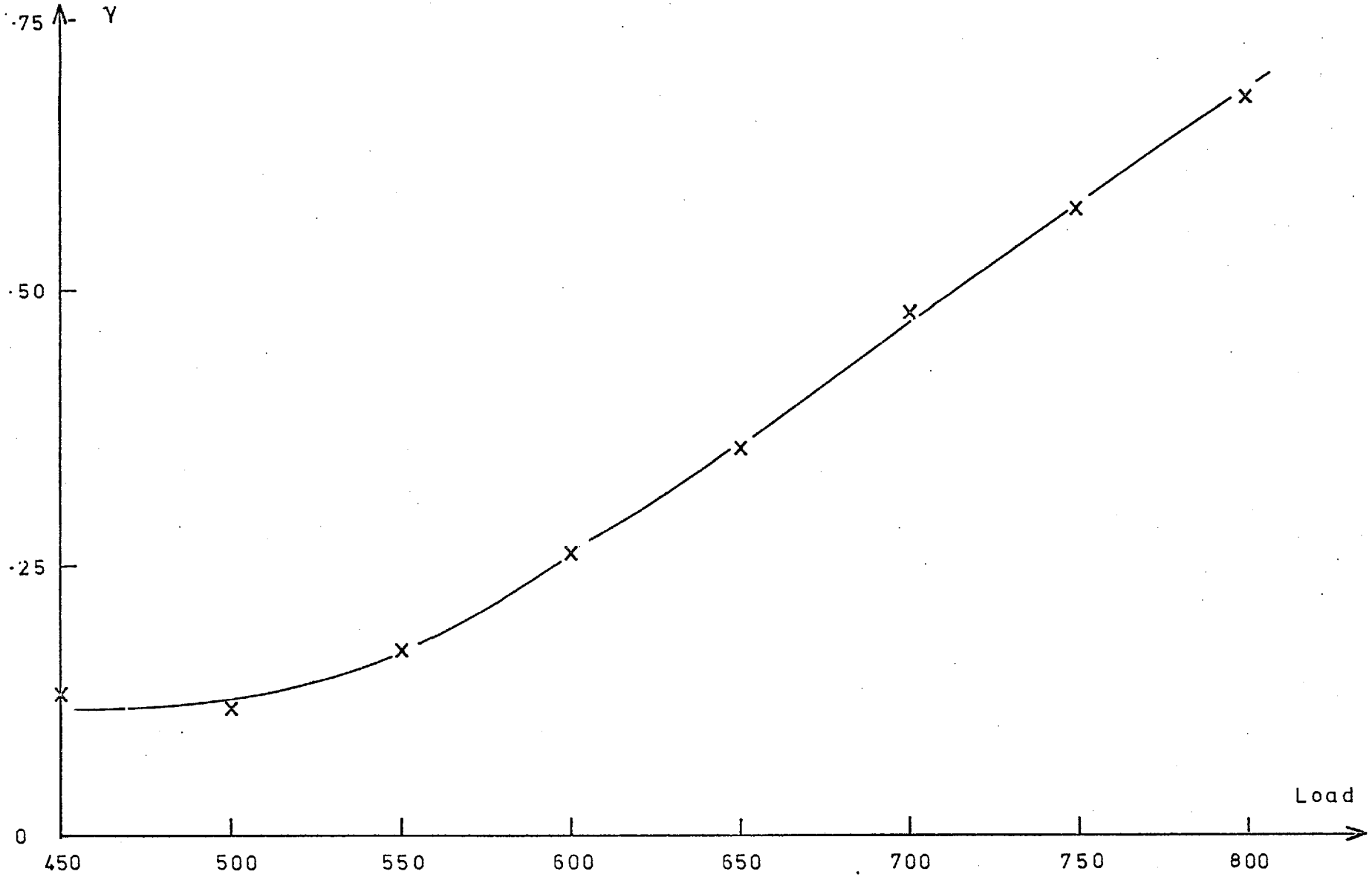


Fig. 5.3 Typical variation of γ with load.

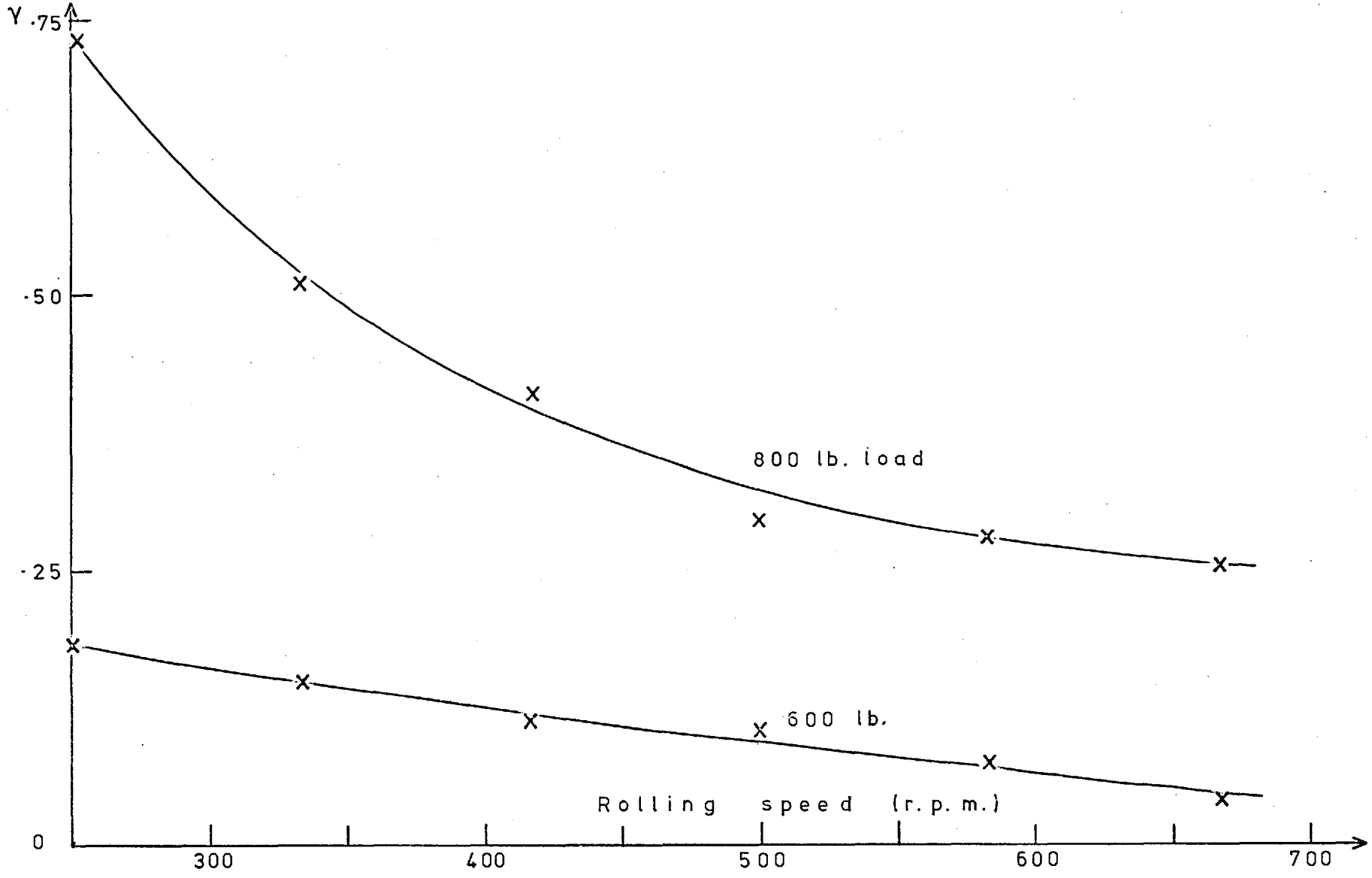


Fig. 5.4 The variation of γ with speed for two loads.

5.3.2 Load and speed variation.

The range of working loads and speeds was very restricted, as mentioned above, and it was not possible to cover the full range of values of contact time fraction by controlling a simple variable. A typical load variation is shown in Fig. 5.3, and Fig. 5.4 shows the variation with speed at two fixed loads.

A problem encountered was the heating up of the test oil by the high pressure pump, particularly at high loads. In order to reduce this effect, measurements were taken as quickly as possible starting from the highest loads and working down. The actual value measured was taken as the average over the second half minute at the test conditions. Local fluctuations did occur but were generally small compared with the values measured. A small amount of electrical mains hum was also present on the recorder input.

At low loads a surprisingly high amount of contact was often observed. This appeared to be peculiar to this test machine and might have been due to slight misalignment artificially increasing the local contact pressures. This effect was not investigated further but loads were kept sufficiently high to avoid this region.

5.3.3 Rate of contact measurements.

Fig. 5.5 shows the mean and standard deviation of rate of contact interruptions. Each point represents the average of fifteen measurements. It can be seen that contacts initially increase with load then rapidly fall off as multiple contact occurrences become more frequent.

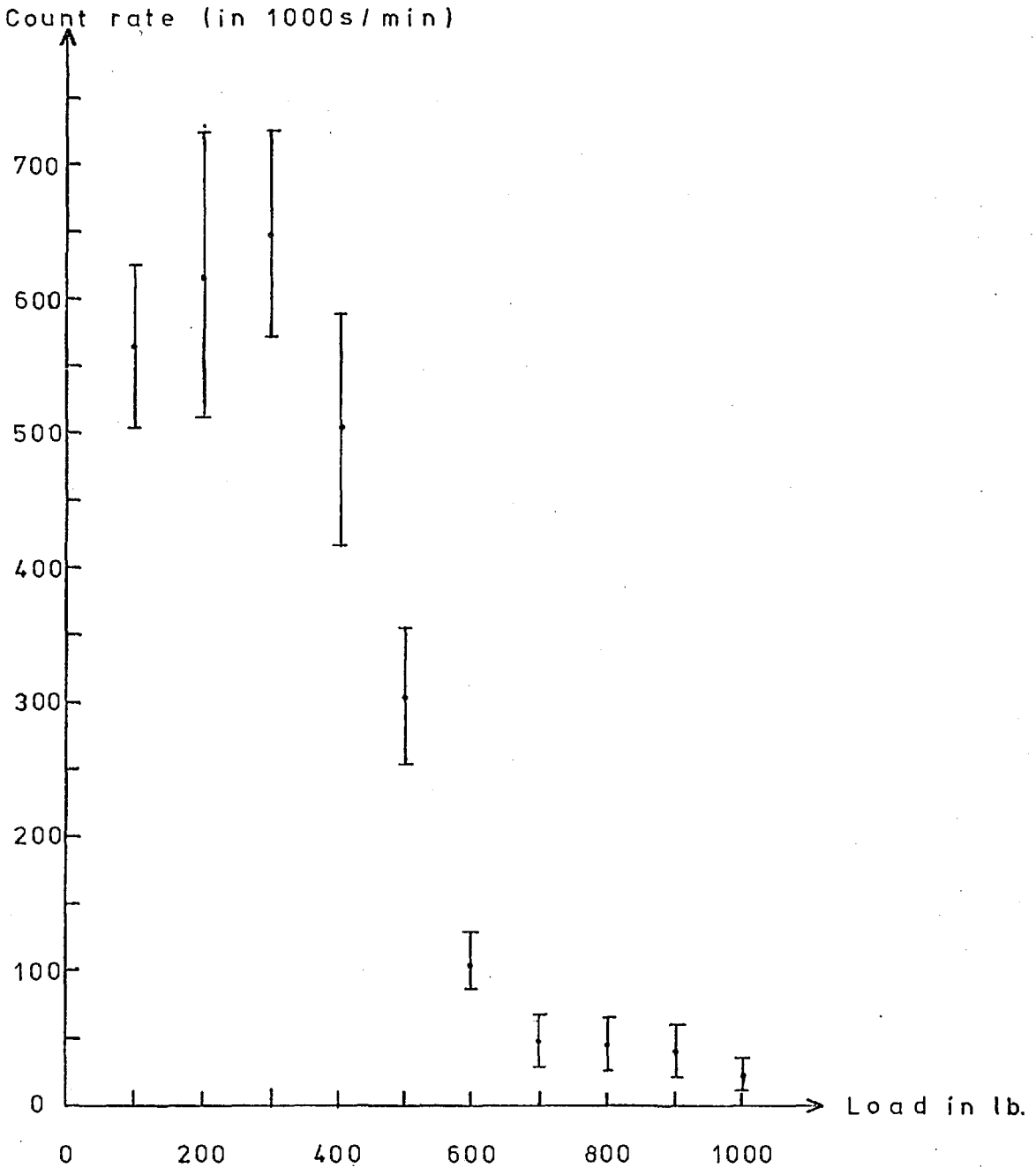


Fig. 5.5 The variation of count rate with load.

The electrical circuit registered a count when the discs switched from a contact to a no-contact state.

5.3.4 Repeatability of results.

Whenever experiments along the lines of those outlined above were performed, the general behaviour of the variables was always as reported. The actual value of contact time fraction varied considerably from run to run under the same nominal conditions. One effect that was consistently observed was that if the discs were left in the oil flooded chamber for a long enough period of time (typically over a weekend) a contact inhibiting film appeared to develop on the discs. After ten or fifteen minutes of running the contact resumed values closer to those normally encountered.

Apart from this effect there was still a variation that could not be ascribed to any obvious cause, such as continued running in of the discs. Possible explanations for this spread were chemical effects, slight changes in relative speed between the discs or temperature variations. As temperature could be varied (though not entirely controlled) a series of experiments was performed to investigate this further.

5.3.5 Temperature effects.

The averaging circuit was connected to one channel of an x-y plotting table, and the other channel was taken to the output of an electronic thermometer. The running conditions were set up and the oil heater switched on before connecting the plotter. Typical outputs are displayed in Fig. 5.6. Several graphs showed interesting behaviour in

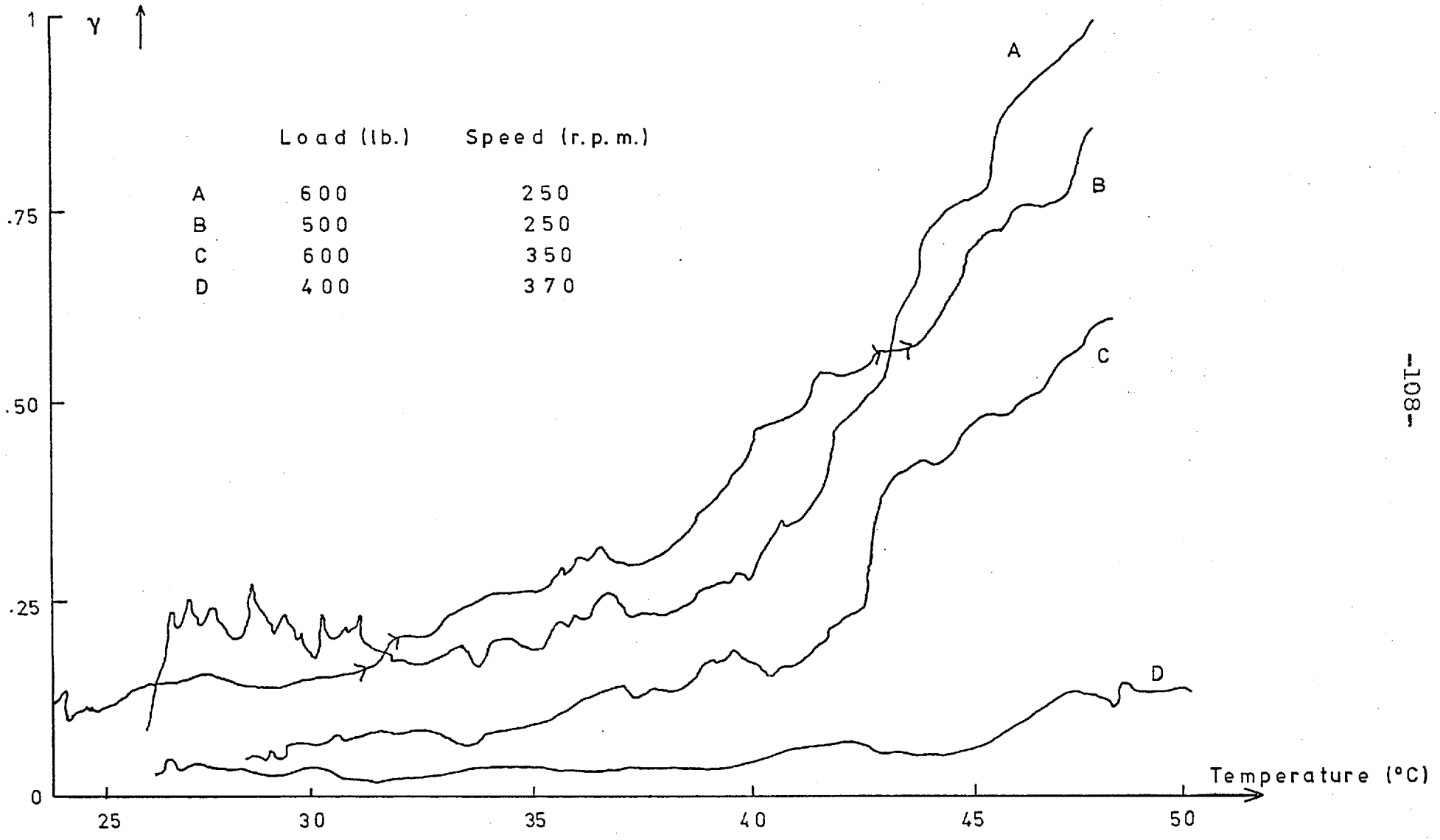


Fig. 5.6 Variation of γ with temperature.

the region of 40°-45°C. A good example of this is shown in Fig. 5.7. It was thought possible that this could be explained by the formation of soap films at these temperatures. It is interesting to note that Macpherson (91) sometimes observed a drop in friction traces in this temperature range.

In an attempt to control temperature as a variable and examine the scatter of results a number of "steady-state" experiments were performed. The disc machine was set running and left under the set conditions for several hours to attain a state of equilibrium.

A sampling circuit was constructed to sample the average contact time fraction and simultaneously trigger the frequency counter. The sampling rate was about four times a minute. A typical output is shown in Fig. 5.8. The count rate for the first thirty readings varied from 56.7 to 19.8×10^3 counts/minute. Towards the end of the experiment the temperature dropped and the contact was reduced.

5.4 Treatment of results.

It was intended to obtain surface topographical characteristics from the discs as these parameters were expected to play a most important part in electrical contact phenomenum. These discs were, however, subject to considerable surface damage when a high rate of sliding was accidentally introduced under heavy loading. It was thought that these surfaces would not be representative of those prevailing when the above programme was carried out. This misfortune precluded a theoretical treatment of electrical contact along the lines adopted by previous workers with all the inherent assumptions

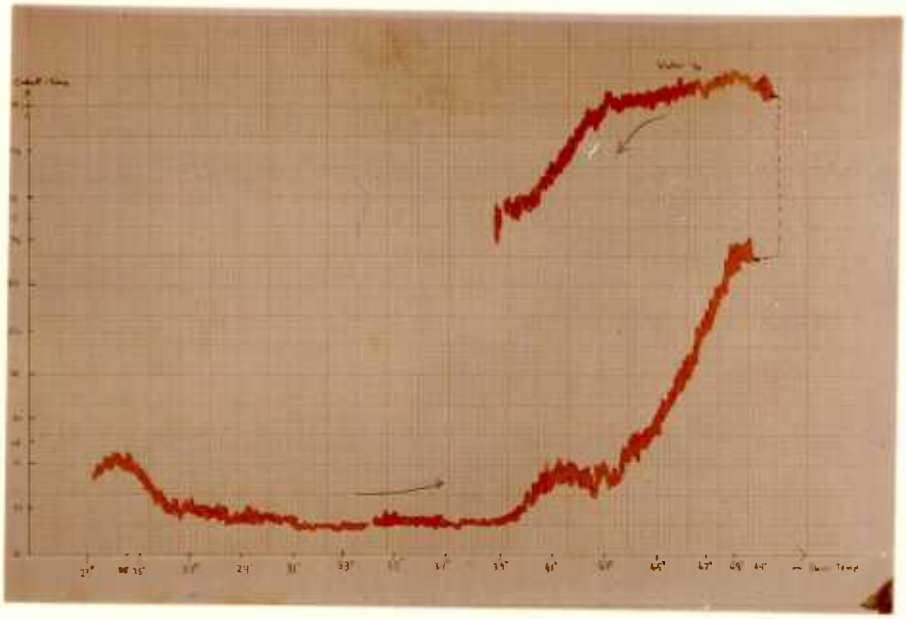


Fig. 5.7 Variation of γ with temperature.

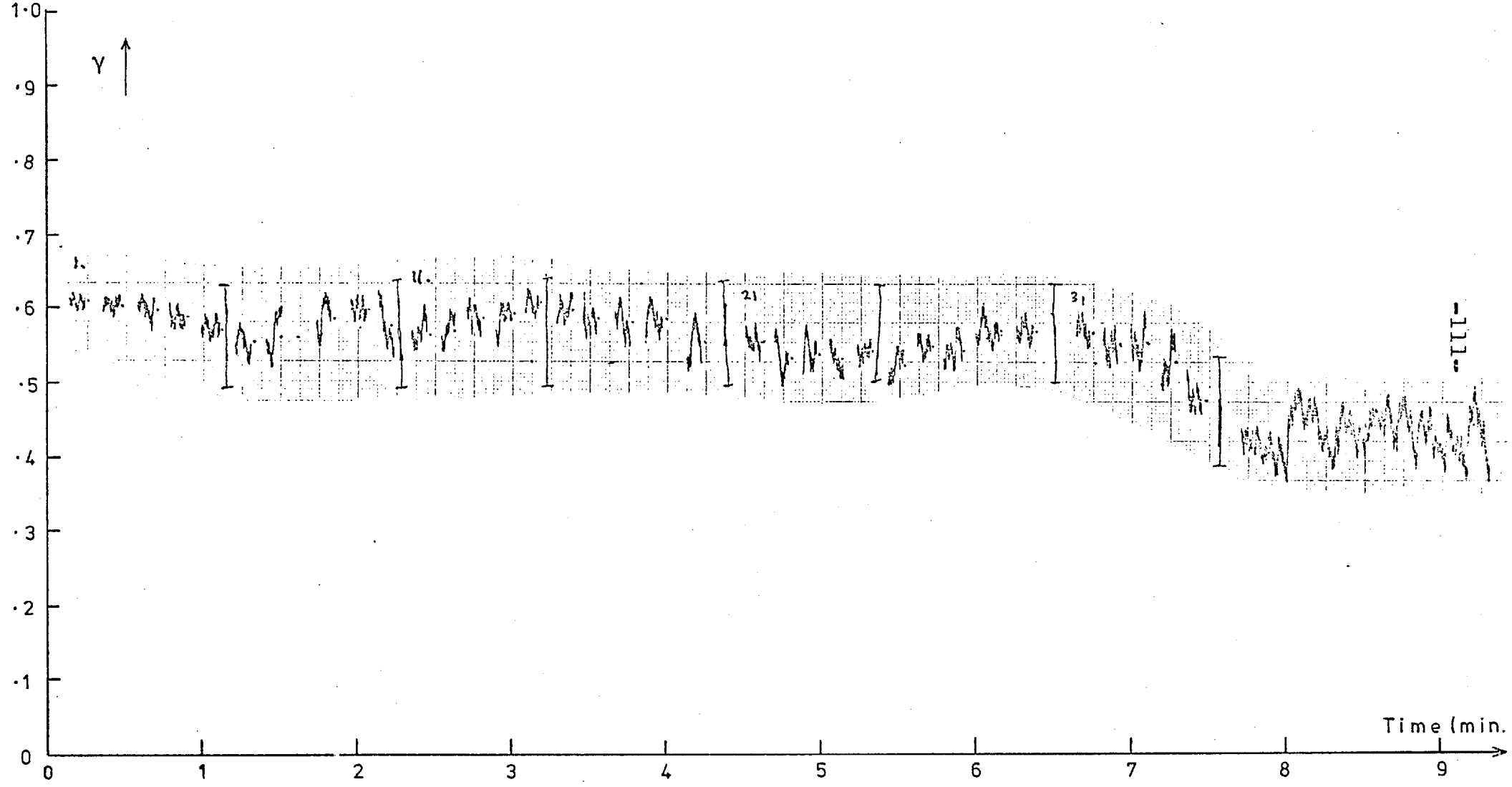


Fig. 5.8 Output of steady state experiment.

that such a theory would necessitate. With regard to the experimental results obtained and the difficulty of accurately reproducing or maintaining a given set of conditions such a theoretical approach would in any case be of very limited value. What is ideally required is a method of analysis which measures some electrical parameter of significance rather than assumes one. The common assumption, for instance, that electrical contacts persist for the Hertzian transit time is particularly susceptible to criticism.

There seemed to be no promising line of research from these results at the time so further work on the disc machine was discontinued. However, theoretical progress made at a later date and described in chapter nine, reveals some interesting characteristics of electrical contact phenomena. This theory was subsequently applied point by point to the steady state experimental results, and the mean contact duration was calculated according to the model described in section 9.4. Some of these results are shown in Figs. 5.9 and 5.10. The calculated values of dwell time were widely spread and in general much shorter than the Hertzian transit time.

5.5 Summary.

This series of experiments was performed with a common lubricant and typical engineering surfaces. The chemistry of the system has been largely neglected and it was not deemed fruitful to study this at any length without knowing the detailed composition of the oil. It seemed necessary to test this method of studying lubricated contacts

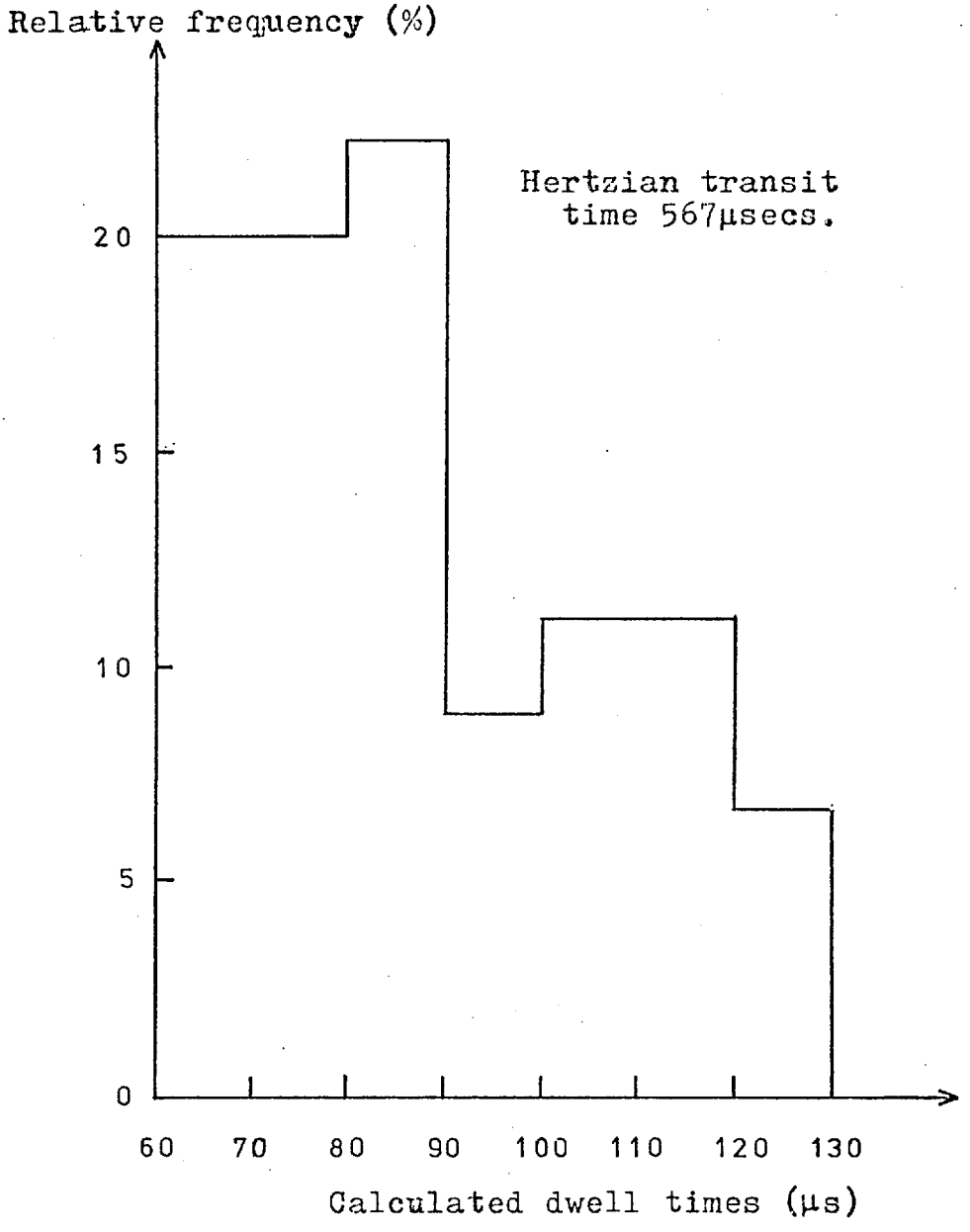


Fig. 5.9 Distribution of calculated dwell times for 600lb. load and 250 r.p.m. speed.

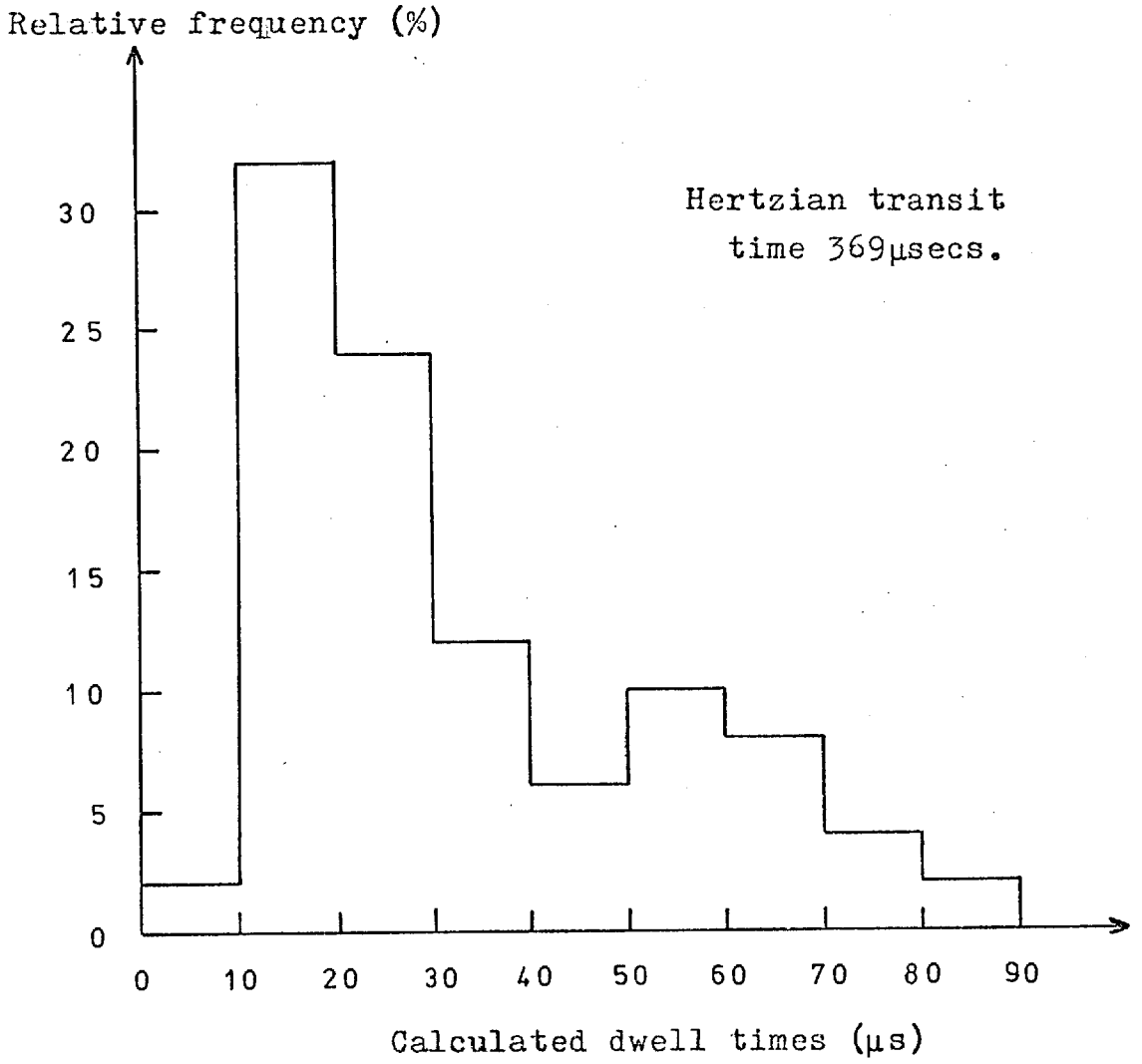


Fig. 5.10 Distribution of calculated dwell times for 650lb. load and 400 r.p.m. speed.

under typical working conditions if it was to be eventually of practical use. The temperature effect shown in Fig. 5.7, the soaking effect described in section 5.3.4 and the results of Millns (22) are a direct consequence of this philosophy. Otherwise the results were in line with previous work in that factors which increased contact width and reduced film thickness increased the severity of electrical contact and vice versa.

Increasing the load during an "asperity counting" experiment caused the expected effect of an initial rise of contact rate followed by a decrease as multiple contacts dominate, although it is not known whether this is an inherent property of the system or possibly caused by the increase in the number of intervals between contacts becoming shorter than that detectable by the equipment employed.

Running-in effects were noted in that an initial increase in load could reduce the asperity count rate to zero, but after a short time counts began to appear and to increase rapidly.

Problems were encountered with reproducibility of experimental conditions, which seem to limit the application of this electrical technique to qualitative uses only.

A theoretical approach, subsequently developed and described in chapter nine, is able to obtain quantitative data from some of the experiments performed above.

It is interesting to note that Leaver et al. (26) have more recently observed a similar anomalous effect at

low loads as that described in section 5.3.2. A possible explanation for this is concentration of load at one end of the line contact.

5.6 Direction of future work.

The experimental results seemed to shed no light on the detailed process of electrical contacts and a new line of research was undertaken. Work was transferred to a tapered roller line contact rig designed by Wymer (40). This rig was previously used for optical interferometric studies and it was thought useful to retain this feature in subsequent work. It was necessary to develop surface coating methods in order to produce the necessary equipment and this work is described in chapter seven. It was also necessary to find a method of obtaining statistically significant data from the electrical contact signals. This is described in chapter six together with the electronic equipment that was designed and built to do this job.

Chapter 6. Experimental approach.

6.1 Introduction.

The aim of this series of experiments was to study the electrical contact signal in greater detail than has been reported previously. The emphasis was on the examination on a statistical scale of changes in electrical contact resistance. No particular physical significance is placed on actual resistance values but large changes of resistance are regarded as events. It was then hoped to relate any statistically significant quantities to the mechanical parameters of the contact.

The test machine in these experiments was designed by Wymer (40) and comprised a tapered roller loaded against a flat glass disc. Provision was made for observation of film thickness by optical interferometry and although not of primary importance to these experiments this useful facility was retained.

In order to obtain the most physically consistent electrical data an attempt was made to simplify the contact to avoid end effects at the extremities of the roller. A special glass test disc was constructed with a conducting annulus contacting the centre of the roller. The very controllable technique of radio frequency (r.f.) sputtering was employed in the fabrication of this disc to ensure the best possible match between the surface of the disc and the top of the conducting annulus.

A technique for processing the electrical data on a statistical scale is described in section 6.4 and

essential electronic equipment was constructed to perform this function.

The oil used in these experiments was a paraffinic type, L74/1113, kindly supplied by B.P. Ltd.

6.2 Tapered roller test machine.

The general plan of the tapered roller thrust bearing assembly is given in Fig. 6.1. For electrical insulation the two needle roller bearings were pressed into Tufnol inserts. The drive sprocket was pressed onto a Tufnol sleeve and the lower race of the ball thrust bearing was insulated from the base plate. The cage retaining the tapered roller was also insulated by means of a Tufnol insert. An insulated wire passing through the cage made contact with a mercury pool inside the roller. A hollowed rod containing a mercury pool was fitted to the shaft of the test disc and an electrical contact was taken off by means of a co-axial cable. Thus the roller and rotor assembly and the disc shaft were at a floating potential with respect to the main frame which was earthed to provide shielding. This system was tested by lightly running a chromium coated disc without test oil which indicated continuous electrical contact.

The system was loaded by a hydrostatic bearing supported in a piston assembly which was pressurised by compressed air. This is shown in Fig. 6.2. The test disc was able to tilt with respect to its shaft as a soft rubber insert connected the two pieces.

The thrust bearing was driven by a $\frac{1}{3}$h.p. shunt wound d.c. motor which was mounted on resilient rubber

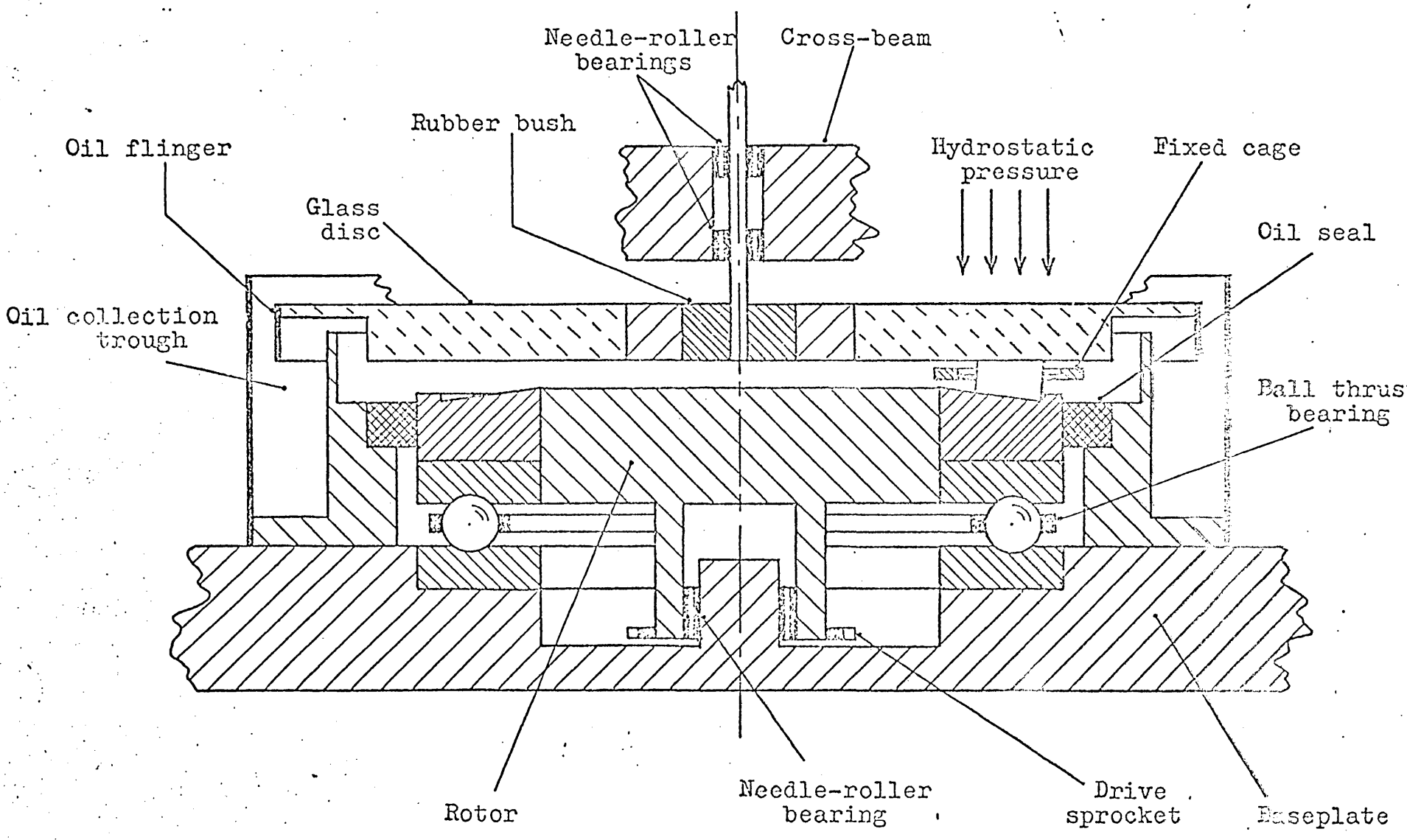


Fig. 6.1 Thrust bearing assembly (loading system omitted).

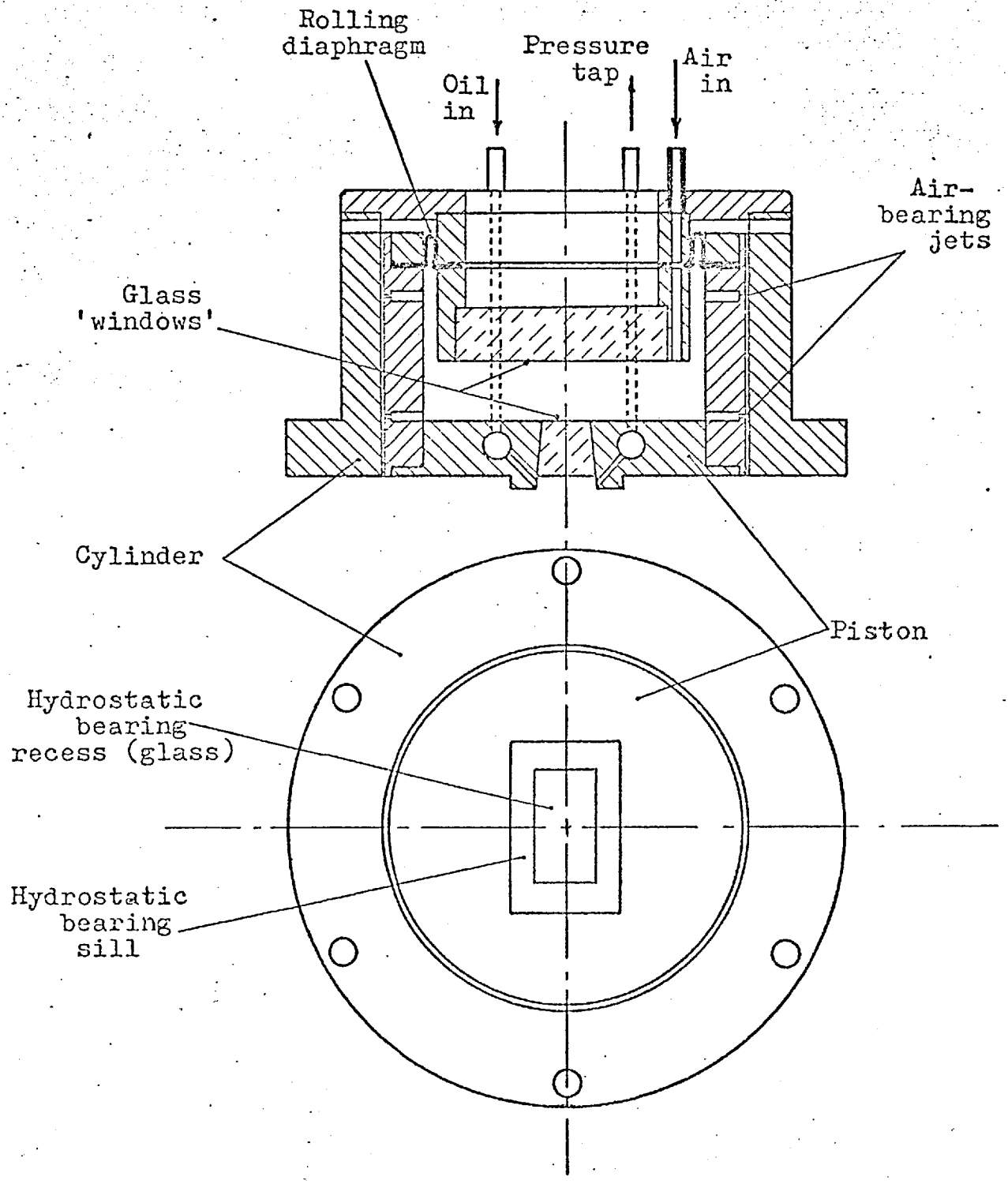


Fig. 6.2 The loading system.

supports to minimise vibrations. A speed reduction of 12.5:1 was effected by a worm gearbox and speed was measured by an electronic tachometer triggered by a magnetic transducer in close proximity to a 120 tooth gear. A thyristor circuit gave speed control of better than 2%.

Temperature was measured by means of a chromel/ alumel thermocouple fitted through a hole in the cage assembly behind the contact.

The dimensions of the blended roller used in these experiments are as follows:

length	1.366cm. (0.538in.)
maximum radius	0.457cm. (0.180in.)
minimum radius	0.365cm. (0.144in.)

The roller was superfinished by British Timken Ltd. to a roughness of about 0.025μm. (1μin.) c.l.a.

Further details and a description of the optical system may be found in (40).

6.3 Specification of glass test disc.

It was thought that the interpretation of electrical contact measurements would be considerably simplified if the effective contact could be reduced essentially to the two dimensional case. The means of achieving this would be to provide a conducting annulus on the glass test disc. This conducting band must also be very thin to avoid any edge effects on the oil film. Very good conductors such as gold or copper tend to smear (92) when highly stressed. Harder materials (e.g. chrome) have poorer conductivity and a thick layer is required. The ideal solution is to excavate an

annular trough into the disc and then fill it with a conducting substance. To avoid serious stress concentration at the edges the conducting material must be extremely close to the disc surface. It is possible to deposit substances to within measurable limits by the r.f. sputtering process. It was felt that the thinnest possible film should be employed but it was hoped that the sides of the trough would prevent a gold film from smearing. The adhesion of gold onto glass is generally poor but an interposing layer of titanium or chromium can considerably improve it. A cross section through the conducting annulus is shown in Fig. 6.3.

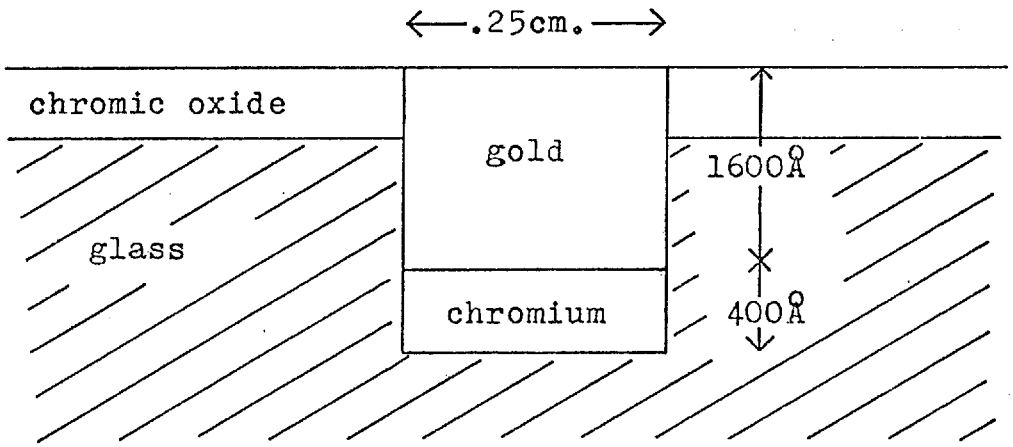


Fig.6.3 Cross section through proposed disc. (Not to scale.)

The position of the annulus on the test disc is shown in Fig. 6.4.

The etching and filling of the trough are separate operations and it is essential that the same mask be used for both. An attempt was made to etch the trough by chemical methods. The disc surface was covered with a

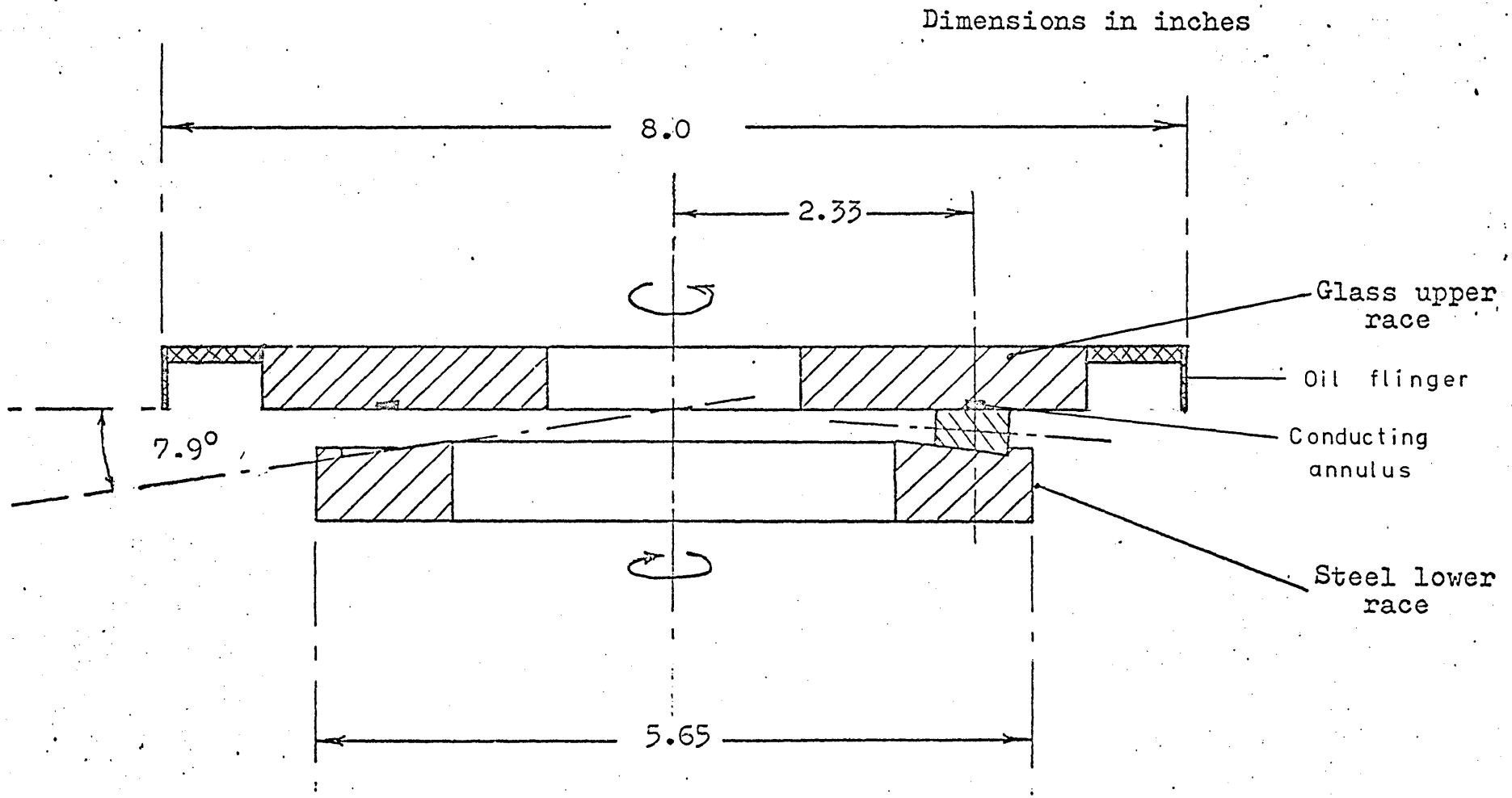


Fig. 6.4 Thrust bearing configuration

silicon based dry film photo-resist. A suitable mask was prepared, the film was developed and the annulus removed. The top surface of the disc was then etched in hydrofluoric acid. The etching was extremely uneven and the annular region took on a clouded appearance. In some areas the protective coating was penetrated. Fig. 6.5 illustrates the etched glass disc. In view of this failure a more ambitious approach was adopted. By reversing the power connections in an r.f. sputtering system it is possible to sputter away the substrate. The use of a suitable mask enables patterns to be etched. This line of development was pursued although fraught with practical difficulties and is described in chapter seven.

6.4 Processing of electrical data.

6.4.1 General approach.

The two quantities that are being examined in this experiment are the duration of contacts and the intervals between the onset of contacts. If, for instance, electrical contacts persisted for the time taken for an asperity to travel through the Hertzian zone then one would expect to find a large number of contacts having this duration. One might also expect the intervals between the onset of contacts to be related to some surface topographical quantity such as the average wavelength. Thus the problem falls into two parts. Firstly, statistically significant parameters relevant to contact duration and spacing must be established. Secondly, these must be related to the physical conditions inside the contact.

It is necessary to obtain electrical signals

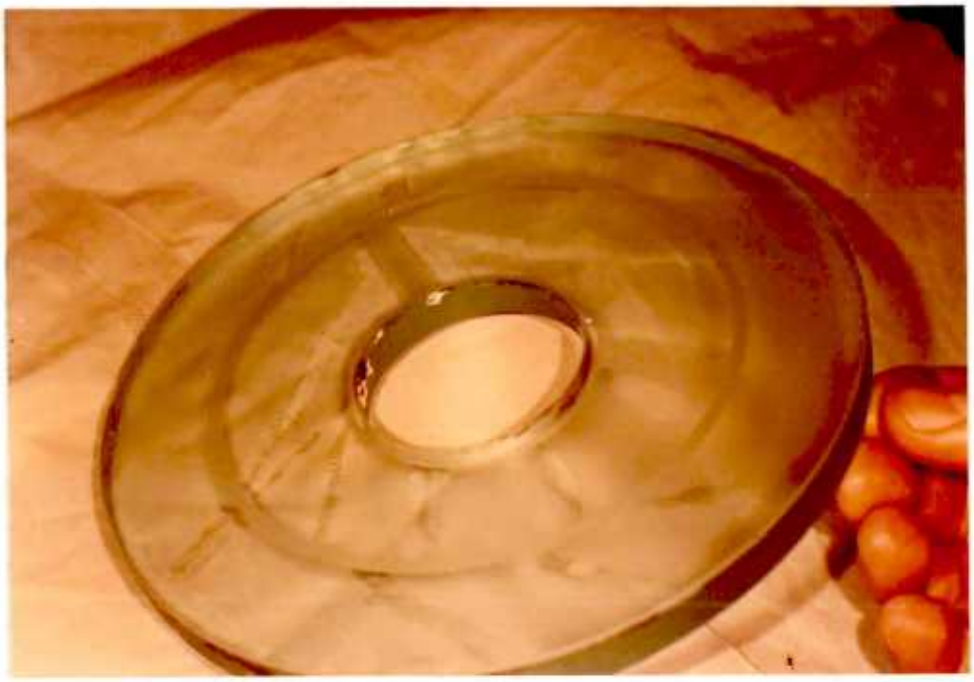


Fig. 6.5 Test disc with chemically etched annulus.

corresponding to the start and termination of contact resistance drops before further analysis is possible. The resistance measurement circuit is connected to a high frequency d.c. amplifier and thence to a trigger circuit which cleans up the signal to give sharp switching between two levels (Fig. 6.7). This signal is differentiated by a circuit with a short time constant to give a signal comprising alternate positive and negative pulses, idealised in Fig. 6.8. The diode circuit of Fig. 6.6 splits the signal into two parts, one consisting of positive pulses, the other of negative pulses only. The negative pulses are inverted and then both signals are fed into similar monostable circuits to give each pulse a fixed width.

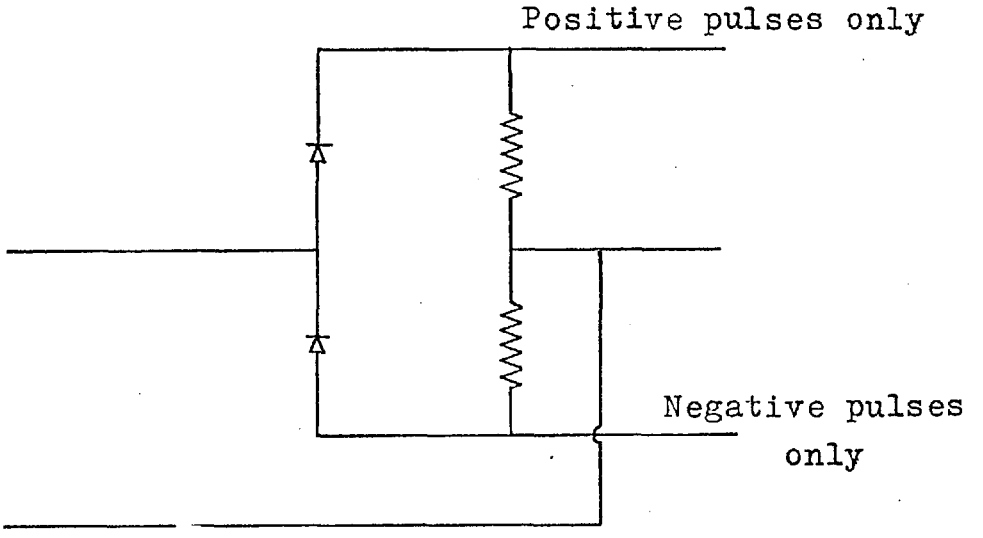


Fig. 6.6 Signal dividing circuit.

The width of these pulses can be altered simultaneously by means of a ganged potentiometer. The twin signals are illustrated in Fig. 6.9. The "A" signal corresponds to the onset of a contact and the "B" signal represents the

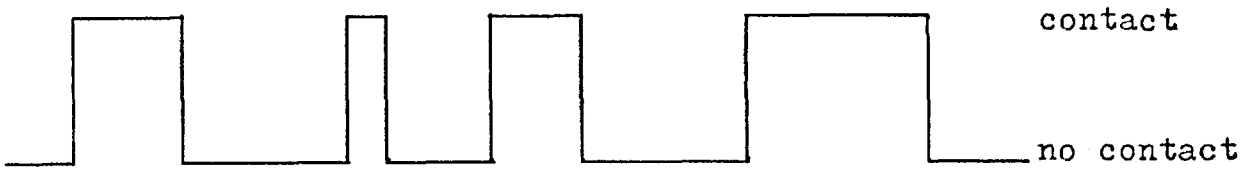


Fig. 6.7 Idealised waveform of asperity contacts.

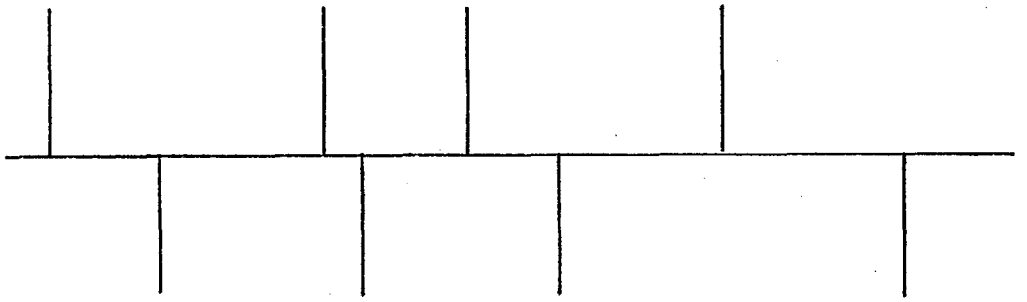


Fig. 6.8 Differentiated signal.

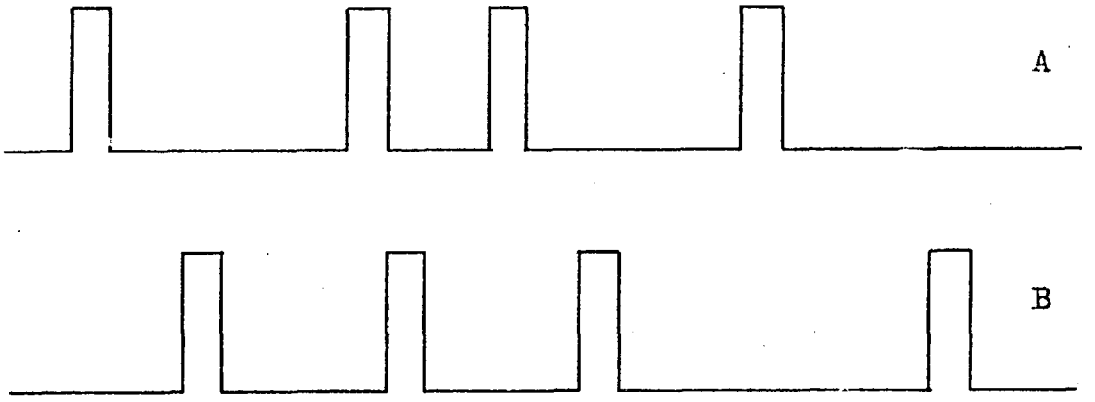


Fig. 6.9 Outputs A and B.

termination of a contact.

The most effective way to examine the spacing of these pulses is to use the signals to start and stop a high frequency oscillator connected to a pulse counter. These counts would have to be stored in a memory and subsequently collated. Unfortunately the sophisticated equipment necessary to perform these functions was not available. However, the main interest at the time was not in the complete distributions of dwell times and intervals but simply those values most commonly occurring which could be taken as representative of the contact conditions. Such information on dwell times can be obtained by cross correlating signal A with respect to signal B. Similarly predominant interval times will be revealed in the autocorrelation of signal A. A Hewlett-Packard correlator was available at the time of these experiments and this was used in the manner described above.

The width of the pulses is an important parameter in these experiments. When the pulse is long the frequency response of the system is considerably reduced. If the pulse is very short the power of the correlated signal is extremely small. Then small deviations superimposed on a constant recurring signal could render undetectable a peak in the correlated output. Some examples are described here to illustrate how the output varies with different inputs. EXAMPLE 1. Contacts occur randomly but persist for a fixed length of time τ. The pulse width in signals A and B is τ'. The cross correlation of A delayed with respect to B is illustrated in Fig. 6.10. If the pulse width τ' is reduced, the peak becomes narrower and shorter.

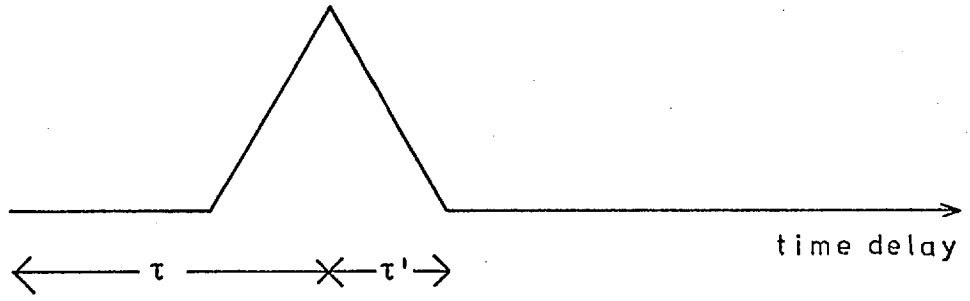


Fig. 6.10 Cross correlation function for example 1.

EXAMPLE 2. Contacts occur at constant intervals $\varkappa$ and persist for a constant time τ. The value of τ' is unchanged from example 1. The cross correlation of A with respect to B now shows peaks repeated at intervals of $\varkappa$. (Fig. 6.11.)

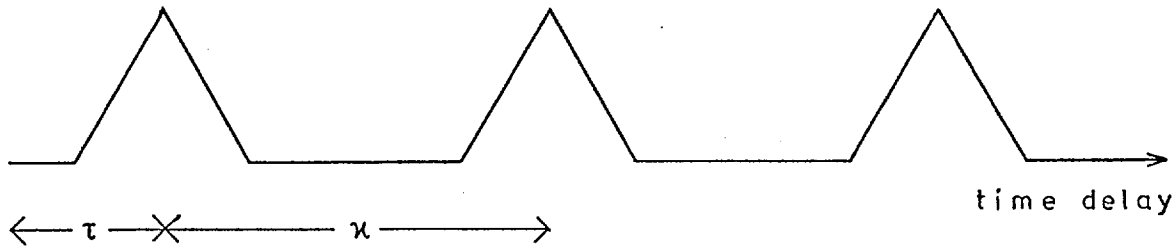


Fig. 6.11 Cross correlation function for example 2.

If the intervals between contacts have a small variation about $\varkappa$, successive peaks will be smaller and more diffuse.
EXAMPLE 3. Contacts occur at constant intervals $\varkappa$ and persist for a random length of time. The autocorrelation function of A is shown in Fig. 6.12. Essentially peaks are repeated at intervals of $\varkappa$ but each successive peak gradually diminishes,

partly because of the finite duration of the signal and partly because of dwell times greater than κ obscuring the next contact.

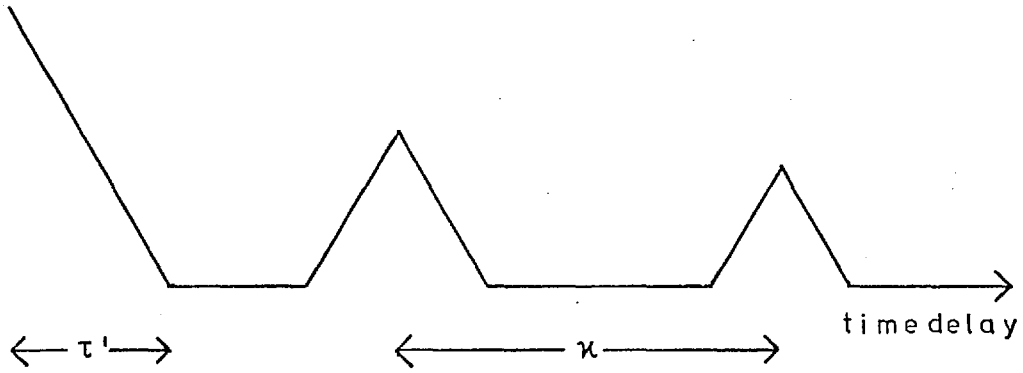


Fig. 6.12 Autocorrelation function for example 3.

6.4.2 Electrical circuitry.

The electrical circuitry is described in three stages. The first stage amplifies the signal from the line contact and then converts it into clean square waves. The second stage differentiates this waveform which is then divided into two signals, A and B. The B signal is inverted so that both waveforms consist of a series of positive pulses. Both A and B are fed into a similar third stage which gives each pulse a fixed width. The Schmidt trigger ensures fast switching between voltage levels. The pulse widths of the A and B signals may be varied together by a ganged potentiometer and the heights of the pulses are controlled by similar means.

The first stage circuit is shown in Fig. 6.13. A 1.5V dry cell powers a potential divider which supplies

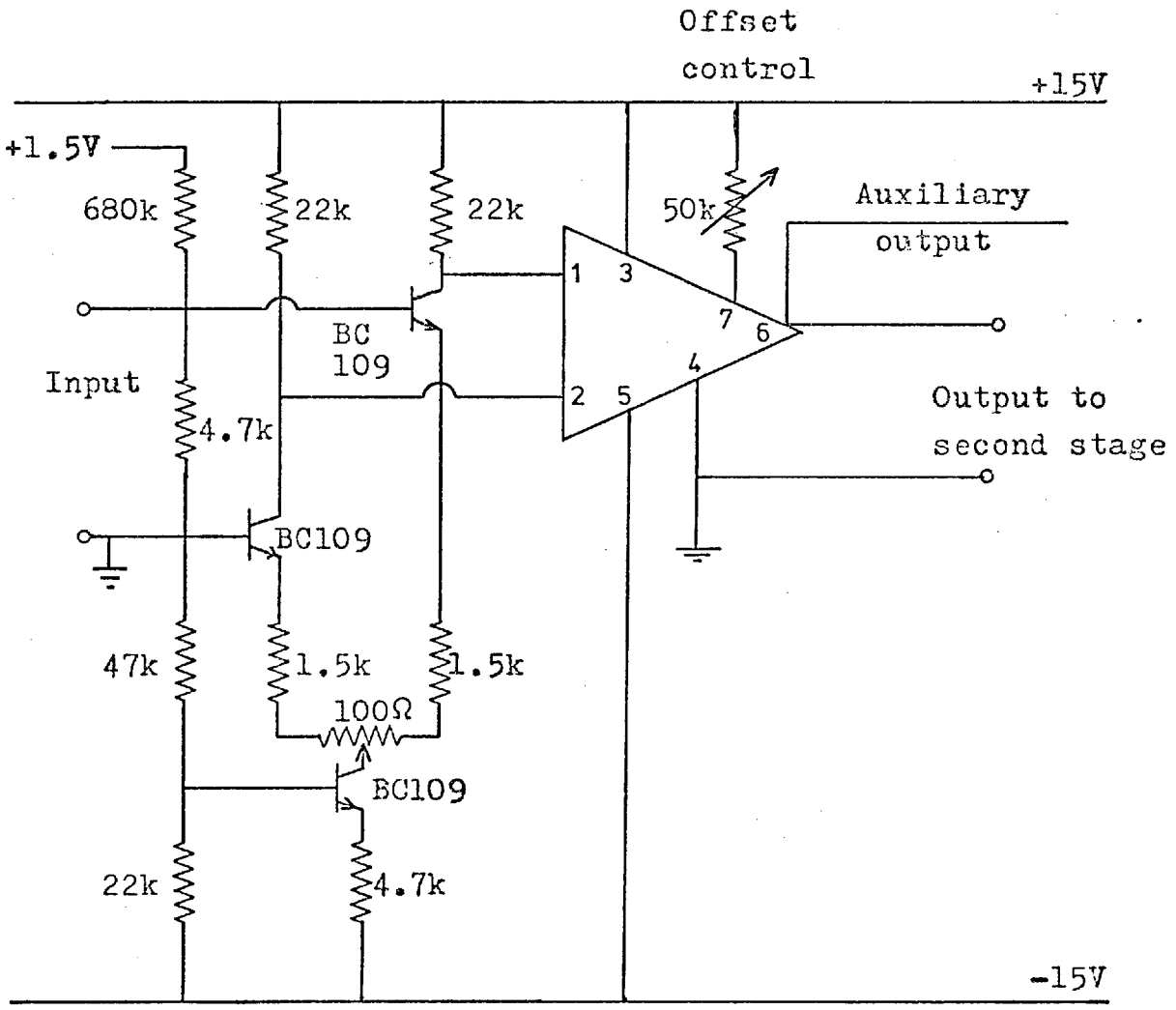


Fig. 6.13 First stage of signal processing.

10.3mV to the contact. This is amplified by a differential amplifier in a long-tailed configuration supplied from a constant current source. The advantage of this type of circuit is its high rejection of common mode signals. The two low noise transistors in the amplifier were matched from a large selection with the aid of a transistor curve tracer. The voltage comparator is a commercially available unit (Ancom 15A-1a) selected for its high slew rate and good frequency response. The two potentiometers were adjusted for good switching from d.c. to 100kHz. A contact registered when the electrical contact resistance fell below about 1000Ω. The output from the voltage comparator is fed into the second stage and an auxiliary output can be connected to a pulse counter.

The second stage is shown in Fig. 6.14. The workings of this circuit have been described above. The diodes used to divide the signal were of a fast germanium type. Two outputs are obtained from this network, each of which is connected to a third stage circuit, shown in Fig. 6.15. The first two transistors are connected into a monostable configuration. On receiving an input pulse, an output pulse is produced of fixed duration. The length of this pulse is predetermined by the 100kΩ variable resistor and can be set from 5 to 500μs.

The final pair of transistors are connected up as a Schmidt trigger which improves the switching characteristics of the circuit. The ganged potentiometer used to vary pulse widths was not of a high precision type and a small

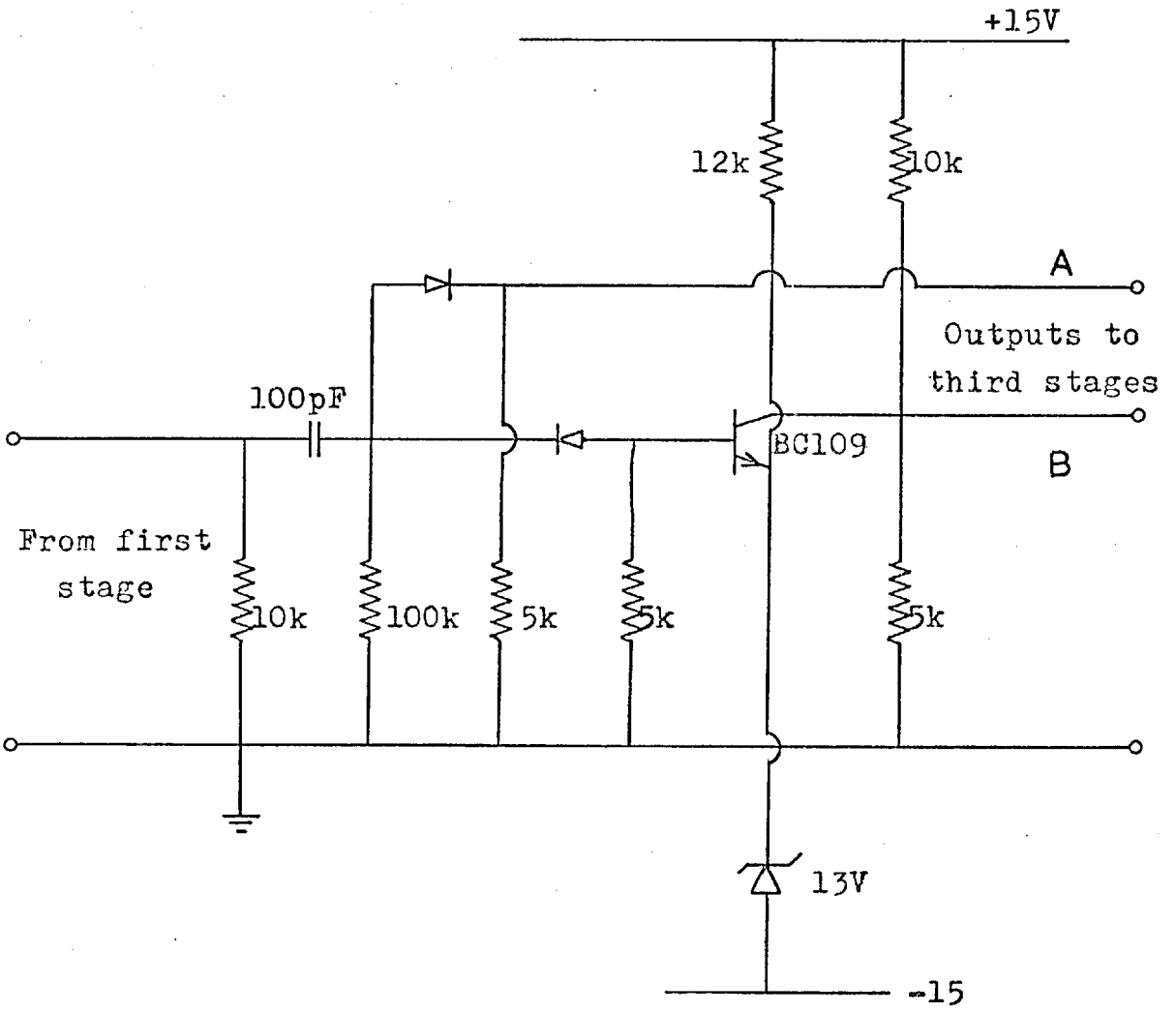


Fig. 6.14 Signal dividing and inverting circuit.

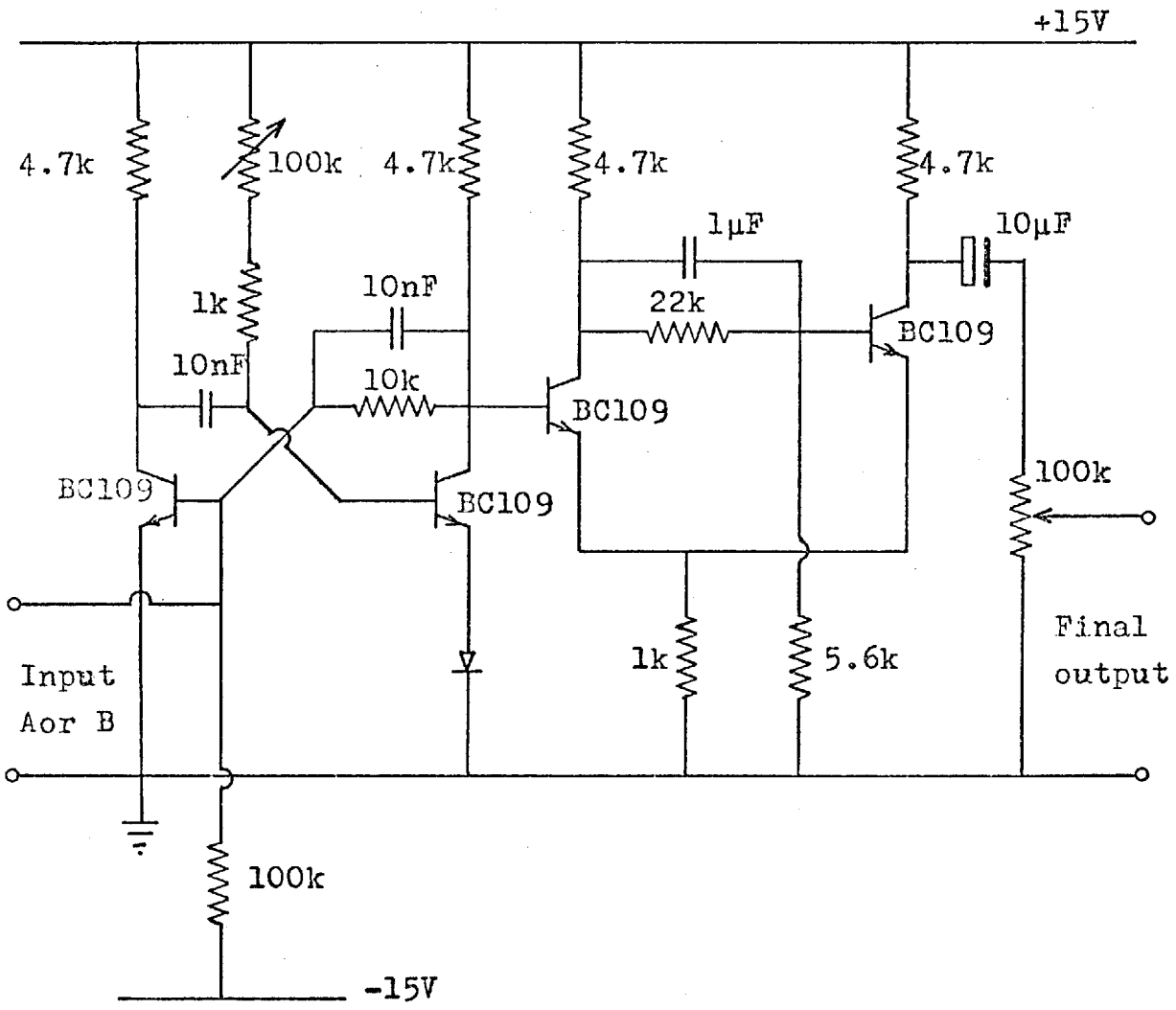


Fig. 6.15 Monostable circuit and Schmidt trigger.

discrepancy was observed when long pulse widths were employed.

Each circuit separately exhibited satisfactory behaviour up to 100kHz. The behaviour of the complete system connected to the correlator is described in chapter eight.

6.5 Summary.

A method has been described of searching for dominant periods in the duration of asperity contact dwell times and in the intervals between the onset of such contacts. Electrical interface equipment was designed and constructed in order to convert the signal into a form suitable for analysis. An existing tapered roller test machine was adapted to provide an essentially two dimensional contact.

Chapter 7. Vacuum deposited coatings.

7.1 Introduction.

7.1.1 Vacuum systems.

Most vacuum systems need to be pumped down to a pressure of at least 10^{-4} Torr (1 Torr = 1mm. of mercury $\simeq$ 10^{-3} atmospheric pressure) in order to ensure that air does not play a noticeable part in the deposition process. It is good practise to pump well below this for systems which deposit at a slow rate. The pumping is normally done in two stages. A mechanical rotary pump can reduce the system pressure from atmospheric to at best about 10^{-2} Torr. Before that stage is reached one would normally switch to an oil or mercury diffusion pump backed off by the rotary pump. Deposition may proceed when the residual pressure is low compared with the expected vapour pressure of the coating substance.

7.1.2 Measurement of deposited films.

Various methods are employed to measure the thickness of deposited material. These include optical methods, the microbalance, stylus profilometry and accurate weighing. Measurement of optical reflection or transmission has the advantage of being possible while the deposition is in process but is limited to thin and transparent films. Steps must be taken to prevent contamination where the light beam enters and leaves the vacuum chamber. Transmission measurements require that the substrate material and support system be transparent.

The frequency of oscillation of a quartz crystal

is utilised in the microbalance. The resonant frequency of the crystal depends on its total mass. From the resonance frequency change the mass accumulation can be calculated and hence the film thickness. This can be used as a continuous monitor during evaporation, but the electric fields associated with sputtering interfere with the sensitive electronic circuitry connected to the crystal.

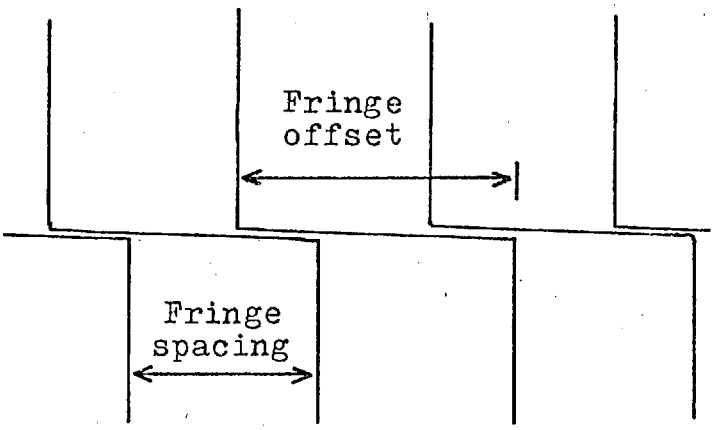
Weighing of a substrate before and after coating is a reasonably accurate means of assessing thickness. For best results a large substrate is necessary and any local variations in thickness are not detectable by this technique.

The following two methods require the production of an edge bounding the coating. This is accomplished by either masking the substrate prior to deposition, or by subsequently etching away part of the film.

Stylus profilometry can detect films 100Å thick but difficulties were encountered with gold films, which were scratched by the instrument head. For these measurements the Rank Taylor Hobson Talysurf IV was used.

The optical interference microscope is a very useful instrument for guaging film thickness. The particular model used was the Varian A-scope interferometer which has a resolution of about 40Å. In this device a monochromatic light beam reflected from the sample interferes with a beam reflected from a specially coated Fizeau plate. The test piece is at a slight angle to the Fizeau plate and this angle is varied to obtain the most suitable fringe spacing. This

spacing is compared with the fringe displacement at the step. Fig. 7.1 illustrates the process.



$$\text{Film thickness} = \frac{\text{Fringe offset}}{\text{Fringe spacing}} \times \text{half wavelength}$$

Fig. 7.1 Interferometric measurement of film thickness.

The production of a good edge is of paramount importance here, and care is necessary to avoid an edge with excess deposited material. The samples were coated with either gold or aluminium to make them sufficiently reflective to obtain good fringe visibility.

7.1.3 Coating techniques.

There are two main methods of vacuum coating, evaporation and sputtering, and a process known as ion plating which is a combination of both. Evaporation of metals is normally performed by suspending loops of evaporant in the form of wire from a coiled heating element. It is found necessary to "flash evaporate" alloys to preserve stoichiometry, as the partial pressures of their constituents are not necessarily in the desired ratios. This is done by dropping grains of evaporant on to a plate at a high temperature.

Non-metals and powders are normally placed in a conducting boat which is again heated by electrical current. For materials with a high melting point resistive heating is impractical and some alternative is necessary, usually an electron beam source.

The mean free path of a particle is normally longer than the source to substrate distance, so that evaporation is essentially a line of sight process.

Reactive evaporation is a useful method for depositing compounds. For instance, titanium dioxide can be deposited by evaporating titanium in an atmosphere of oxygen.

The adhesion of coatings depends very greatly on the cleanliness of the substrate but generally evaporated films are regarded as having poor adhesion at ambient temperature.

Sputtering relies on removal of material by momentum transfer rather than by thermal agitation. A d.c. sputtering system can only sputter conducting materials. The normal mode of operation of such a system is as follows. The system is pumped down to an acceptable pressure then back-filled with an inert gas (usually argon) to about 10^{-1} Torr. A negative potential is applied to the target and a glow discharge occurs between the target and the earthed substrate. Positively charged argon ions are accelerated into the target and knock out material which moves through the plasma. Some of these particles strike the substrate where they may remain. If a dielectric target were to be placed in the

system the negative charge on one side would be quickly neutralized. Thus the electric field would be concentrated inside the dielectric making further sputtering impossible.

For this reason radio frequency (r.f.) sputtering is needed. The dielectric target is exposed to a high frequency oscillation. When the signal goes positive, electrons are attracted towards the target and the vacuum side of the dielectric gains a negative charge. When the applied signal goes negative, the electric field reverses and positively charges ions are attracted to the target. If the alternating voltage were varying slowly enough the system would simply undergo this cycle of charge transfer. At the high frequencies used in sputtering there is not sufficient time for the more massive ions to neutralize the negative potential. The electrons are much more mobile and the target acquires a net negative bias. Thus the situation is now comparable with the d.c. system and the same sputtering process occurs.

A difficulty arises in the case of exchanging the insulating target for a conducting one. Now there is no net bias as the target simply follows the applied voltage. It is therefore necessary to include a blocking capacitor between the target and the power supply. This allows a conducting target to acquire a negative bias without changing the situation appreciably for an insulating target.

Conditions in the r.f. plasma are such that ionisation occurs more easily and it is possible to work at lower pressures than d.c. sputtering allows. Typically these

are in the range 10^{-3} to 10^{-2} Torr.

In order to obtain the most effective transfer of power into the plasma it is necessary to include a matching network between the power supply and the target. This normally takes the form shown in Fig. 7.2. The variable transformer T_1 and capacitor C_1 are adjusted so that a purely resistive impedance is presented to the power supply. Furthermore, this resistance should equal the output impedance of the supply. C_2 is the blocking capacitor.

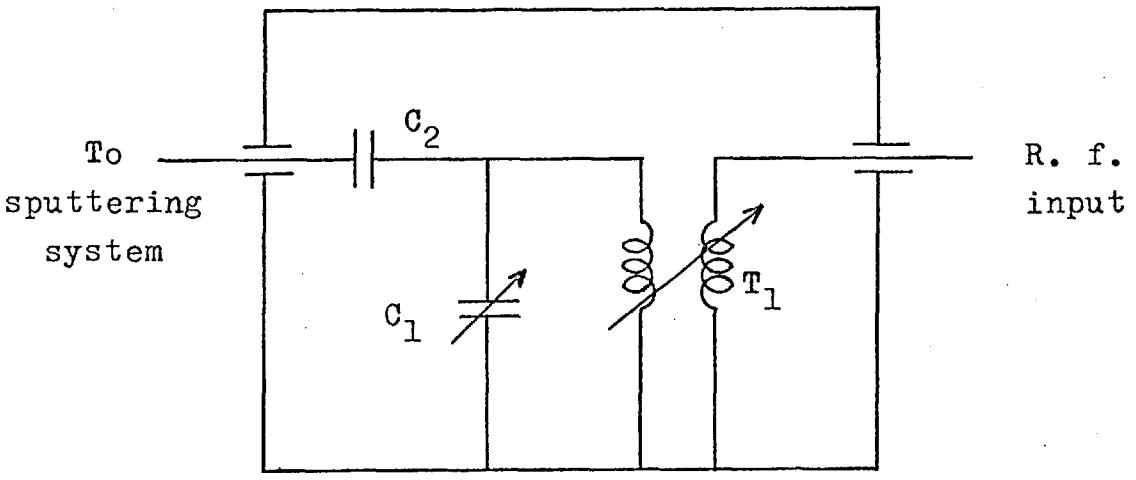


Fig. 7.2 R. f. matching network.

Sputter etching occurs when the electrical connections to a sputtering system are reversed so that the substrate comes under ionic bombardment. This is extremely useful as a cleaning process and as such is widely used.

Another variation is called bias sputtering, where a small negative bias is applied to the substrate. A combination of sputtering and sputter etching now occurs. If

the bias is too high, of course there can be no deposition.

The application of a magnetic field causes charged species moving in the plasma to follow a helical path, thus increasing the probability of ionising collisions and making the process more efficient. This is referred to as magnetron sputtering and there is a great deal of commercial interest in it.

As mentioned above ion plating is a hybrid process. The substrate is negatively biased and a plasma is induced. The coating material is evaporated and passes through the plasma before reaching the substrate. Ion plating combines the higher rates of deposition of evaporation with the good adhesion of sputtered films.

If an ion is accelerated through a much higher voltage than is used in sputtering, it tends to embed itself in the surface quite firmly. This is the basis of ion implantation which is used for doping semiconductors in integrated circuits.

7.1.4 Comparison of sputtered and evaporated films.

Generally sputtered coatings have good adhesion to the substrate whereas evaporated coatings are usually described as poor. This can be ascribed to two possible causes. Firstly the charged particles in the plasma may increase the likelihood of chemical action. Secondly the peak energy of sputtered species is considerably higher than that for evaporation. Fig. 7.3 shows a comparison between the velocity distributions corresponding to silver at 1800^{o}C

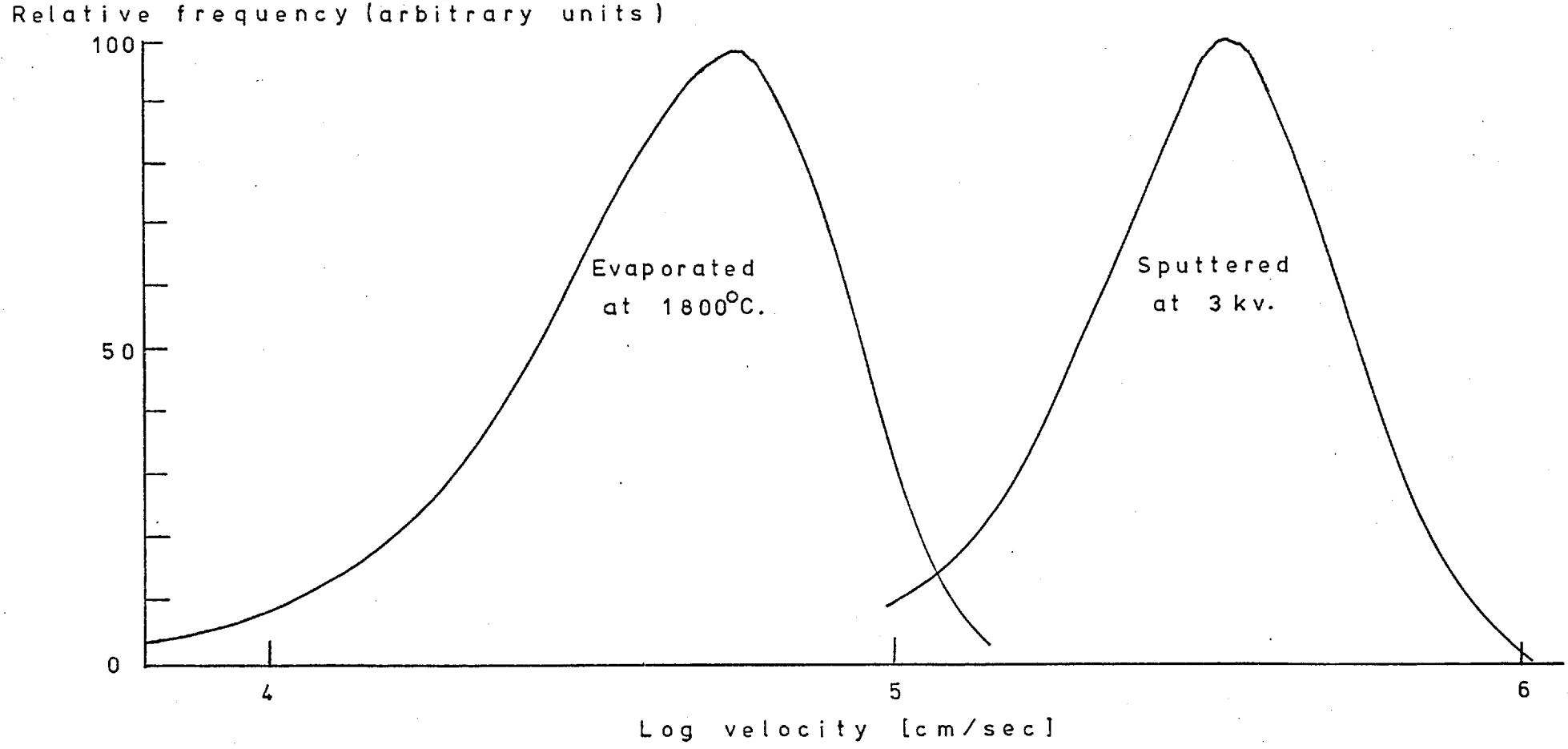


Fig. 7.3 Velocity distributions in sputtering and evaporation of silver.

and silver sputtered at 3kV. From the graph the ratio of velocity at the peak of the distribution is $10^{5.55}/10^{4.75} \simeq 6.3$. The corresponding ratio of energies is about 40 to 1.

Bias sputtering often improves the adhesion of coatings as weakly adhered material is likely to be re-sputtered or more firmly entrenched by "knock-on".

Sputtering is a very controllable process as deposition rates vary with applied power and are constant with time. This makes it most suitable for optical coatings.

One of the practical difficulties of sputtering patterns is to obtain a good edge. Particles moving through the plasma have random velocities so that material tends to be deposited underneath masks if they are very slightly raised above the substrate.

7.1.5 General.

More complete descriptions of surface coating processes may be found in Maissel and Glang (93) and Anderson and Chapman (94). The ion implantation process is not strictly surface coating. It has been mentioned, however, because of its possible importance in tribology in the surface hardening of metals and improvement in fatigue life (95 and 96).

The rest of this chapter is concerned with r.f. sputtering and sputter etching. Many of the difficulties encountered were mundane but it is emphasised that this technology is new, and most of the important developments in r.f. sputtering have taken place in the last decade (97).

7.2 Description of sputtering equipment.

The experimental work was done on a Materials Research Corporation sputtering module no. R.D.2467. This is shown in Fig. 7.4. Two targets may be in the machine at any time, and a rotating "J" platform allows any of the substrate tables to be positioned beneath either target. Fig. 7.5 shows a detail of the vacuum chamber. The three substrate tables are a heated stage, a water cooled bias or etch platform and an earthed plate usually used for pre-sputtering. This is necessary for cleaning the target prior to deposition and for raising it slowly to operating power without undue thermal shock. A rotating shutter shields the unused target during sputtering and the substrate during pre-sputter. Power was supplied from an M.R.C. r.f. power supply with a capability of 1.25kW.

The normal sequence of operations was as follows,

(i) etch substrate
(ii) pre-sputter to clean target
(iii) bring substrate into plasma
(iv) post-sputter if using high powers to cool target gradually.

The gas normally used for sputtering was argon. In a few experiments a mixture of 95% argon and 5% oxygen was used. When using pure argon the gas was passed through a hot titanium getter to remove the last traces of oxygen.

7.3 Optical coatings in lubrication research.

7.3.1 Introduction.

Prior to the use of surface coatings good

Fig. 7.4 General view of sputtering system.

Fig. 7.5 Vacuum chamber.

visibility of interferometric fringes depended upon the choice of a suitable combination of oil and moving element.

Kirk and Archard (98 and 99) studied the contact between crossed perspex cylinders. The fringes obtained were very feint as the refraction indices of perspex and oil are very similar.

Better results were obtained by Gohar and Cameron (100, 101, 102, 103 and 104) using materials with a high refractive index. These were diamond, sapphire and special glasses.

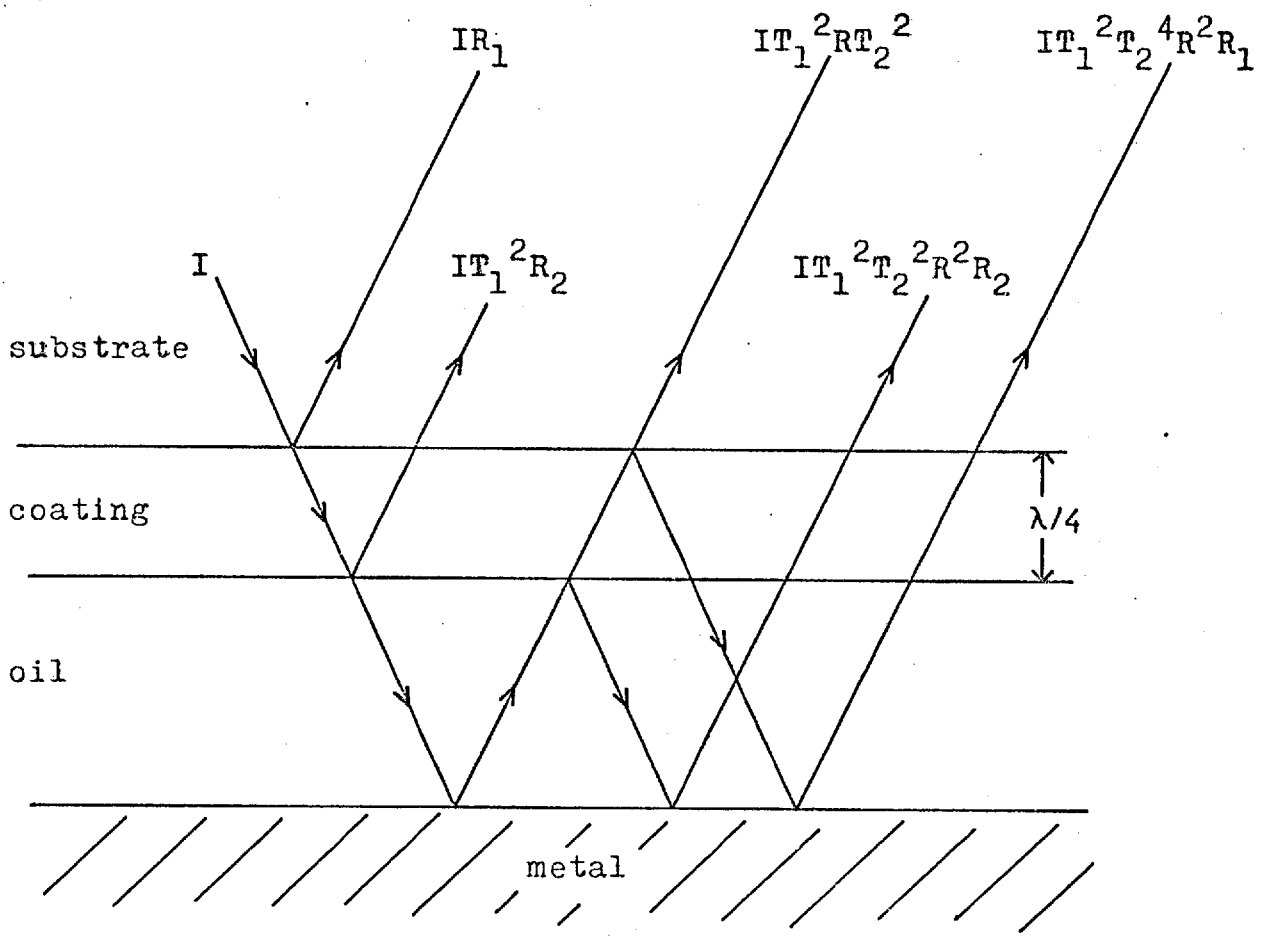
Much better visibility was obtained by Foord et al. (105, 106 and 107) who used a semi reflecting chromium coating in order to match the intensities of beams reflected from both oil interfaces.

Dowson and Jones (108) employed dielectric coatings in a study of lubricant entrapments. At the same time Westlake and Cameron (109, 110, 111, and 112), using a multiple beam interferometric arrangement, also adopted titanium dioxide coatings. Very clear fringes were obtained although some problems were encountered with the coatings peeling off under high loads (112).

7.3.2 Simple optics of dielectric coating for normal incidence.

Fig. 7.6 represents a substrate coated with a quarter wavelength thick dielectric coating. This is separated by an oil film from a metal roller.

A simple optical model is described for use in



R_1, T_1 represent the coefficients of reflection and transmission at the substrate/coating interface.

R_2, T_2 pertain to the coating/oil interface.

R is the coefficient of reflection of the metal surface.

Fig. 7.6 Ray diagram of simple optical model.

choosing suitable materials for obtaining the best visibility of fringes. All rays of magnitude R_1R_2 or higher order have been neglected.

The intensity of the reflected beam from a dielectric interface at normal incidence is given by

$$R = \left(\frac{\mu_1 - \mu_2}{\mu_1 + \mu_2}\right)^2.$$

If the incident ray is travelling from a rare medium to a denser medium there is a phase change of π radians on reflection.

The visibility V_s of fringes is defined as

$$\frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

where I_{max} and I_{min} have the obvious meanings. It can be seen that the maximum value of this function is unity in the case where the intensity drops to zero in the dark fringes.

In the situation above the largest possible value of I is $I_{coating} + I_{roller}$ and the smallest possible value is the difference between $I_{coating}$ and I_{roller}. Thus the visibility is given by

$$V_s = \frac{I_{coating}}{I_{roller}}$$

or its reciprocal, whichever is less than unity.

In order to find the coating which gives the clearest fringes it is necessary to compare $I_{coating}$ and

I_{roller}. In the following calculations the light source is assumed to be monochromatic and the effect of the finite bandwidth of the light is ignored. Also there is assumed to be negligible absorption in the dielectric media. Then we have

$$V = \frac{IR_1 + IT_1^2R_2}{IT_1^2T_2^2R(1 + RR_2 + T_2^2RR_1)}$$

Glass and oil have similar refractive indices of about 1.5 so we now put $R_1 = R_2 = R'$.

Then $$V = \frac{R'[1 + (1 - R')^2]}{R(1 - R')^4[1 + RR'(1 + (1 - R')^2)]}$$

Fig. 7.7 shows V plotted against R' for various values of R.

The highly polished steel specimens used in typical lubrication experiments have a value for R of about .69 (40). From Fig. 7.7 the best visibility is then obtained for $R' \simeq 0.207$ interpolating between the R = 0.6 and R = 0.7 curves.

Thus $$\left(\frac{\mu_1 - \mu_2}{\mu_1 + \mu_2}\right)^2 = 0.207.$$

If $\mu_2 \simeq 1.5$ then $\mu_1 \simeq 1.5 \frac{1 + \sqrt{.207}}{1 - \sqrt{.207}} \simeq 4.00$.

No common material has a refractive index as large as this (except germanium in the infra red) and a compromise is necessary. Titanium dioxide and chromic oxide (Cr_2O_3) have refractive indices of 2.65 and 2.55 respectively. The corresponding values of R' are 0.077 and 0.067 and these values yield visibility of .25 and .21 respectively. A compound coating could be used to increase the visibility

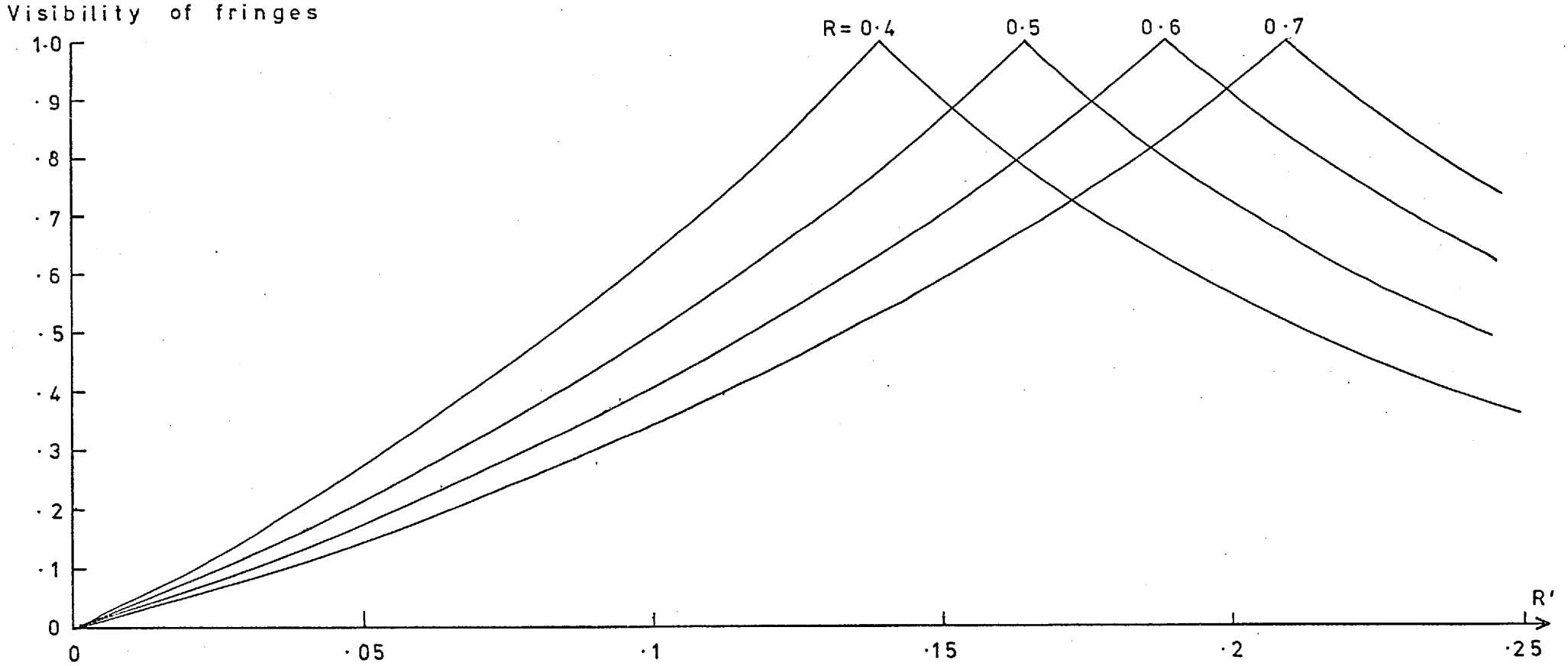


Fig. 7.7 Visibility of fringes as R and R′ vary.

of fringes but would increase the likelihood of mechanical failure. As some work on TiO_2 has been reported, the optical coating used in this work is Cr_2O_3 and a sputtering target was fabricated by the arc plasma spraying process.

7.3.3 Practical details of sputtering Cr_2O_3.

This material was deposited at the powers of 300W to 400W in an argon or argon/oxygen atmosphere. There was no observable difference in the quality of coatings deposited under this range of conditions and rates of deposition of 3000Å to 7,500Å per hour were obtained. Thin coatings had a slight yellowish colouration and thicker coatings displayed the green colour of the bulk material. Adhesion was generally good but for best results it was necessary to sputter etch the substrate, particularly in the case of sapphire. Prior to being placed inside the vacuum system samples were degreased in toluene and washed in acetone.

The maximum diameter of specimen that could be fitted inside the vacuum system was about 6½". The coating was not entirely even, being thickest in the centre and thinnest at the periphery. The variation was less than 15% up to 4½" diameter and rather more at greater diameters. This was not a serious problem in practice as the thickness at any radius was constant so that the deposition time could be varied to give the correct value for any particular working distance.

Allchurch (113) was the first to use a semi-reflecting chromic oxide coating in lubrication research.

The substrate was a grooved sapphire disc used to study spinning conformal contacts. Subsequently the coating was applied to many types of test specimens including sapphire windows fitted inside a roller bearing race (114) and a tapered roller bearing race. In all these test arrangements the coating stood up remarkably well. Perhaps the most severe test of its adhesion was performed by Hedley (115) who ran a coated sapphire disc loaded up to 300,000 p.s.i. with 1% traction. In this test machine the sapphire cracked with the coating still intact.

The excellent performance of this coating (far in excess of the author's own modest requirements) has played an important part in extending the working range of optical EHL test machines.

7.4 Sputter etching of disc.

Certain difficulties arise when masking substrates prior to sputter deposition. The mask must be in intimate contact with the substrate, otherwise sputtered material will undercut the mask and a diffuse edge will result. The same problem occurs when sputter etching a pattern. Precision masks can be constructed of thin metal foil but these distort under ionic bombardment because of the heat generated. To circumvent this problem, two masks were constructed from a single piece of duralumin of half-inch thickness. They were arranged to cover a six inch diameter circle except for an annulus of $4\frac{1}{2}$ inches diameter and .1 inch in width. A thin strip was cut out along a radius of the inner mask to provide for an electrical connection from the annulus to the

centre of the disc. In order to reduce shadowing of the substrate the masks were tapered towards the annular region. A spigot was also constructed to centralise the inner mask on the glass test disc. The bases of the masks were finished by lapping, to obtain the best possible contact with the very flat glass substrate. Fig. 7.8 shows a cross-section through the assembly. Masks as bulky as those described above are rarely used in vacuum systems and practical problems arose from their large charge and thermal capacity.

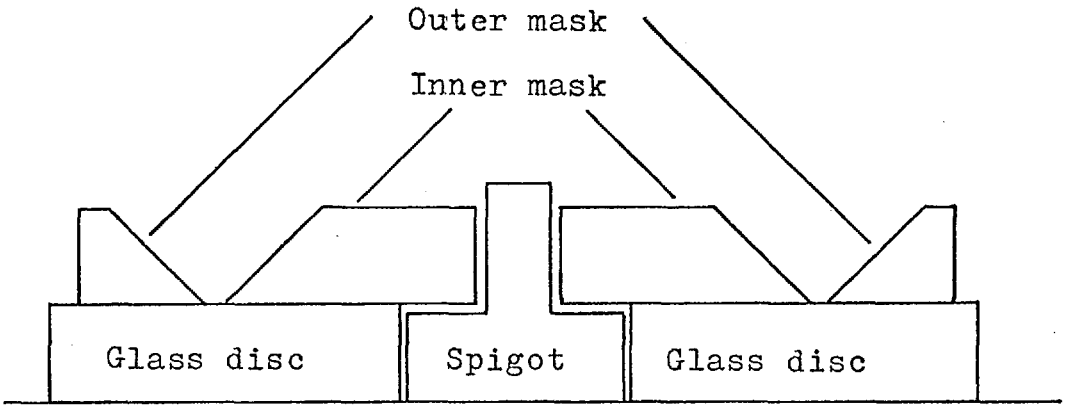


Fig. 7.8 Cross-section through mask and substrate assembly.

At powers below about 200W no etching of the substrate was discernible. At the highest power (500W) a rate of 150Å/min was obtained. It was noted that the dark space dipped over the annular region at high powers. After a short time of etching the plasma became unstable and it was increasingly difficult to match the system satisfactorily. This was ascribed to charge storage on the masks. The first disc was cracked when a discharge occurred between the

substrate platform and a metal lug used to centralise the outer mask. Subsequently the outer mask was aligned by eye.

Several trials followed, each of which was terminated by the shattering of the glass substrate. Annealing of the discs did not alleviate the problem. The masks required about an hour to cool sufficiently to be handled. Clearly this was a thermal problem, either thermal shock due to the application of high power to one surface, or simply too high a temperature gradient between the hot masks and the cooler substrate platform. In an attempt to differentiate between these two effects a second disc was placed beneath the test disc. It was argued that if the trouble arose because of the rapid heating of the top surface, only the top disc would be damaged. In fact it was the bottom disc only that shattered, indicating that the cool metal support platform was at least partially responsible. In all further trials the test disc was raised above the metal platform by small ceramic washers.

When, at last, specimens could be recovered intact the etching was found to be uneven, presumably because the outside mask assumed a positive charge and repelled some of the bombarding ions. The inside mask was earthed by the centre spigot and thus remained electrically neutral. Further trials were performed without the centre spigot so that both masks could assume the same potential. A side effect of this change was to focus the plasma through the hole in the centre of the inner mask. A ceramic bead, inadvertently left on the platform beneath this hole, glowed orange when 500W of power was applied and this gives some indication of

the local power concentration. It was now possible to etch an even annulus into the disc, although at somewhat lower rates than previously obtained.

The problem with matching the power still occurred after about an hour of continuous running. Attempts to rectify this had previously been disastrous and the mismatch was taken as a warning sign. The sputtering process was halted for a length of time and subsequently resumed. Rather than run at full power continuously the system was held at maximum power for a minute and then at low power for four minutes. This process worked satisfactorily.

7.5 Sputtered chromium and gold coatings.

No problems were experienced with the adhesion of chromium coatings which were sputtered in argon at 400W power. Gold sputtered directly on to glass had rather poor adhesion and could be removed with adhesive tape. The introduction of a 100W bias on the substrate platform led to much better adhesion and a deposition rate some 55% of the normal at 200W power. Gold coatings several thousands of Angstroms in thickness could not be removed with adhesive tape when bias sputtered. In order to obtain the best possible adhesion the gold coating was bias sputtered on the final test discs.

The three operations of etching, chromium deposition and gold deposition were performed without letting the system up to air and Fig. 7.9 shows the completed specimens.

Fig. 7.9 Two completed test discs.

7.6 Summary.

A number of untypical practical problems were encountered during sputter etching of the glass discs. A disc was prepared and the gold coatings survived for about ten hours of running time under light loads.

The chromic oxide dielectric coating, although not so important to the author's work, has proved extremely successful in optical EHL research.

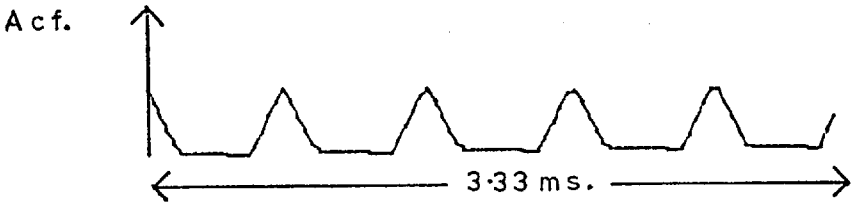
Chapter 8. Results of correlation tests.

8.1 Testing of equipment.

Each of the circuits described in chapter six was tested for frequency response and was found to be satisfactory up to about 100kHz. At frequencies higher than this the system still worked but the amplification in the first stage was reduced. This caused changes in the switching levels so that the resistance of a high frequency contact would have to be less than the normal 1kΩ level in order to register. The system was assembled and connected to the correlator and the response to an applied square wave was measured. The output of channel A was autocorrelated and the result is shown in Fig. 8.1. Strictly, the output should be a succession of dots as the correlator functions on digital principles. No facility was available for automatically raising the pen from the plotting table and the dots were connected.

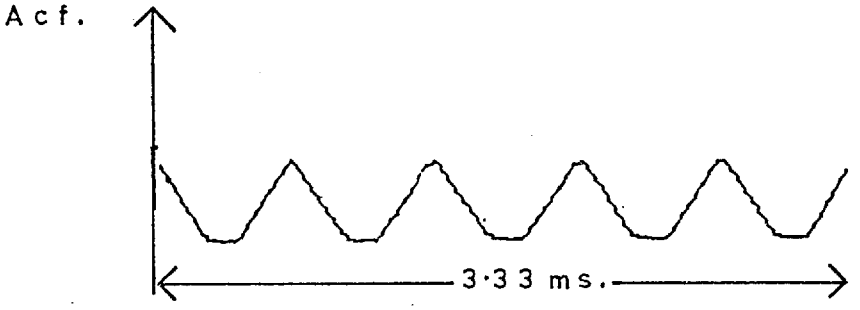
For Fig. 8.2 the output pulse was lengthened and it can be seen that the peaks are higher and wider. These results may be compared with Fig. 6.12. The reduction in the heights of successive peaks referred to in section 6.4.1 is hardly perceptible as the sampling time was long compared with the pulse width.

Fig. 8.3 shows the final output when the A and B channels are cross correlated. The input was again a square wave. The slight flattening of the peaks of the cross correlation function illustrates slight mismatch between the pulse widths from channels A and B. This Fig. may be compared with Fig. 6.10.



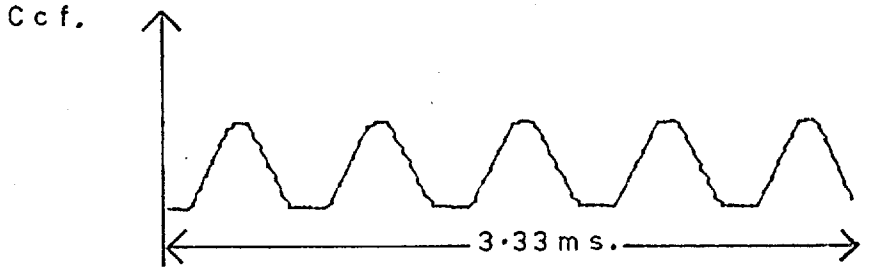
Test frequency 1.5kHz. Pulse width 150μs.

Fig. 8.1 Autocorrelation test with narrow pulse.



Test frequency 1.5kHz. Pulse width 250μs.

Fig. 8.2 Autocorrelation test with wide pulse.



Test frequency 1.5kHz. Pulse width 250μs.

Fig. 8.3 Cross correlation test with wide pulse.

8.2 Slip measurement.

Practical difficulties make the accurate measurement of roller speed impossible. A measure of total slip can be achieved by comparing the speed of the test disc with that calculated from the driving flange. An attempt to do this directly was made by sputtering a hundred chromium strips around the edge of a test disc. A coaxial light source and photocell was directed towards these strips, and the output was taken together with the output from the driving shaft speed transducer to the electronic tachometer. This device compares two signals and finds their ratio. This arrangement did not work reliably, probably because of the thin covering of oil from the hydrostatic bearing on the chromium coating. An estimate of slip was then obtained in the following manner. The disc was allowed to run for a large number of counted revolutions. The output from the tachometer was recorded for this period. The number of disc revolutions was compared with that calculated from the average of the measured driving flange speeds. Agreement was always better than half of one percent. This does not obviate the possibility that instantaneous sliding could have occurred but this was only observed on a large scale when starting up the system. Using a pulley driven system on the test disc, Wymer found that a wide range of mineral oils gave less than one percent slip. Although slip measurement is extremely important in asperity contact research, it appeared that the approximate method described above would be adequate, considering the small Hertzian width employed in these experiments, and the criterion developed in section 4.8.

8.3 Experimental procedure.

The test disc, roller and driving plate were washed in toluene to remove oils and then rinsed in acetone. About ten cubic centimetres of test oil were injected by means of a syringe onto the driving plate and roller and the system was assembled. The oil pump was started and the load adjusted by an air pressure control valve. The system was slowly brought up to speed which was adjusted to give a recognisable interferometric fringe colour. After a short time optical measurements became difficult because of oil contamination inside the piston assembly. However, no great importance was attached to optical methods in the early stages as the first requirement was to establish statistical parameters of significance under any operating conditions.

Load values were calculated from a calibration by Wymer (40), using recess pressure measurements of the hydrostatic bearing.

The contact half width, b, was calculated from Hertz formula (116)

$$b = \sqrt{\frac{8WR}{\pi E'L}}$$

where W is the load, L the length of the contact, R the mean radius of the roller and E' the reduced elastic modulus (defined on page 59) for the glass and steel combination. For this system

$$E' = 1.198\times10^{11}\,N/m^2 \quad (17.38\times10^{6}\,lb/in^2)$$

and

$$R = 4.11\times10^{-3}\,m \quad (0.1618in)$$

Speed and temperature measurements were recorded and the electrical signal was then studied. The variable electrical parameters were the duration of the pulses produced by channels A and B and the time between samples on the correlator. This sampling time was varied between 3.33μs. and 100ms. Finally, in order that the results be sufficiently representative of the test conditions, the correlator was set to display the average of a number of sampling periods, usually one thousand.

8.4 Results of cross correlation analysis.

The actual power present in the output signal depends on the degree of correlation and also the input pulse widths from channels A and B. It was found that the output power was quite low compared with the test signals shown in Figs. 8.1 to 8.3. Fig. 8.4, for example, has a vertical magnification of one hundred compared with the earlier Figs. In subsequent Figs. where the initial peak is suppressed, the vertical magnification is typically a further four or five orders of magnitude.

Fig. 8.4 shows a characteristic output of the cross correlation analysis. The initial peak decays away and no comparable peaks follow. This indicates that a large number of contact dwell times are shorter than the pulse width. Fig. 8.5 has a more irregular "tail" and reflects the effect of reducing the number of samples. In order to examine the correlation after the initial peak the amplification was considerably increased. In subsequent Figs. the initial peak is off the scale. Figs. 8.6,8.7 and 8.8 are typical

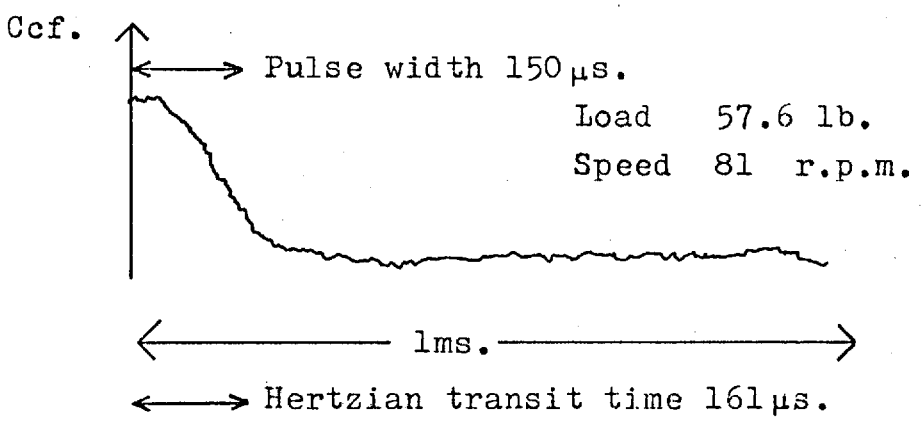


Fig. 8.4 Cross correlation analysis.

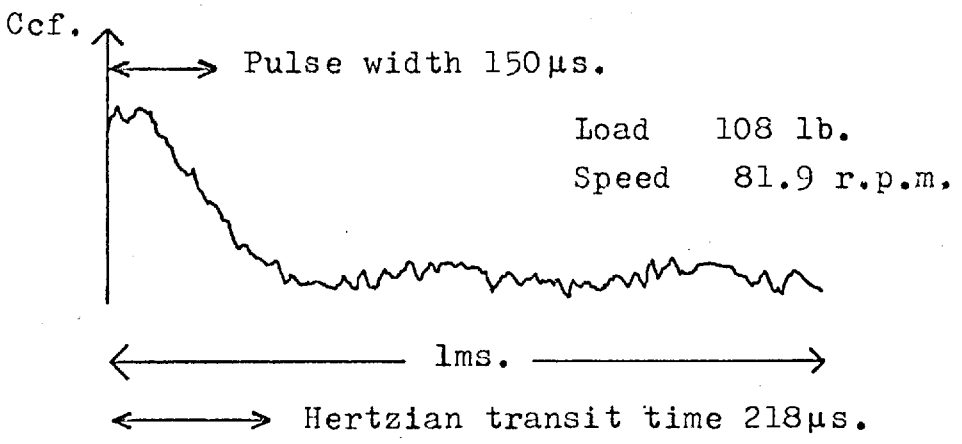


Fig. 8.5 Cross correlation analysis.

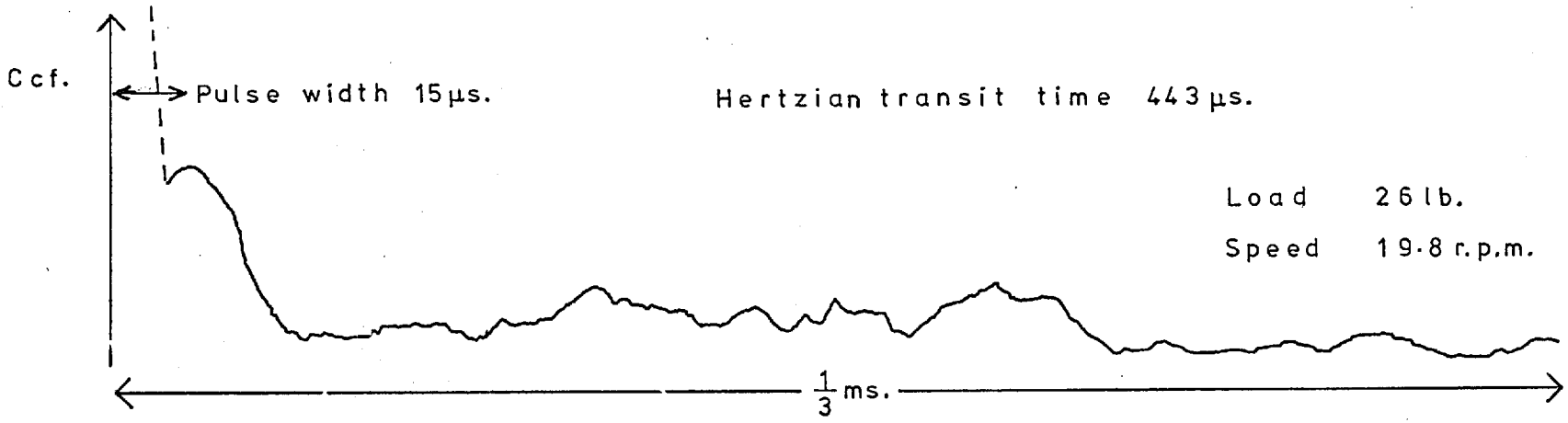


Fig. 8.6 Cross correlation analysis.

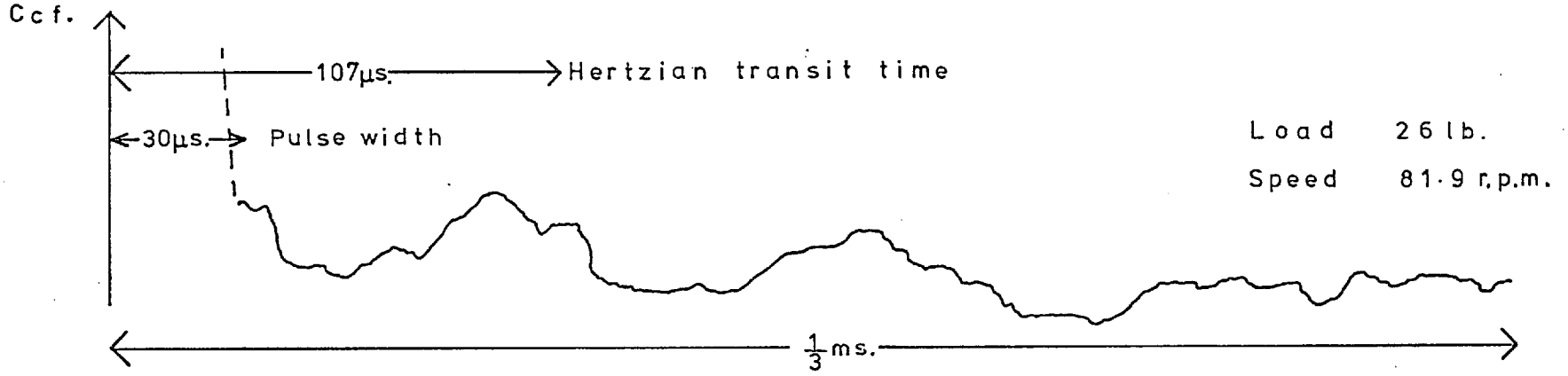


Fig. 8.7 Cross correlation analysis.

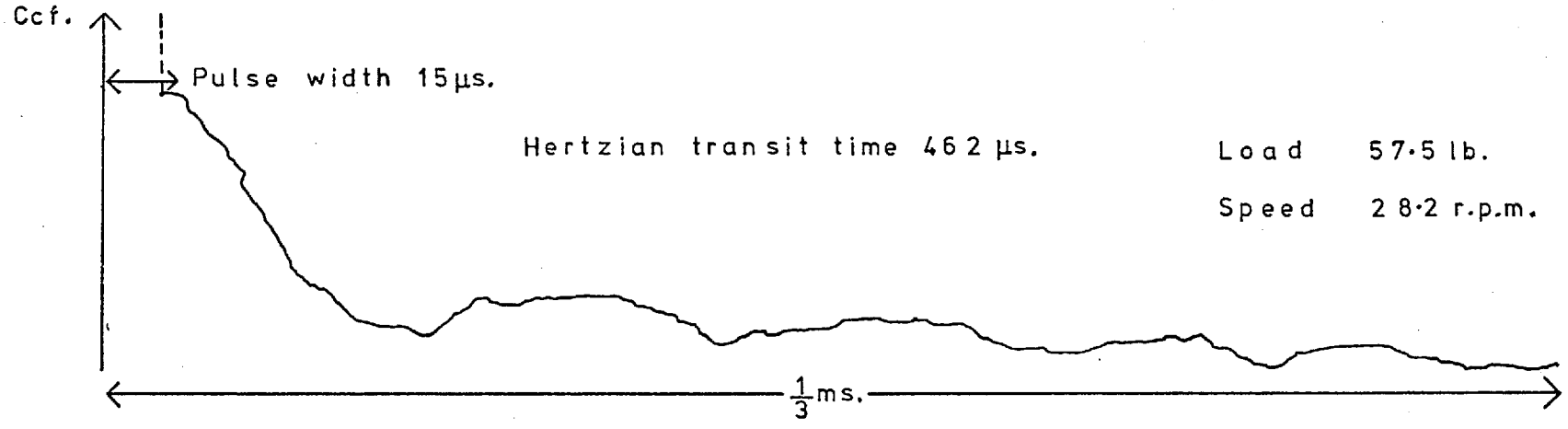


Fig. 8.8 Cross correlation analysis.

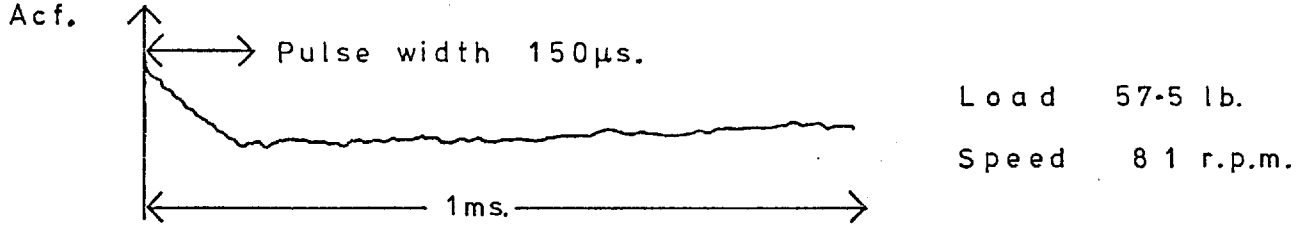


Fig. 8.9 Autocorrelation analysis.

of results obtained for the range of conditions investigated with short pulse durations and sampling intervals. From these results there is no evidence to suggest any correspondence between Hertzian transit time and asperity contact dwell time.

8.5 Results of autocorrelation analysis.

Fig. 8.9 shows the behaviour that was found for the autocorrelation experiments. The decay time for the initial peak corresponds exactly with the pulse duration. This must be interpreted as indicating random behaviour over a wide range of delay times although there may be dominant frequencies with a period very short compared with the pulse width. This possibility was investigated for a range of conditions and the results are shown in Figs. 8.10 to 8.13. Some periodic behaviour is evinced but appears to be related to the pulse width rather than load or speed.

A few experiments were performed with direct autocorrelation of the electrical resistance signal obtained from the auxiliary output (Fig.6.13). It was expected that some periodic component would appear on a suitable time scale because of the short periods of higher contact that occurred when the roller ran over the gold band connecting the conducting annulus to the centre of the disc. This band did not interfere to any large extent with the other measurements as it represented only about 1½% of the working perimeter. The period of the peaks in Fig. 8.14 is very close to the time for one revolution of the test disc. Fig. 8.15 shows the second half of the first period on an expanded time scale. There is no obvious component representing the

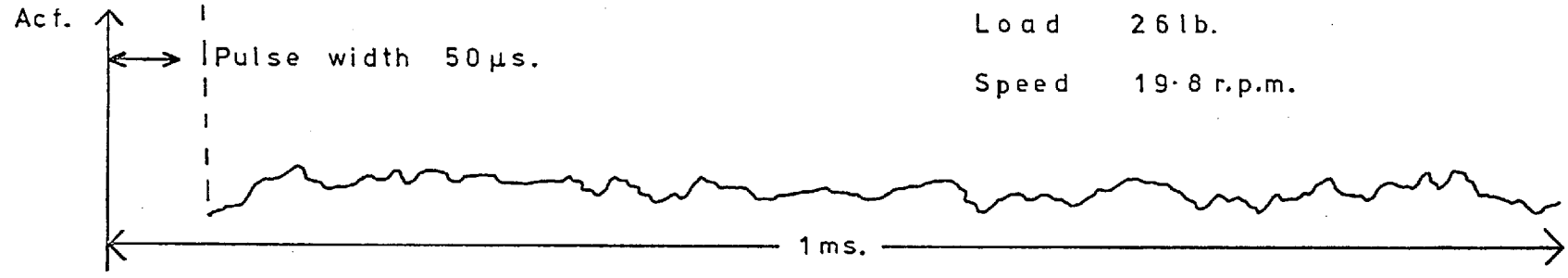


Fig. 8.10 Autocorrelation analysis.

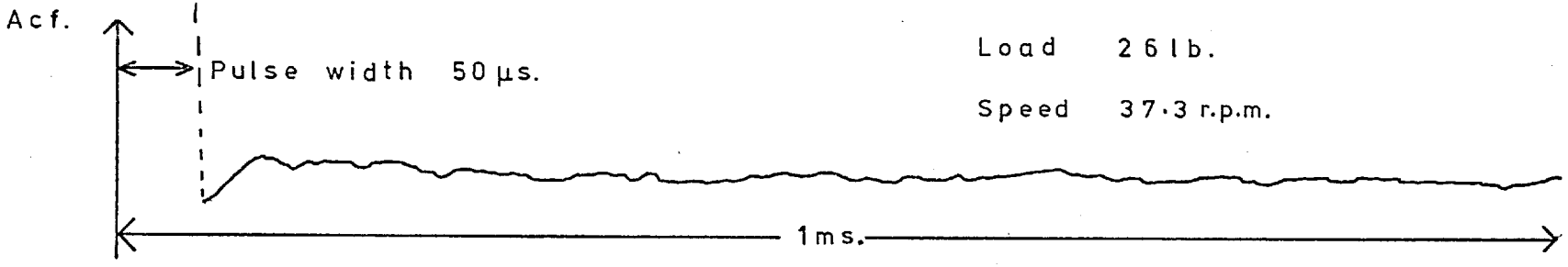


Fig. 8.11 Autocorrelation analysis.

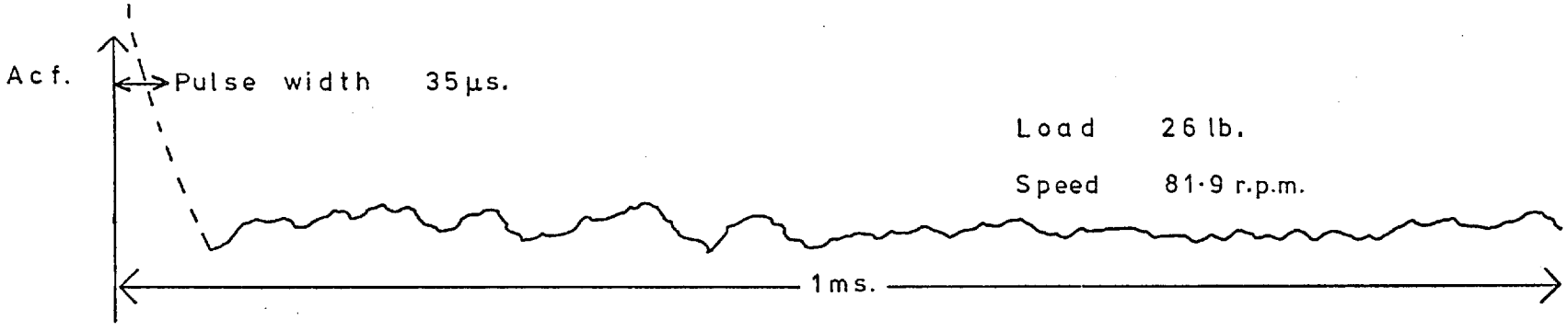


Fig.8.12 Autocorrelation analysis.

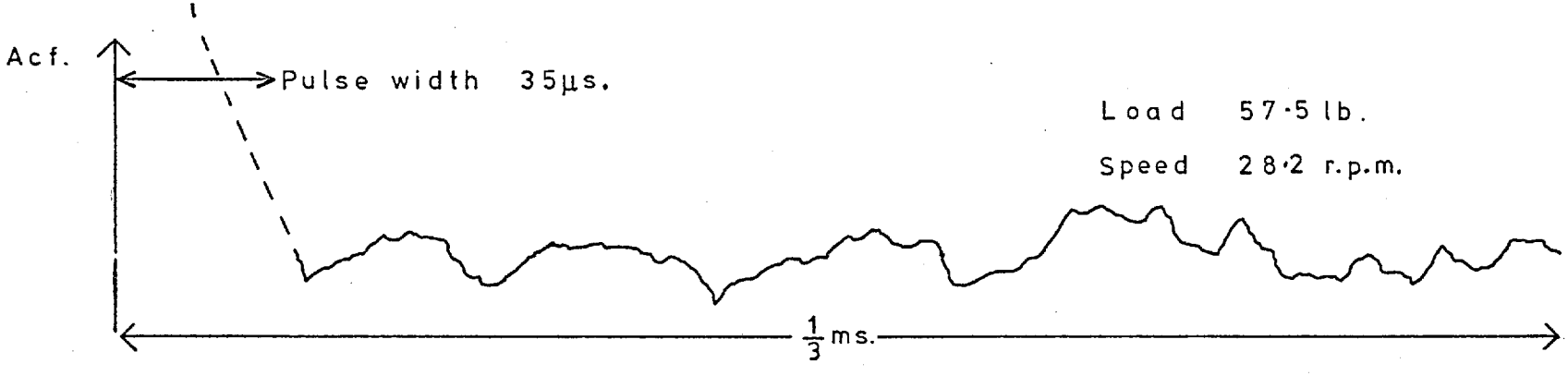


Fig.8.13 Autocorrelation analysis.

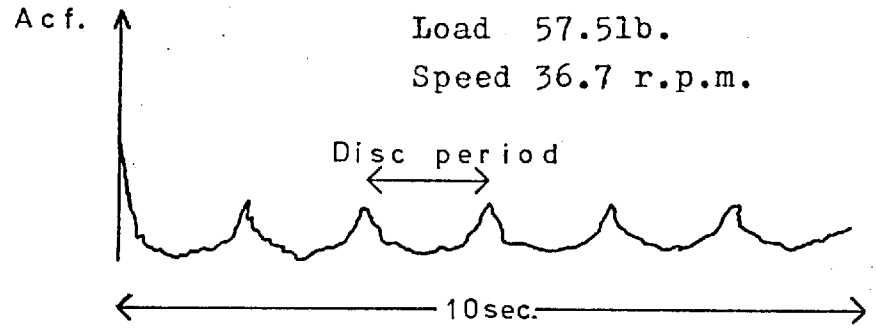


Fig. 8.14 Direct autocorrelation of electrical resistance signal.

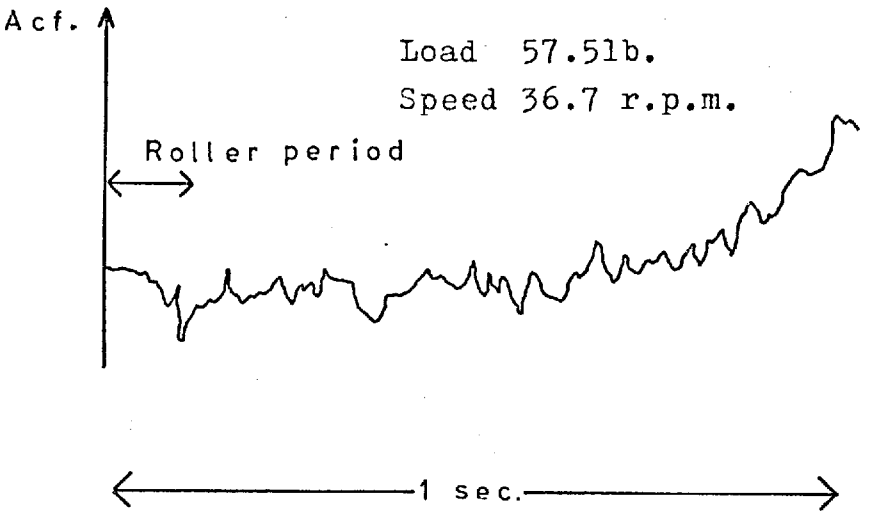


Fig. 8.15 Direct autocorrelation of electrical resistance signal.

time of one revolution of the roller.

8.6 General.

Incidental measurements of the count rate were made by means of the auxiliary output and rates of up to 7,700 counts per sec. were observed. The variation of count rates under the same nominal conditions was less than for the disc machine experiments reported in chapter five. A subjective estimate of the observed variation is about 15% between the highest and lowest values.

The test disc was used for about ten hours of running time. Under low loads a small amount of wear of the gold film took place as small flakes of gold were found on the rubber gasket when cleaning the system. This deterioration became more marked as the load was increased.

8.7 Discussion and conclusion.

Several criticisms must be made of the foregoing experimental work. Continuous monitoring of all the experimental conditions was extremely difficult as it was necessary to record many parameters by hand. Where possible it is far better to record automatically the electrical resistance signal and subsequently several methods of analysing data may then be employed.

The idea of annulling the end constrictions of the roller and studying a simplified contact may or may not be justified in a practical context. It may be that dominant frequencies are more apparent at the ends of the roller.

It is not thought that the gold wear particles

exerted a significant effect on the electrical signal. After the end of a short experiment under higher load most of the gold was found on the rubber oil seal and very little inside the contact itself.

A problem with the analytical methods employed in this work is that the answer has to be guessed approximately before performing the experiment, because the location of important frequencies is extremely dependent on the time scale chosen to examine the signal. Apart from the direct autocorrelation of the electrical contact signal, no dominant frequencies were observed. This suggests that both the dwell times of contacts and the intervals between the occurrence of contacts exhibit random behaviour, at least in the ranges examined. This supports the results of Christensen (48) shown in Fig. 2.13, that for nominal rolling line contact, the dwell times of electrical contacts are widely spread. This must compromise the common assumption of flat parallel Hertzian areas and one speculates as to the meaning of pure rolling conditions with rough surfaces in a practical context.

In retrospect, the lack of any dominant frequencies related to the onset of contacts is perhaps not surprising. For example, consider a series of equally high peaks on a lubricated rolling surface, in a straight line parallel to the direction of rolling. These peaks are nearly equispaced at interval k but small variations in separation occur. Then the distribution of separations may be of the form shown in Fig. 8.16.

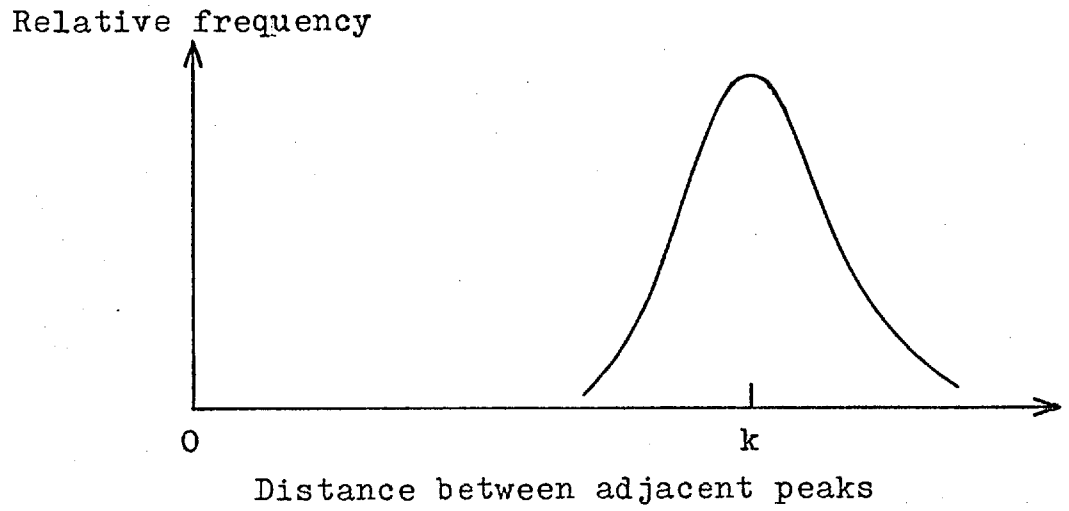


Fig. 8.16 Separation of peaks for model surface.

A similar train of peaks is now added, adjacent to the first, but on average 180 degrees out of phase. An element rolls against this surface and is short compared with k in the rolling direction and wide in the transverse direction. Then one expects the mean distance between contacts to be $\frac{k}{2}$, but a second random process is involved and the variance of the intervals must be doubled. This behaviour is indicated in Fig. 8.17.

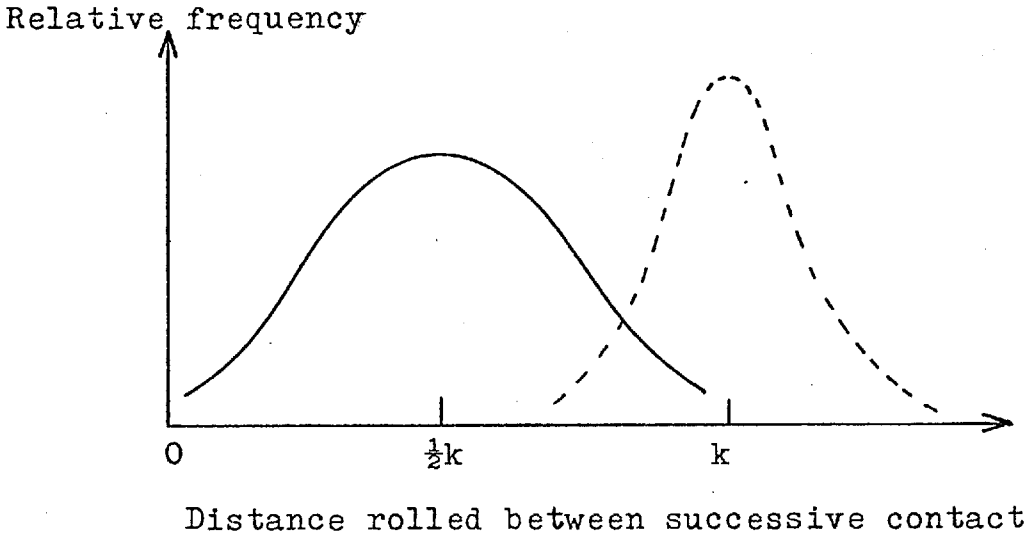


Fig. 8.17 Distance between contacts for double row of peaks.

If this process of adding extra peaks is increased the distribution will increase in deviation and the most commonly occurring separation will tend towards the origin.

If it is true, under a given set of conditions, that both the contact dwell time and the intervals between contacts are randomly distributed, then it is possible that a basis exists for reconstituting the original electrical data from the observed electrical contact resistance measurement. Such a basis is formulated in chapter nine and its application to published results has most interesting consequences. Thus, although the results of the experimental work are negative in nature, a line of theoretical development is suggested that proves extremely promising.

Chapter 9. A basis for the interpretation of electrical contact data.

9.1 Introduction.

It has been shown in chapter three that the low electrical resistance state observed in lubricated contacts does not necessarily correspond to the collision of two opposing asperities. It is likely however, that it does correspond to the close approach of the two surfaces.

The work discussed in chapter two shows that the fluctuation rate and the contact time vary in a systematic manner with the operating variables despite differences in applied potential and current. No basis exists for predicting the changes caused by a deliberate variation of electrical parameters.

The theories of Tallian (42) and Johnson et al. (52), although extremely valuable, do not lead to a practical device for monitoring conditions inside the contact area. Both theories are confined to rolling contact, and the assumption that the duration of electrical contact is equal to the Hertzian transit time is particularly suspect. The exponential behaviour of no-contact time with the expected number of asperity collisions is a most interesting insight on the part of Johnson et al. It does not explain the data, for instance, of Furey (18) shown in Fig. 9.1. The plot of log $\ln(1-\gamma)$ against log load indicates a systematic deviation from any simple dependence that one might expect.

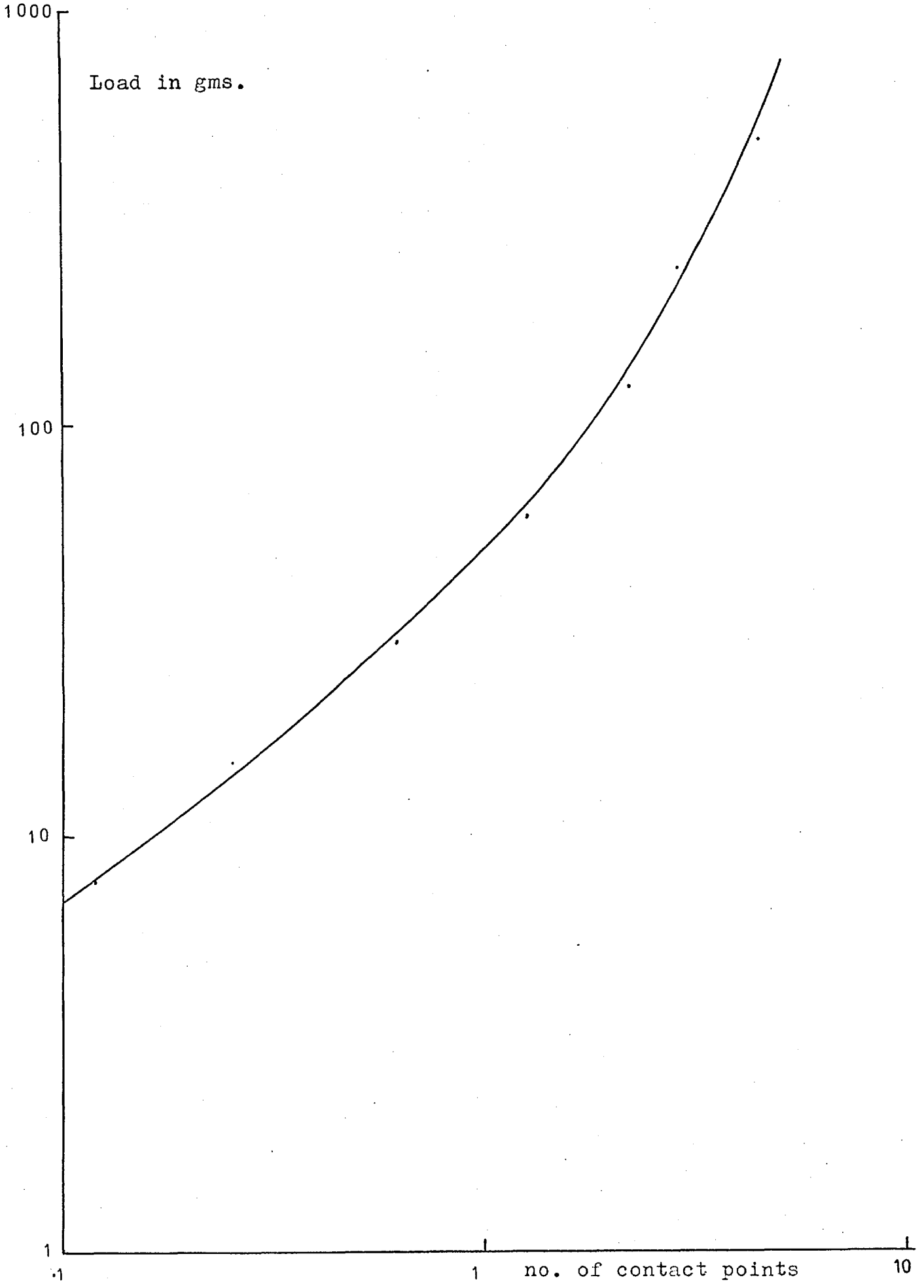


Fig. 9.1 Plot of Furey's data on Johnson basis.

The development that follows treats the electrical signal itself as a random variable, constructed of individual contacts which may or may not overlap. By this means, the results above and other published results are reinterpreted.

9.2 Theoretical approach.

It is necessary to consider an imaginary experiment in which it is possible to detect every individual asperity-asperity contact as it occurs. The word "contact" is used loosely, to describe a drop to the low resistance state. The lines marked in Fig. 9.2 represent the onset of these contact events. They are taken to be infinitesimally thin so that simultaneous events are impossible. With ideal surfaces of regularly spaced identical asperities, one might expect these contacts to be evenly spaced. With real surfaces which have random components both parallel and perpendicular to the direction of motion this assumption is unreasonable, and the intervals between the onset of contacts must be given a statistical distribution $p(\kappa)$. Under a given set of physical conditions the number of individual contacts that occur in a sufficiently long time interval, t, is denoted by C. If these contacts are now assigned a finite width, τ, governed by a distribution $q(\tau)$ then it is possible to derive expressions for the measured number of contacts, C_m, and the no-contact time, $(1 - \gamma)$, taking into account overlapping contact events. This process is illustrated in Figs. 9.3(a) and 9.3(b). In Fig. 9.3(a) the contact levels for simultaneous events are separated for clarity.

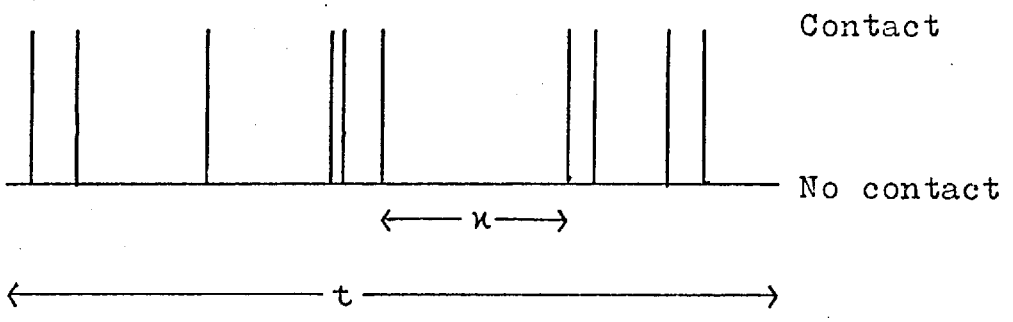
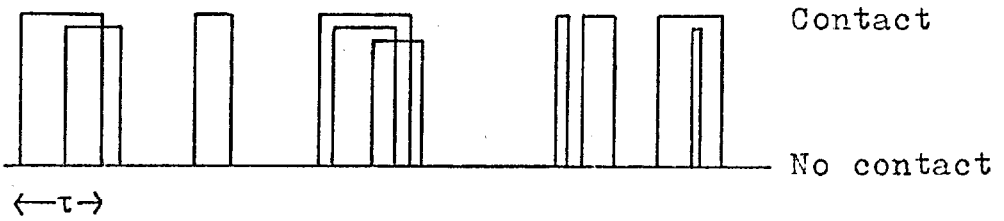
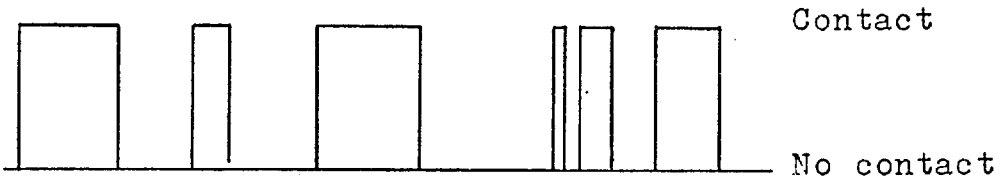


Fig. 9.2 Onset of contacts.



(a) C individual contacts.



(b) C_m compound contacts.

Fig. 9.3 Idealisation of multiple contact occurrences.

9.3 No-contact time and count rate.

In the sampling time t, C events occur. These are separated by C - 1 intervals, but there are two interrupted intervals, at the start and termination of the sampling period. For simplicity the number of intervals is taken as C. In practical sampling times, C is a large number and the effect of this approximation is negligible.

As the probability density of the distribution of κ is $p(\kappa)$, the number of intervals between κ and $\kappa + d\kappa$ in length is $Cp(\kappa)d\kappa$. For these intervals the no-contact time is the sum of the products of all durations τ, less than κ, and the amounts by which κ exceeds τ, i.e.

$$Cp(\kappa)d\kappa \int_0^{\kappa} (\kappa - \tau)q(\tau)d\tau .$$

Summing this quantity over the whole range of κ,

$$t(1 - \gamma) = C\int_0^t p(\kappa)\int_0^{\kappa} (\kappa - \tau)q(\tau)d\tau\, d\kappa \qquad 9.1$$

Similarly, the number of contact durations less than any interval length κ, i.e. the number of counts that register, is given by

$$Cp(\kappa)d\kappa\int_0^{\kappa} q(\tau)d\tau .$$

This is now summed for all intervals κ, thus:

$$C_m = C\int_0^t p(\kappa)\int_0^{\kappa} q(\tau)d\tau\, d\kappa \qquad 9.2$$

The functions $p(\kappa)$ and $q(\tau)$ must be identified before any further development is possible. One might expect that the intervals between C events occurring at random in time t would be described by an exponential

distribution. This was verified by a Monte Carlo method described in Appendix A.

The general form of this dependence is

$$p(\kappa) = A_1 \exp(-A_2\kappa).$$

In order to be an acceptable probability distribution function $p(\kappa)$ must satisfy the condition

$$\int_o^t p(\kappa)d\kappa = 1.$$

Thus $$A_1 \int_o^t \exp(-A_2\kappa)d\kappa = 1$$

i.e. $$-\frac{A_1}{A_2}\exp(-A_2\kappa)\Big|_o^t = 1$$

i.e. $$\frac{A_1}{A_2}\left(1 - \exp(-A_2 t)\right) = 1.$$

If $A_2 t$ is sufficiently large, the second term in the brackets may be neglected, and $A_1 = A_2$. Now, the sum of all the intervals must equal the total sampling time, t. Thus

$$\int_o^t CA_1\kappa\exp(-A_1\kappa)d\kappa = t.$$

i.e. $$-C\exp(-A_1\kappa)(\kappa + \frac{1}{A_1})\Big|_o^t = t$$

$$\frac{C}{A_1}\left(1 - \exp(-A_1 t)[A_1 t + 1]\right) = t$$

Again, if $A_1 t$ is sufficiently large, then $A_1 = \frac{C}{t}$.

Thus $$p(\kappa) = \frac{C}{t}\exp(-\frac{C\kappa}{t}). \qquad 9.3$$

It is now possible to obtain a general relationship between C_m and γ.

Let $$Q(\kappa) = \int_o^\kappa q(\tau)d\tau$$

and let $R(\kappa) = \int_o^\kappa Q(\tau)d\tau$.

Clearly $R(0) = Q(0) = 0$.

From equations 9.2 and 9.3

$$C_m = \frac{C^2}{t}\int_o^t \exp(-\frac{C\kappa}{t})Q(\kappa)d\kappa$$

and $$1-\gamma = \frac{C^2}{t^2}\int_o^t \exp(-\frac{C\kappa}{t})\int_o^\kappa (\kappa-\tau)q(\tau)d\tau\, d\kappa.$$

Now $\int_o^\kappa \tau\, q(\tau)d\tau = \kappa Q(\kappa) - R(\kappa)$, integrating by parts.

Therefore $\int_o^\kappa (\kappa-\tau)q(\tau)d\tau = \kappa Q(\kappa) - \big(\kappa Q(\kappa) - R(\kappa)\big) = R(\kappa)$.

Thus $$1-\gamma = \frac{C^2}{t^2}\int_o^t \exp(-\frac{C\kappa}{t})R(\kappa)d\kappa$$

and integrating by parts,

$$1-\gamma = -\frac{C}{t}\Big(\exp(-\frac{C\kappa}{t})R(\kappa)\Big|_o^t - \int_o^t \exp(-\frac{C\kappa}{t})dR\Big).$$

C is sufficiently large for $R(t)\exp(-C)$ to be negligible compared with the second term.

Thus $$1-\gamma = \frac{C}{t}\int_o^t \exp(-\frac{C\kappa}{t})dR$$

$$= \frac{C}{t}\int_o^t \exp(-\frac{C\kappa}{t})Q(\kappa)d\kappa = \frac{C_m}{C}.$$

Thus $$C_m = C(1-\gamma) \qquad 9.4$$

This holds for any simple function $q(\tau)$.

Three special cases for the distribution of contact durations, $q(\tau)$, are now considered. They are described by the headings "constant τ", "circular distribution" and "exponential distribution". They correspond naively to rolling line contact, rolling point contact and sliding, with or without rolling.

9.4 Constant τ.

In this case the individual contacts are of constant duration, τ', although measured contacts may be longer. The distribution of τ is zero everywhere except at $\tau = \tau'$ and is given by the delta function

$$q(\tau) = \delta(\tau - \tau').$$

The delta function is infinitely high and infinitesimally thin at zero argument, in such a way that its integral is unity. Thus when the range of integration includes zero argument the value of the integral is unity. Otherwise the integral is zero.

Thus $$\int_o^p q(\tau)d\tau = 0 \text{ for } p<\tau'$$

$$= 1 \text{ for } p>\tau'.$$

Thus equation 9.2 becomes

$$C_m = C\int_o^t p(\kappa)\int_o^\kappa \delta(\tau - \tau')d\tau\, d\kappa$$

$$= C\int_{\tau'}^t p(\kappa)d\kappa$$

and from equation 9.3

$$C_m = C\int_{\tau'}^t \frac{C}{t}\exp(-\frac{C\kappa}{t})d\kappa$$

$$= C\left(\exp(-\frac{C\tau'}{t}) - \exp(-C)\right).$$

Since C is large, then to a good approximation

$$C_m = C\exp(-\frac{C\tau'}{t}) \qquad 9.5$$

Similarly equation 9.1 becomes

$$1 - \gamma = \frac{C}{t}\int_o^t p(\kappa)\int_o^\kappa (\kappa - \tau)\delta(\tau - \tau')d\tau\, d\kappa.$$

The only non zero values of this expression occur when $\tau = \tau'$ and $\kappa > \tau'$, thus

$$1 - \gamma = \frac{C}{t}\int_{\tau'}^{t} p(\kappa)(\kappa - \tau')d\kappa$$

$$= \frac{C^2}{t^2}\int_{\tau'}^{t}(\kappa - \tau')\exp(-\frac{C\kappa}{t})d\kappa$$

$$= \frac{C^2}{t^2}\int_{o}^{t-\tau'}\kappa'\exp(-\frac{C\kappa'}{t})\exp(-\frac{C\tau'}{t})d\kappa'$$

where $\kappa' = \kappa - \tau'$.

Thus $$1 - \gamma = \frac{C^2}{t^2}\exp(-\frac{C\tau'}{t})\int_{o}^{t-\tau'}\kappa'\exp(-\frac{C\kappa'}{t})d\kappa'$$

$$= -\frac{C}{t}\exp(-\frac{C\tau'}{t})\Big(\kappa'\exp(-\frac{C\kappa'}{t}) + \frac{t}{C}\exp(-\frac{C\kappa'}{t})\Big)\Big|_{o}^{t-\tau'}$$

$$= \exp(-\frac{C\tau'}{t}) - \Big(1 + \frac{C}{t}(t - \tau')\Big)\exp(-C).$$

The second term may be neglected since C is sufficiently large, thus

$$1 - \gamma = \exp(-\frac{C\tau'}{t}) \qquad 9.6$$

Comparison of equations 9.5 and 9.6 shows that 9.4 is obeyed. The expression $\frac{C\tau'}{t}$ occurs frequently and has a physical interpretation. $C\tau'$ is the total persistence of asperity contact, with each contact counted separately whether occurring multiply or singly. Thus $\frac{C\tau'}{t}$ represents the average number of asperities simultaneously in contact. As an average $\frac{C\tau'}{t}$ may take any positive value or zero.

Equation 9.6 compares directly with the expression derived by Johnson et al. (52) described in section 4.4.3, that $1 - \gamma = \exp(-m)$.

The exponent, m, is the expected number of

asperities contacting at any time. If the contact width is increased in the direction transverse to that of rolling, one would expect C to increase proportionately. In this ideal situation, contact only occurs at the leading edge of the Hertzian area and ceases at the trailing edge. Similarly, if the contact dimension parallel to the rolling direction be changed then τ' will vary accordingly and C will remain the same. Thus, for line contact $\tau' \propto W^{\frac{1}{2}}$ and C is constant. Clearly m will vary as $W^{\frac{1}{2}}$. It is shown in section 9.15 that for a wide range of γ the constant model may be extended to point contact. Then $C \propto W^{\frac{1}{3}}$, $\tau' \propto W^{\frac{1}{3}}$ and m will change as $W^{\frac{1}{3}}$. Thus the dependence of micro-contact points on load is not the same as that found for dry contact by Greenwood and Williamson (70). These results are in agreement with Johnson's conclusion in section 4.4.3 that for rolling contact m is proportional to Hertzian area at fixed film thickness.

The behaviour of C_m is worthy of further study as it appears to go through a maximum as C increases. This maximum may be found by differentiating equation 9.4 with respect to C and equating to zero.

Thus $$\frac{dC_m}{dC} = \exp\left(-\frac{C\tau'}{t}\right) - \frac{C\tau'}{t} \times \exp\left(-\frac{C\tau'}{t}\right) = 0$$

i.e. $$\frac{C\tau'}{t} = 1.$$

Thus the maximum measured count occurs when $C = \frac{t}{\tau'}$. Then, from equation 9.6, $1 - \gamma = \exp(-1) \simeq .368$. The behaviour of γ and C_m with increasing C is displayed in Fig. 9.4.

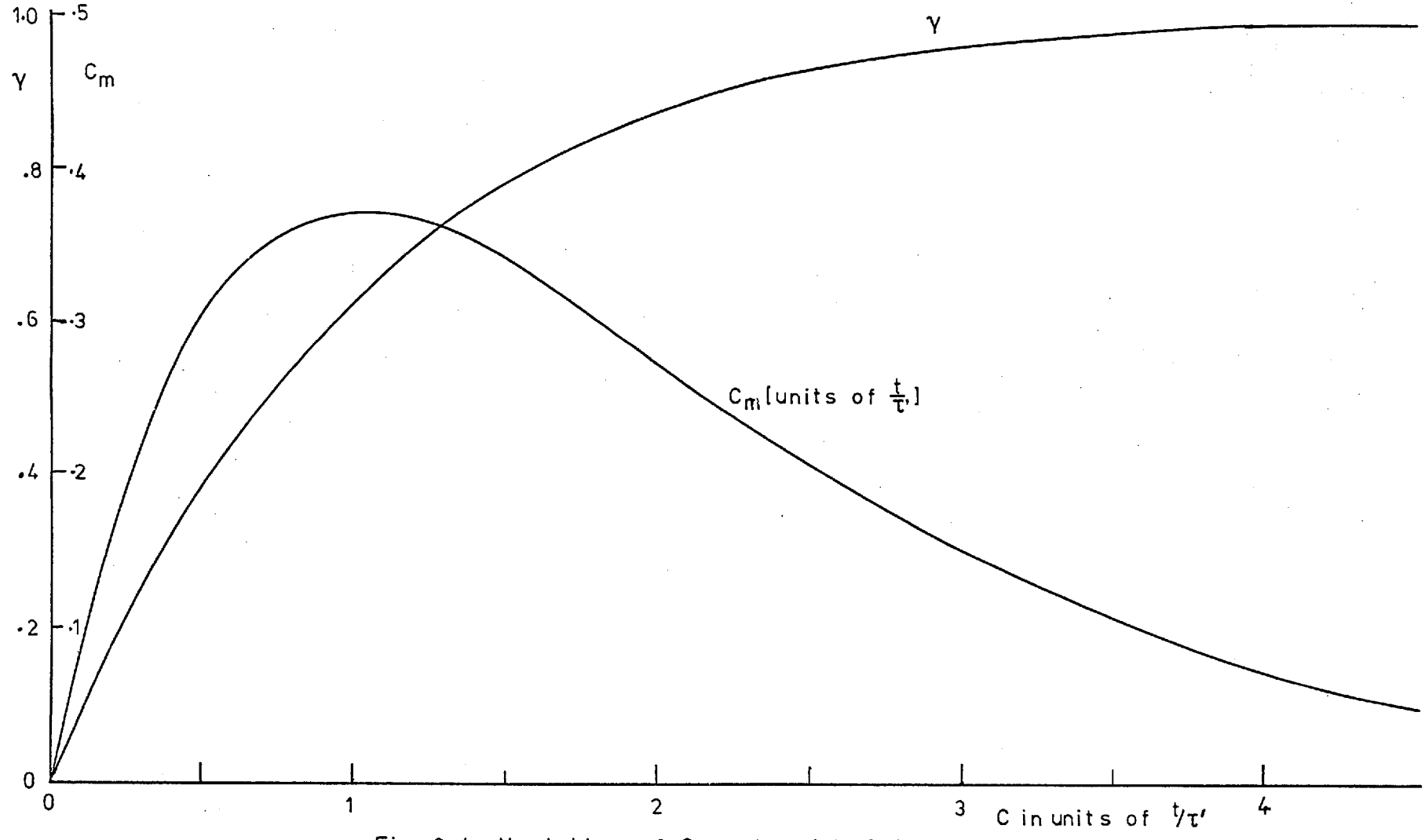


Fig. 9.4 Variation of C_m and γ with C for constant duration.

9.5 "Circular" distribution.

This distribution is formulated to describe the probability density function of the durations of random asperities through a circular Hertzian area. It is derived as follows, with reference to Fig. 9.5 which represents a nominal contact area of radius r. The distance 2L is the Hertzian transit length for an asperity contact occurring at a distance x from the centre line. This distance x, is equally likely to take any value from 0 to r. Thus the distribution of x is given by $p(x) = \frac{1}{r}$. The probability that some value L* lies between 0 and L is the same as that of x lying between r and $(r^2 - L^2)^{\frac{1}{2}}$. Thus if q(L) is the distribution of L, then

$$\int_0^L q(L^*)dL^* = -\int_r^{\sqrt{r^2-L^2}} p(x)dx$$

$$= \frac{r - (r^2 - L^2)^{\frac{1}{2}}}{r}$$

and differentiating with respect to L

$$q(L) = \frac{L}{r(r^2 - L^2)^{\frac{1}{2}}}.$$

This distribution has been derived in terms of the dimensions of the contact area. It applies equally well to durations of contacts since for a fixed speed U, $L = U\tau$. Thus, if L is now interpreted as a contact duration, and r as the transit time through the widest part of the contact, the same formula may be used to give the distribution of dwell times.

From this and equations 9.2 and 9.3,

$$C_m = \frac{C^2}{t}\int_0^t \exp(-\frac{C\mu}{t})\int_0^{\mu/2} \frac{L}{r(r^2 - L^2)^{\frac{1}{2}}} dLd\mu \text{ for real}$$

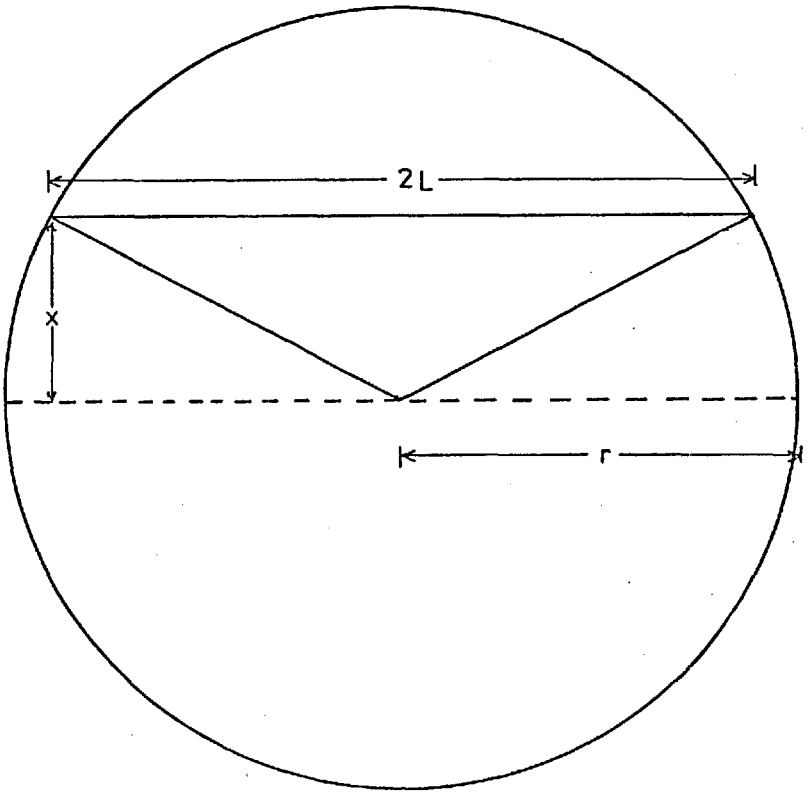


Fig. 9.5 Circular Hertzian Area.

An asperity contact occurring at a distance x from the centre line persists for a distance 2L. The distribution of x is constant since x is equally likely to take any value between 0 and r.

values of square root

$$= \frac{C^2}{rt}\int_o^t \exp\left(-\frac{C\nu}{t}\right)\left[r - (r^2 - \nu^2/4)^{\frac{1}{2}}\right] d\nu$$

$$= C - \frac{C^2}{rt}\int_o^t \exp\left(-\frac{C\nu}{t}\right)(r^2 - \nu^2/4)^{\frac{1}{2}} d\nu.$$

In this expression, if $\nu > 2r$ then the square root term is set to zero. This may be explained by reference to the original derivation of C_m. If $\nu > 2r$ all contacts will occur separately and the term $r - (r^2 - \nu^2/4)^{\frac{1}{2}}$ simply reduces to r. Thus the upper limit of integration may be reduced from t to 2r without changing the real value of the integral.

Thus $$C_m = C - \frac{C^2}{rt}\int_o^{2r} \exp\left(-\frac{C\nu}{t}\right)(r^2 - \nu^2/4)^{\frac{1}{2}} d\nu.$$

If $m = \frac{\nu}{2r}$ and $d\nu = 2r\,dm$

then $$C_m = C - \frac{2C^2 r}{t}\int_o^1 \exp\left(-\frac{2rmC}{t}\right)(1 - m^2)^{\frac{1}{2}} dm.$$

If $A = \frac{2rC}{t}$

then $$\left.\begin{aligned} C_m &= C\left(1 - A\int_o^1 \exp(-Am)(1 - m^2)^{\frac{1}{2}} dm\right) \\ &= C\left(1 - AI(A)\right) \text{ (say)} \end{aligned}\right\} \quad 9.7$$

This integral is not evaluated simply and a series solution for this is described in Appendix B.

From equation 9.4, $\gamma = 1 - \frac{C_m}{C}$

$$= AI(A) \qquad 9.8$$

The quantity A is dimensionless and represents the ratio of the maximum contact transit time 2r to the average time between the onset of contacts, $\frac{t}{C}$. In order

to examine the behaviour of C_m as C varies it is necessary to choose units of $\frac{t}{2r}$. Fig. 9.6 shows γ and C_m as a function of C with C_m and C in units of $\frac{t}{2r}$, so that C has numerically the same value as A. The integral I(A) is tabulated in Appendix B (Fig. B.2)

C	C_m	γ
0.01	.0099	.0078
0.1	.0925	.0753
0.2	.1711	.1445
0.3	.2376	.2081
0.5	.3398	.3204
0.8	.4341	.4574
1.0	.4685	.5315
1.35	.4919	.6356
1.4	.4925	.6482
1.5	.4920	.6720
2.0	.4673	.7664
5.0	.2286	.9542
10.0	.1036	.9896

Fig. 9.6 Variation of γ and C_m with C.

This data is presented graphically in Fig. 9.7. It can be seen that C_m has a maximum value at $C \simeq 0.7\frac{t}{r}$ where $\gamma \simeq 0.648$. It is also clear that, given any value of γ a unique value of A is defined.

The mean value of the circular distribution is shown in section 4.7 to be $\frac{\pi}{4}$ x diameter. In this case it is $\frac{\pi r}{2}$ or $\frac{At}{4C}$. The average number of contact points, m, is given by $\frac{C}{t} \times \frac{\pi r}{2}$ i.e. $\frac{1}{4}A$.

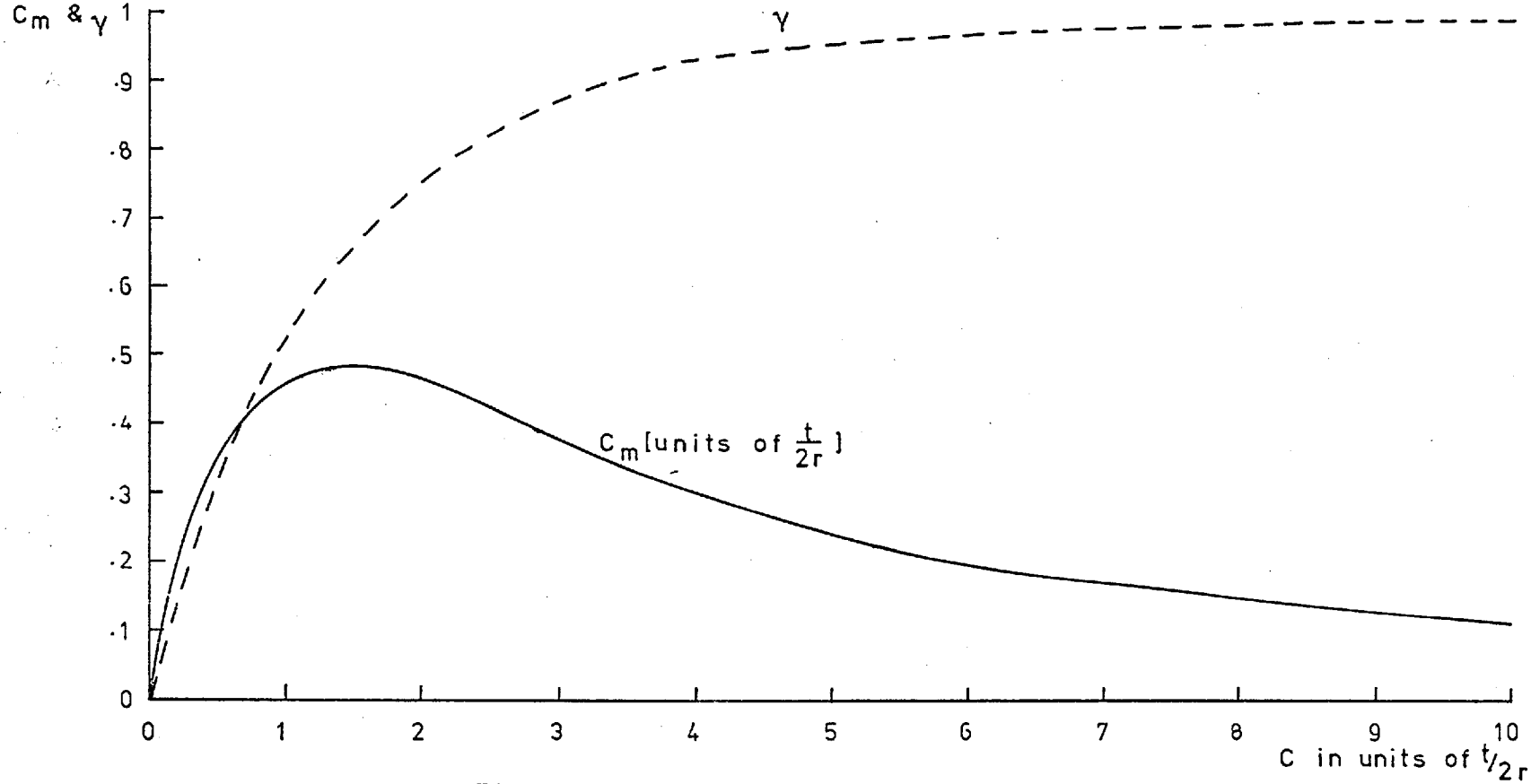


Fig. 9.7 Variation of C_m and γ for circular distribution.

9.6 Exponential distribution.

We now let $q(\tau) = a\exp(-\frac{\tau}{\tau_m})$ where τ_m is the mean value for duration of an isolated contact.

Now $\int_o^t q(\tau)d\tau = 1.$

Thus $1 = a\int_o^t \exp(-\frac{\tau}{\tau_m})d\tau$

$= -a\tau_m \exp(-\frac{\tau}{\tau_m})\Big|_o^t$

If $t \gg \tau_m$ then $a = \frac{1}{\tau_m}$.

Thus $q(\tau) = \frac{1}{\tau_m}\exp(-\frac{\tau}{\tau_m})$.

From equations 9.2 and 9.3 and the above equation the measured count is given by

$$C_m = C\int_o^t \frac{C}{t}\exp(-\frac{C\nu}{t})\int_o^\nu \frac{1}{\tau_m}\exp(-\frac{\tau}{\tau_m})d\tau\, d\nu.$$

The inner integral is the same form as that above and is simply, $1 - \exp(-\frac{\nu}{\tau_m})$.

Thus $C_m = \frac{C^2}{t}\int_o^t \exp(-\frac{C\nu}{t}) - \exp(-\nu(\frac{C}{t} + \frac{1}{\tau_m}))d\nu.$

$= -C\exp(-\frac{C\nu}{t})\Big|_o^t + \frac{C^2}{C + \frac{t}{\tau_m}}\exp(-\nu(\frac{C}{t} + \frac{1}{\tau_m}))\Big|_o^t.$

The upper limits may again be neglected, and

$$C_m = C\left(1 - \frac{1}{1 + \frac{t}{C\tau_m}}\right)$$

$$= C\frac{1}{\frac{C\tau_m}{t} + 1} \qquad 9.9$$

From equations 9.4 and 9.9

$$\gamma = 1 - \frac{1}{1 + \frac{C\tau_m}{t}}$$

$$\gamma = \frac{\frac{C\tau_m}{t}}{1 + \frac{C\tau_m}{t}} \qquad 9.10$$

The expression $\frac{C\tau_m}{t}$ is again identified with the average number of asperities simultaneously in contact, since by definition of τ_m,

$$C\tau_m = C\int_0^t \tau\, q(\tau)\, d\tau .$$

The variation of γ and C_m with C is shown in Fig. 9.8 where C is in units of $\frac{t}{\tau_m}$. When C_m is measured in these units C_m and γ are numerically equal.

9.7 Comparison of models and method of application.

The three types of variation of C_m, each in appropriate units, are shown in Fig. 9.9 as a function of γ.

For all the models C is found from equation 9.4,

i.e. $$C = \frac{C_m}{1 - \gamma}$$

The average number of simultaneous micro-contacts, m, is found in each case from γ by suitable manipulation of equations 9.6, 9.8 and 9.10.

The number of contact events C, occurring in time t has been taken as constant in this analysis. In practise it would be subject to statistical variation. In a mechanical contact which is running unsteadily this variation will become larger. It may be that a better guide to the conditions inside the contact is given by the mean contact duration τ_m. Fig. 9.10 shows m and τ_m expressed in terms of the measured quantities γ and C_m for the three models.

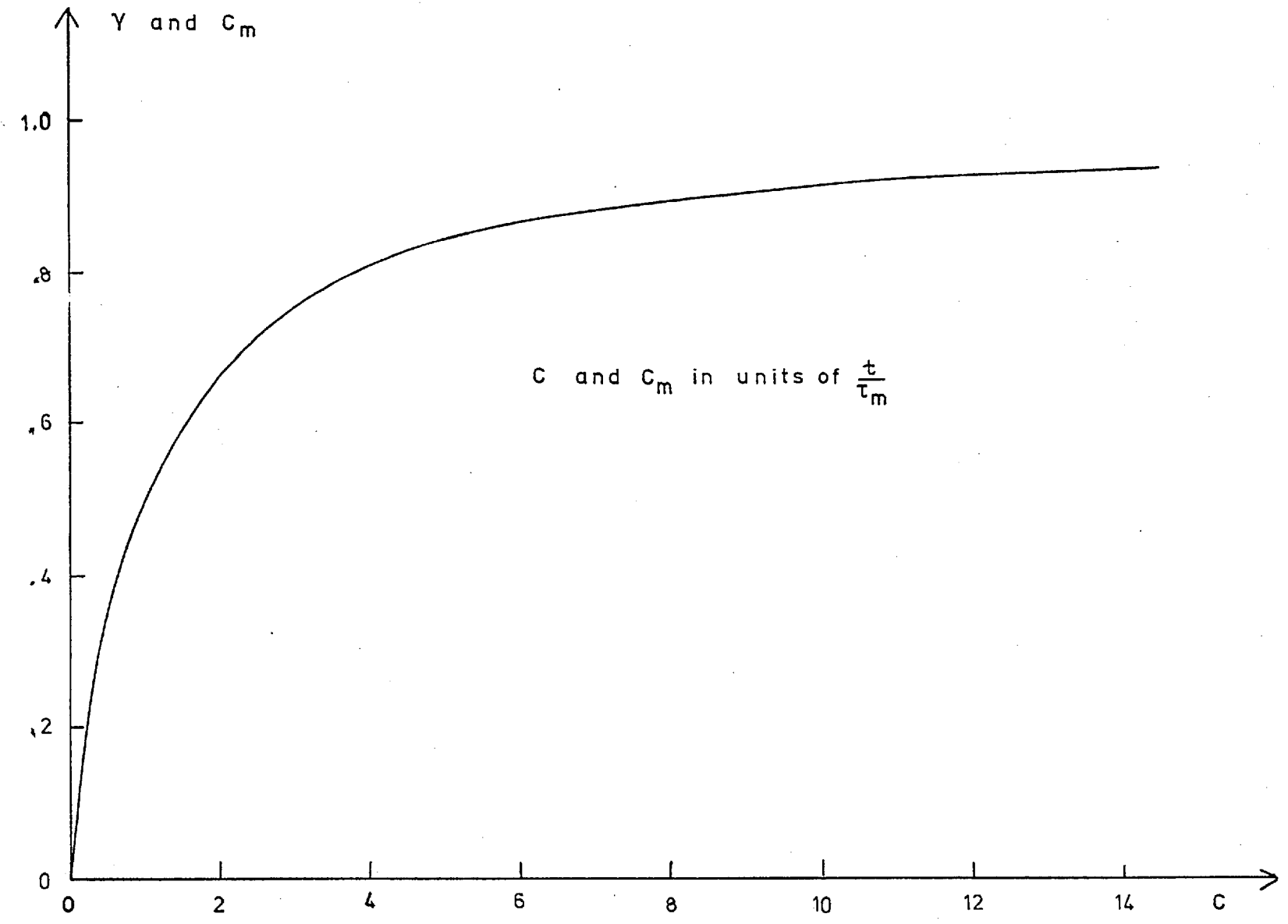


Fig. 9.8 Variation of γ and C_m with C for exponential model.

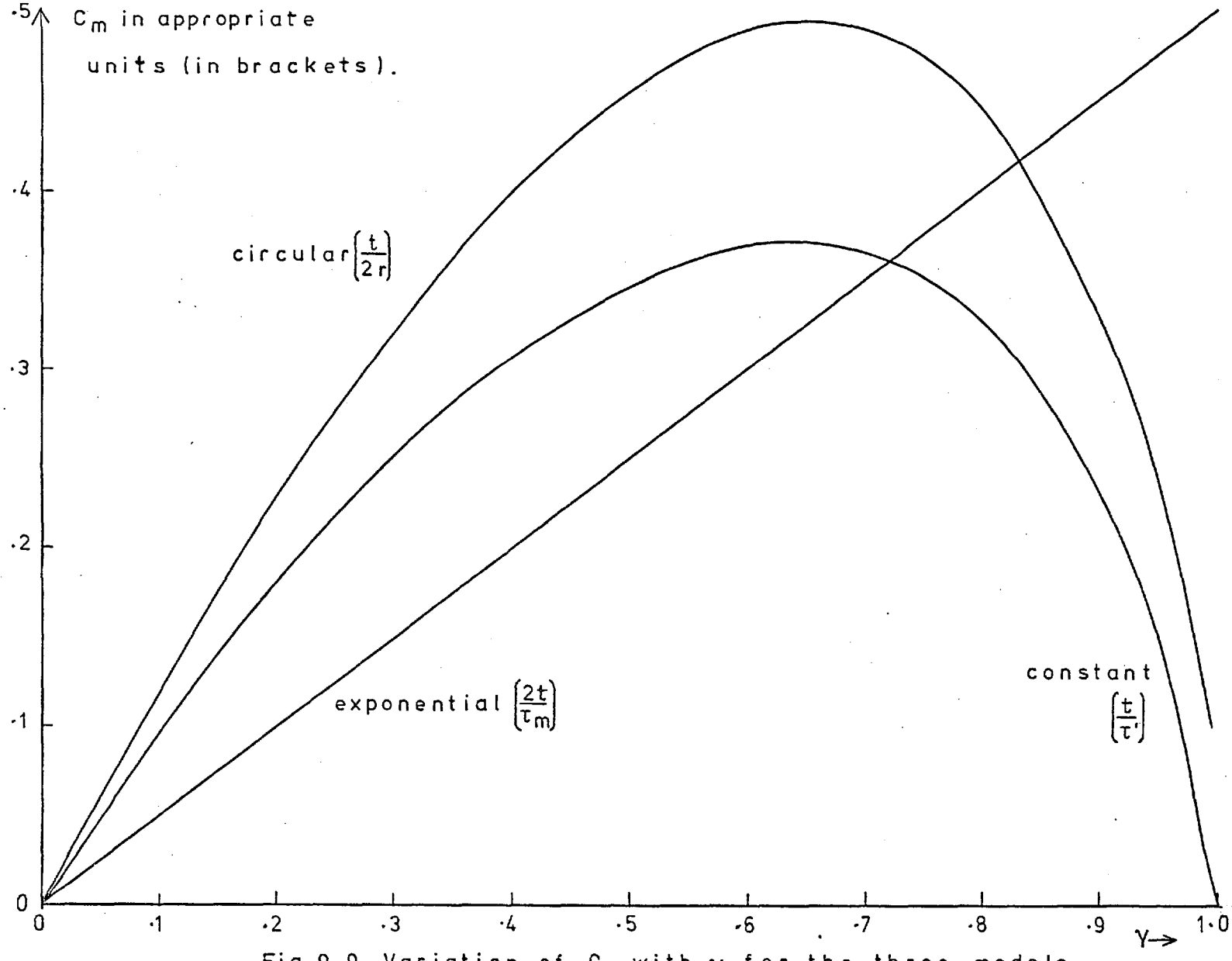


Fig. 9.9 Variation of C_m with γ for the three models.

Model	m	τ_m
(i) constant τ	$-\ln(1-\gamma)$	$\frac{(\gamma-1)\ln(1-\gamma)}{C_{m/t}}$
(ii) circular distribution	$\frac{1}{4}A$	$\frac{A(1-\gamma)}{4C_{m/t}}$
(iii) exponential distribution	$\frac{\gamma}{1-\gamma}$	$\frac{\gamma}{C_{m/t}}$

Fig. 9.10 Contact points and mean duration.

The values of m depend only on γ (and the appropriate model) whereas τ_m also requires simultaneous measurement of count rate. The τ_m, derived from electrical measurements may be used in conjunction with the theoretical model of chapter four to examine surface topography and oil film thickness.

9.8 Physical basis and validity of assumptions.

The physical cause of the low resistance state is discussed in chapter three. If, by the correct choice of circuitry, it is possible to distinguish between actual contact and quasi-metallic contact, then the interpretation of electrical contact data is direct. One would expect however that asperities as close as some fraction of one micro-inch would have some significant interaction as regards, for instance, traction behaviour. More experimental work is needed in this area before it will be possible to link electrical contact data directly with wear or friction.

It does appear that the low resistance state corresponds to the close approach of asperities and is very

sensitive to small separations. Electrical data may then be interpreted in terms of the contact that will occur when the oil film is slightly reduced in thickness. This is extremely important as it implies the existence of a predictive element in the electrical signal. This point is discussed later with relevance to the work of Czichos (49).

The assumption that the contact events are distributed exponentially is valid as long as the contact area is sufficiently large for a reasonable number of asperities to have the possibility of breaking through the oil film. At large values of $\frac{h}{\sigma}$ a small number of dominant surface features may dictate electrical behaviour. Otherwise this is a reasonable assumption, at least as acceptable as the statistical assumptions made of surface topography.

The "constant τ" and "circular distribution" models correspond nominally to rolling line and point contacts and are clearly of theoretical importance. There is some doubt as to whether or not the Hertzian area transit time dominates electrical behaviour (cf. Fig. 2.15).

Christensen's measurement of sliding contact durations (Fig. 2.16) suggests that an exponential distribution of dwell times is appropriate. Fig. 9.11 shows an exponential curve overlaid on Christensen's data. The mean value of this distribution is about 5μs. which gives better agreement with the value obtained in section 4.7. Referring again to the rolling contact data in Fig. 2.15 an exponential distribution is a better approximation than

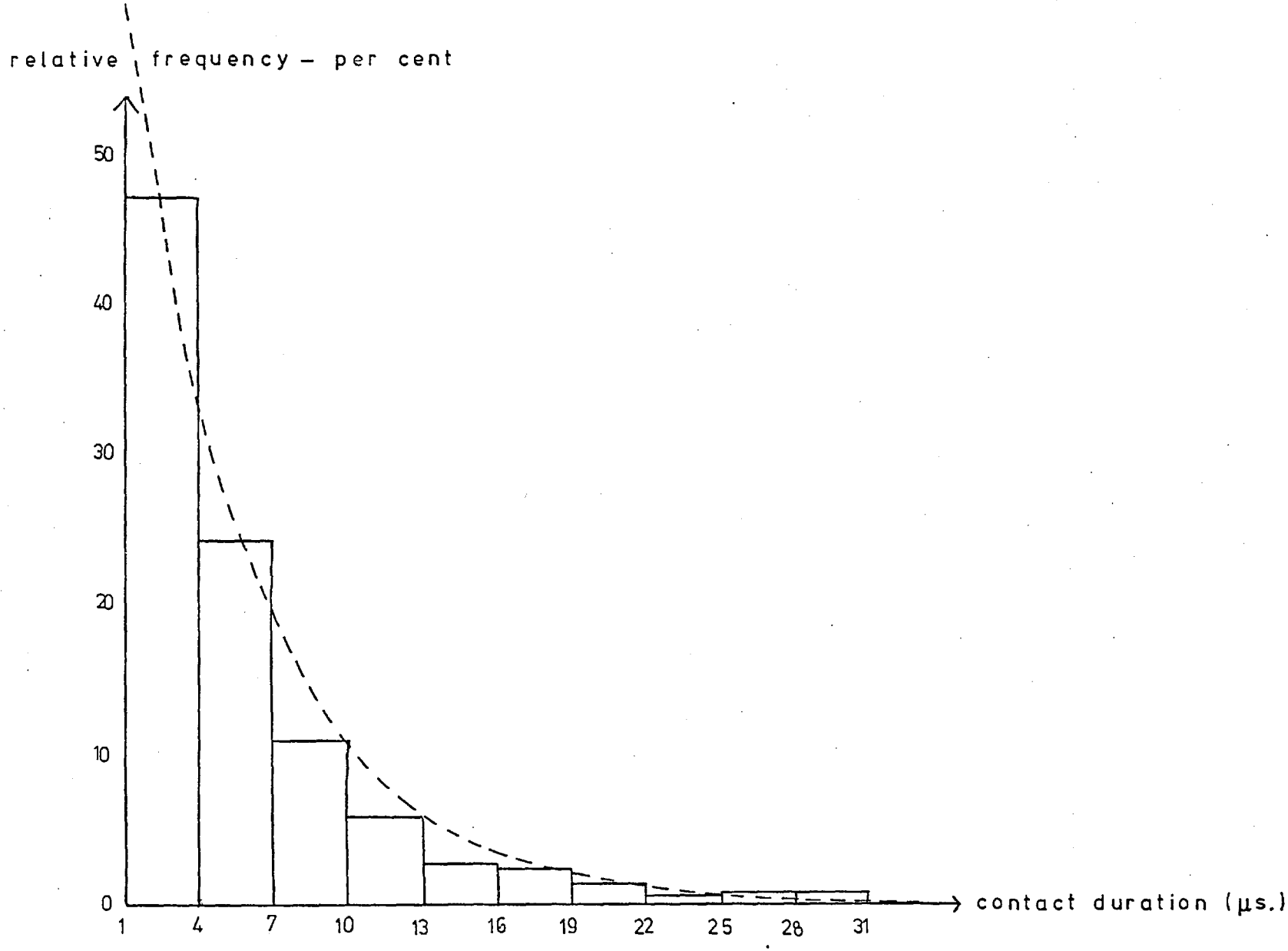


Fig. 9.11 Exponential distribution of contact durations.

the constant duration commonly assumed. In most of the experimental data subsequently re-examined the exponential model is used.

Electrical measurements made across the races of bearings are extremely difficult to interpret. Let I_r, O_r represent an electrical contact between the rth roller and the inner race and outer race respectively. The contact event is then expressed in Boolean Algebra by the composite event $(I_1 \cap O_1) \cup (I_2 \cap O_2) \cup (I_3 \cap O_3) \cup \ldots (I_n \cap O_n)$, where n is the number of rollers. There is also the possibility of roller-cage contacts. The author has made no attempt to solve this problem and the electrical data have been treated as for a simpler case. It may be that the contact durations arising from this complex interaction follow a simple law. Even then it is not easy to relate derived quantities to conditions inside the bearing. However a value of "m" has been calculated from the experimental data of Leaver et al. (26) and Garnell (24) and shows systematic variation.

The frequency response of the counting system has been ignored in the theoretical treatment above. The effect of this will be to reduce the measured value of C_m at high rates of fluctuation. It is possible to incorporate this effect into the models above, but this has not been deemed appropriate at this stage.

The above theory is now applied to some of the results discussed in chapter two. The data for these results have been taken from published graphs by means of

a travelling microscope, or simply a ruler where appropriate. Where graphs have been plotted on probability paper, the published data were scaled up onto full size probability paper and the ordinates thereby obtained. The special cases considered cover rolling and sliding in both line and point contact as well as two types of bearing.

9.9 Roller bearing.

The data of Garnell (24) shown in Fig. 2.5 are not directly appropriate for reinterpretation as a relationship must be found between the average contact resistance, R and γ. This is obtained with reference to Fig. 9.12 which shows a simple potential divider having resistances R_1 and R_2 connected to a bearing with apparent resistance R.

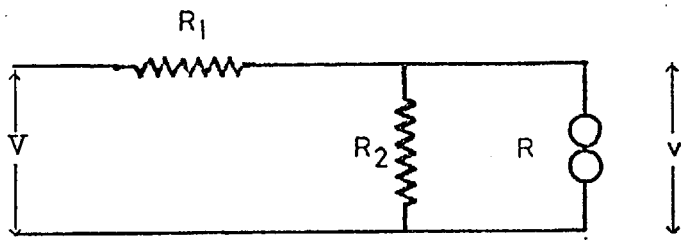


Fig. 9.12 Test circuit (Garnell).

If v is the measured potential across the bearing

then $1 - \gamma = \frac{v(R_1 + R_2)}{VR_2}$ where V is the applied potential.

Now $$v = \frac{\frac{RR_2}{R + R_2}}{R_1 + \frac{RR_2}{R + R_2}} V$$

Thus $$1 - \gamma = \frac{R_1 + R_2}{R_2} \times \frac{RR_2}{R_1(R + R_2) + RR_2}$$

$$= \frac{R(R_1 + R_2)}{R_1R + R_1R_2 + R_2R}.$$

Then $\gamma = \frac{R_1R_2}{R_1R + R_1R_2 + RR_2}.$

For the exponential distribution of contact durations $m = \frac{\gamma}{1 - \gamma}$ (Fig. 9.10).

In this case $m = \frac{R_1R_2}{R(R_1 + R_2)}$

Thus m is proportional to the reciprocal of the apparent contact resistance. The values of R_1 and R_2 have been taken as 200kΩ and 10kΩ respectively in Fig. 9.13 which shows m plotted against λ', the ratio of film thickness to the sum of the c.l.a. roughnesses of race and rollers. The gradients of the lines are different for the rough and smooth cases and the equations derived from Fig. 9.13 are:

$$\lambda' \text{ rough} = .89\ m^{-.174}$$

$$\lambda' \text{ smooth} = .90\ m^{-.254}.$$

9.10 Tapered roller thrust bearing.

The data of Leaver, Sayles and Thomas (26) have been plotted in Fig. 9.14 as "m" $= \frac{\gamma}{1 - \gamma}$ against speed U. One would expect in this case that the oil film between the upper race and rollers is essentially the same as that of the lower race and rollers, and that the surface finish of both races is the same.

It is clear that γ can also be interpreted as the probability that there is contact at any given moment in time and that the probability of two independent systems giving contact at the same time is easily derived.

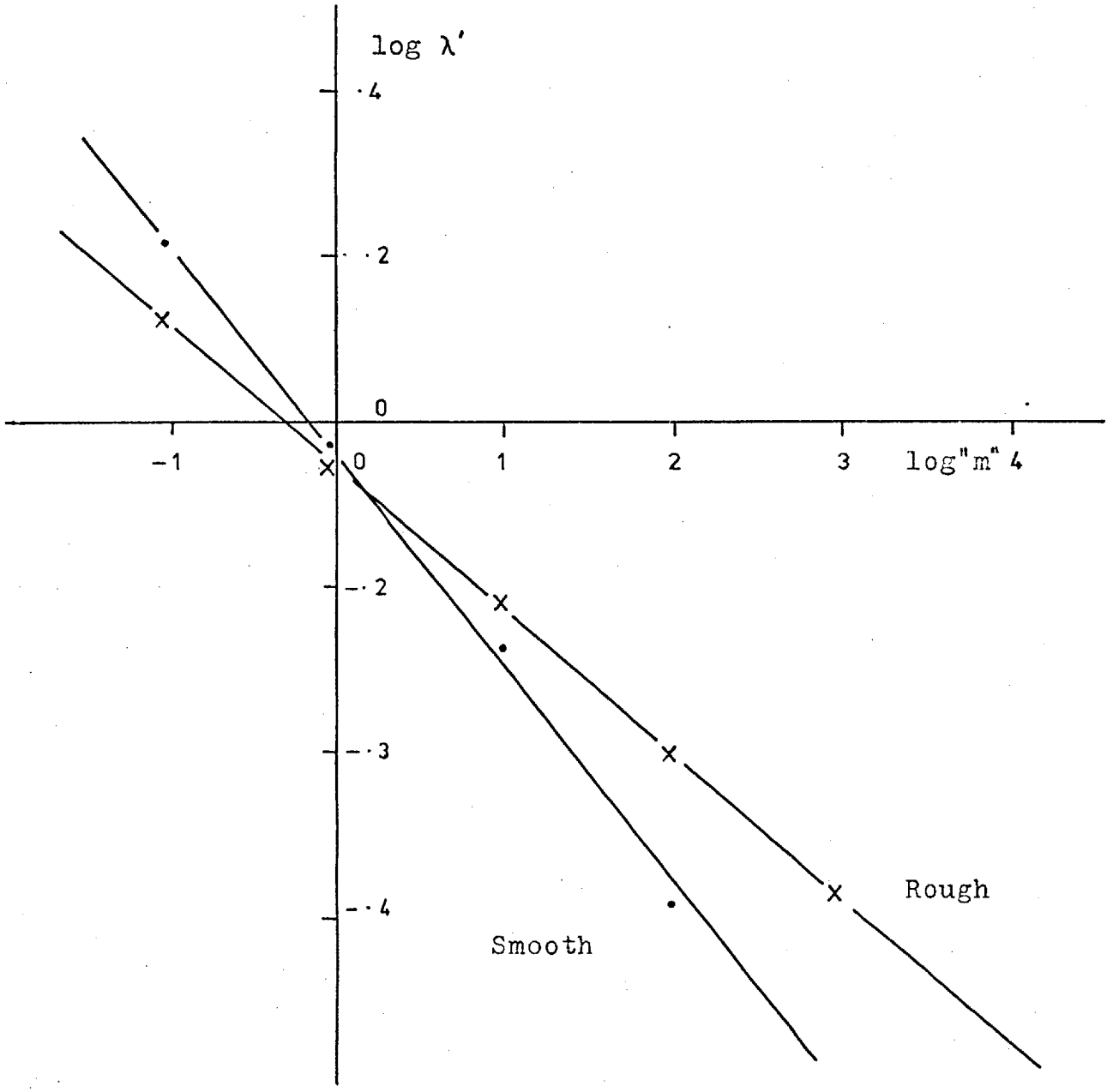


Fig. 9.13 Dimensionless film thickness and "m" for roller bearing.

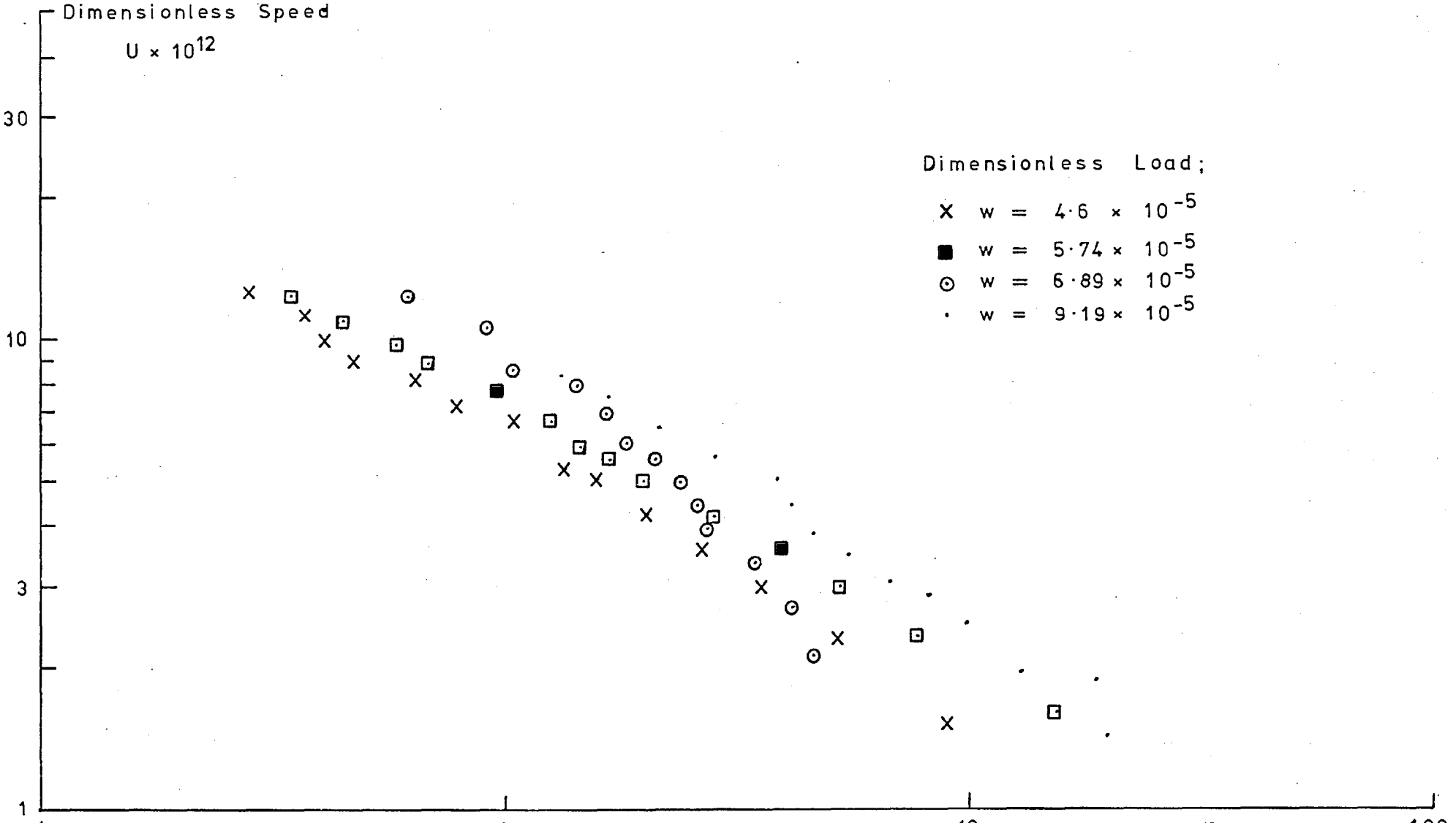


Fig. 9.14 Variation of "m" with speed.

If γ_1, γ_2 represent the average contact between a race and the rollers one would expect the measured contact between races to be $\gamma_1 \times \gamma_2$. Since there should be no difference between γ_1 and γ_2 on a statistical basis the quantity m has also been redefined as $m = \frac{\gamma^{\frac{1}{2}}}{1 - \gamma^{\frac{1}{2}}}$ in Fig. 9.15 and this should be a more realistic parameter to study. It can be seen that the gradients of the highest and lowest loads are similar and show a load dependence of approximately $W^{0.7}$. For the two middle loads the gradients are quite different. The range of behaviour of speed with m is given by $U = km^{-n}$ where n ranges from .60 to 1.06. If film thickness is taken as being proportional to $U^{0.7}$, then $h = k^1 m^{-p}$ where p lies between 0.42 and 0.74, where k and k^1 are some constants.

These differences are no doubt partly explained by the changes in surface roughness during the course of the experiment. Unfortunately only initial and final values are given for surface roughness and an interpretation along these lines would be speculative.

The authors stress the importance of filtering profiles in order to include in the roughness value, σm, those wavelengths relevant to the contact width. A plot of $\frac{h}{\sigma m}$ and γ for two loads has been redrawn in Fig. 9.16 where $m = \frac{\gamma^{\frac{1}{2}}}{1 - \gamma^{\frac{1}{2}}}$. The equations of the lines in this figure are: for the high load $\frac{h}{\sigma m} = 4.2m^{-.50}$ and for the low load $\frac{h}{\sigma m} = 3.1m^{-.45}$.

9.11 Line contact-rolling.

The data of Christensen in Fig. 2.17 have been

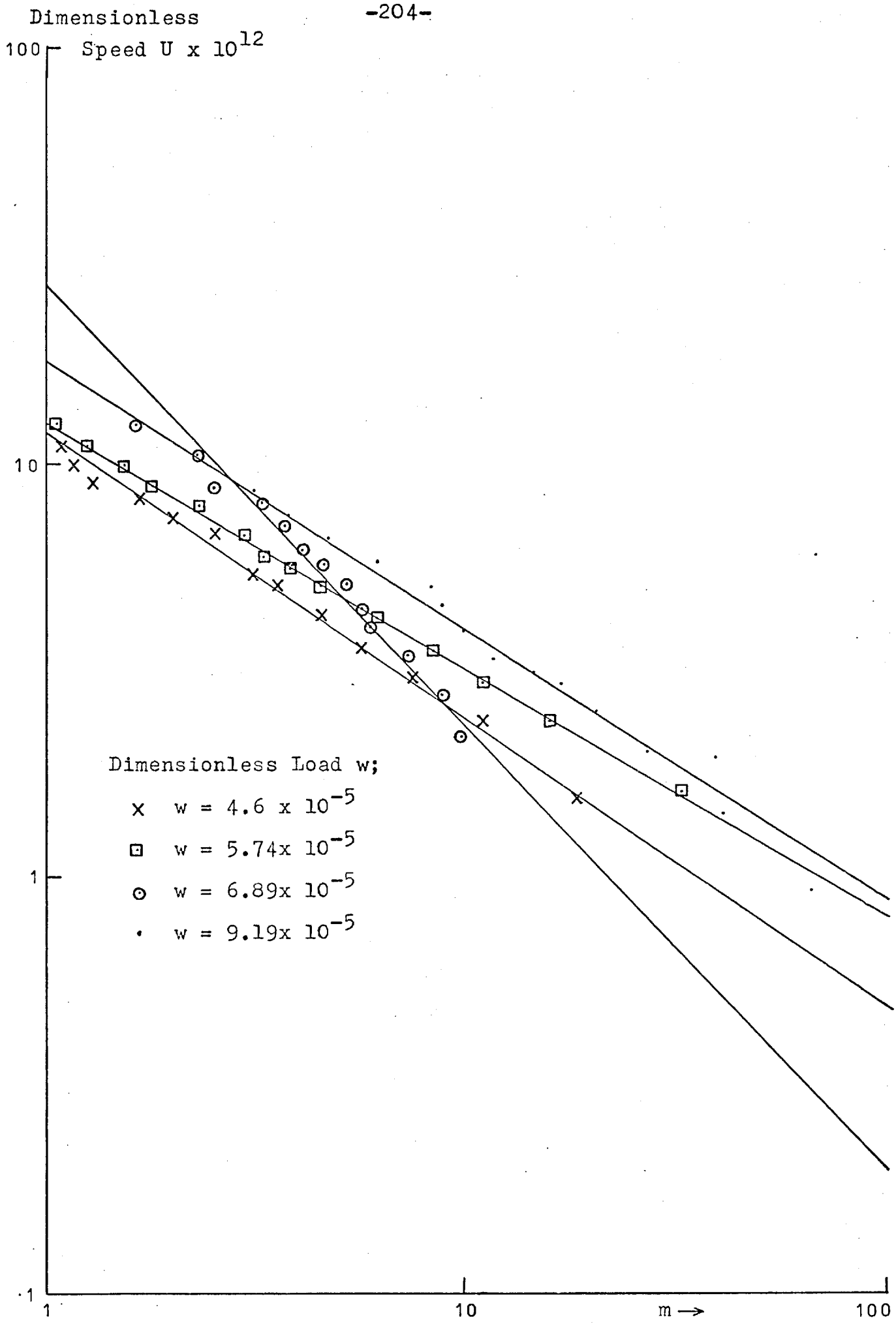


Fig. 9.15 Variation of corrected m with speed.

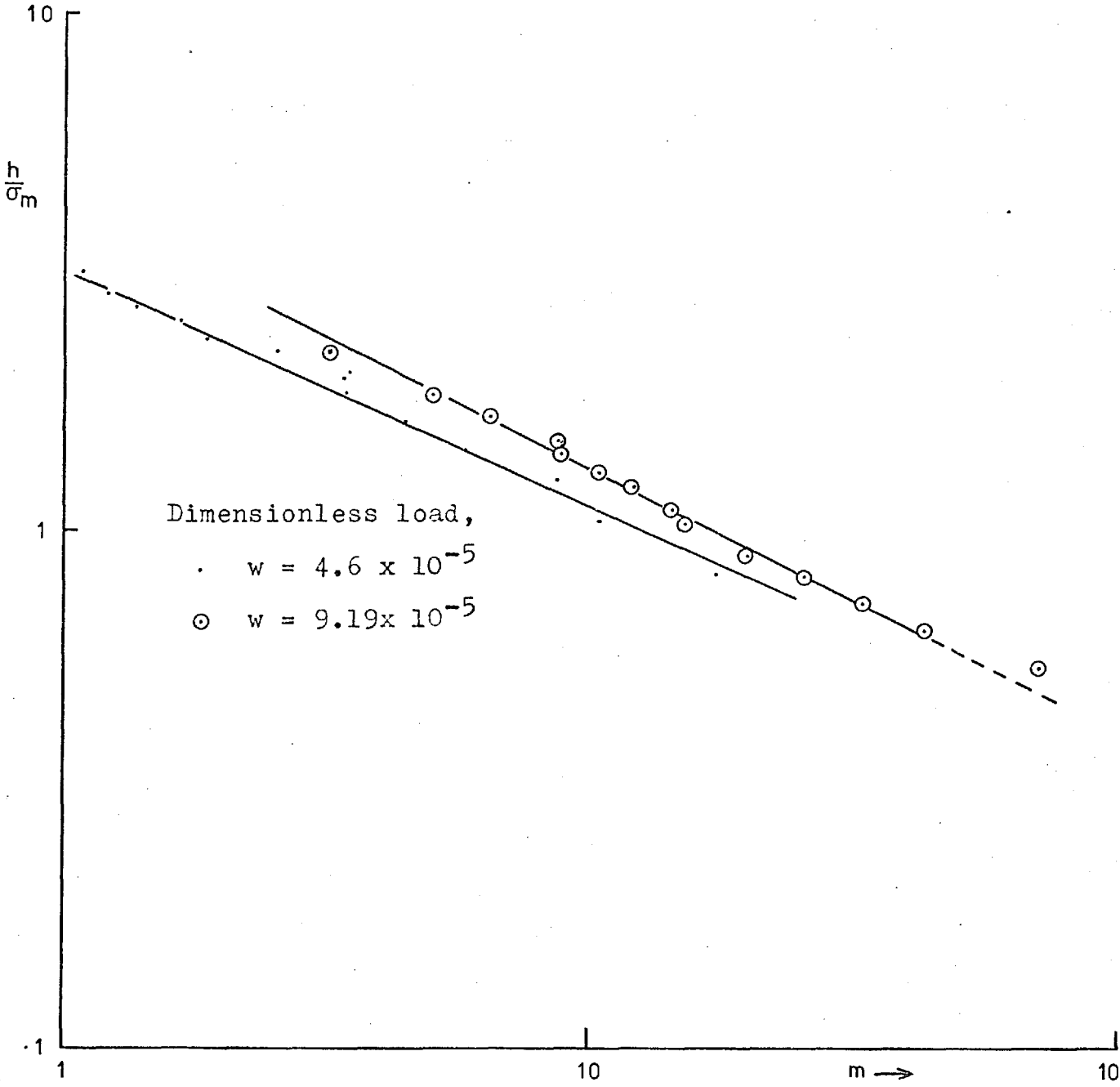


Fig. 9.16 Variation of corrected m with oil film thickness.

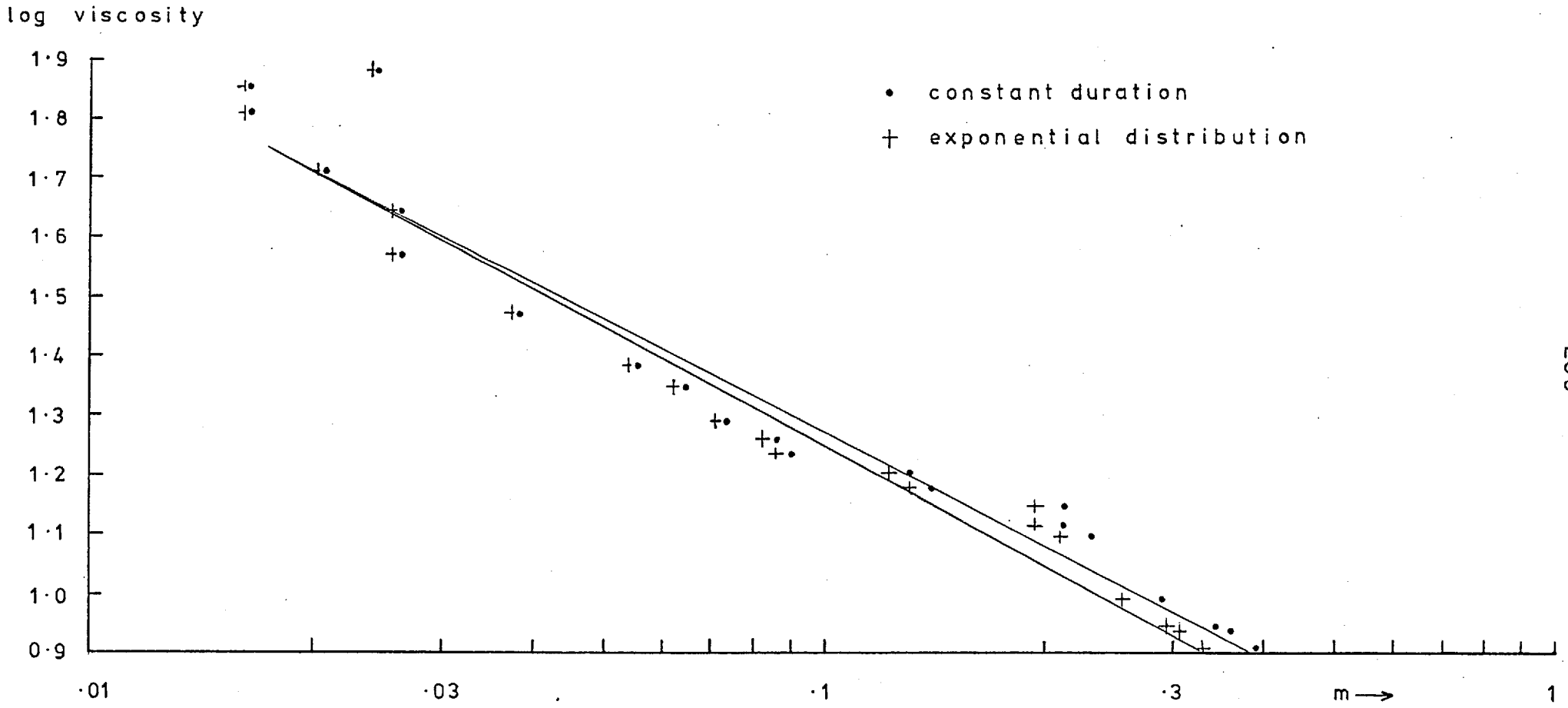


Fig. 9.17 Variation of m with viscosity for rolling line contact.

replotted in Fig. 9.17 as m against viscosity where m is defined by both constant duration and exponential models. The difference between the m values is small since γ is fairly small. In each case the points are better fitted by two parallel straight lines rather than by a single line. There may or may not be reason for systematic deviation, such as dislodgement of a thermocouple or perhaps some chemical effect. A gradient has been obtained for comparison purposes and using an empirical formula of Christensen's in which film thickness is proportional to $(\text{viscosity})^{0.83}$ the following proportionalities between m and film thickness are obtained;

$$h \propto m^{-0.55} \text{ (constant } \tau \text{)}$$

$$h \propto m^{-0.53} \text{ (exponential model).}$$

From the simultaneous measurement of γ and C'_m, values of τ have been calculated as defined by Fig. 9.10(i). These and the measured values of τ are displayed in Fig. 9.18. Both sets of points follow the same general pattern, but it is surprising that the calculated values are generally greater than the measured ones. The calculated quantity represents single contacts whereas the measured τ admits multiple contacts as well. This discrepancy must be explained either by deficiencies of the model or by inaccurate measurements. The large difference that Christensen found between γ and the $\tau C'_m$ product suggests that the fluctuation rate as measured is too low. The fact that this difference was more marked for sliding indicates that the frequency response of the electrical

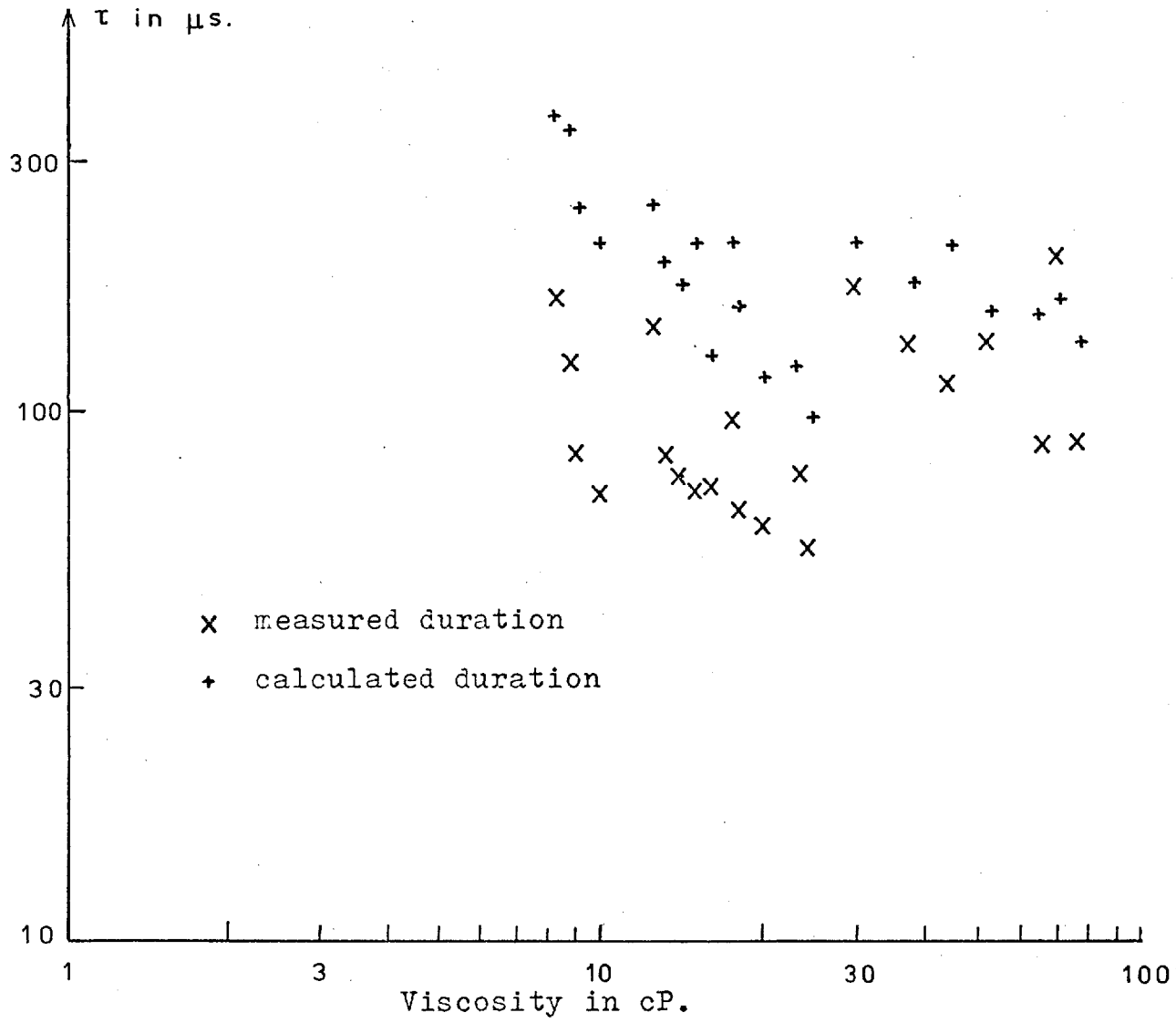


Fig. 9.18 Comparison of calculated and measured contact durations.

system as a whole was insufficient.

9.12 Line contact-sliding.

Czichos (49) investigated the amount of time for which the contact resistance was less than 10Ω and greater than 50kΩ in a sliding cylinder on flat device. From these results it is possible to define γ for contact resistances (a) less than 10Ω

(b) less than 50kΩ.

This has been done in Fig. 9.19 where m (exponential model) and speed variation are shown for cylinders of two different roughnesses. The behaviour is reasonably linear except for the extreme points where errors in obtaining the data from the published work may be significant. The gradients for both rough surfaces are essentially the same and the variation is fairly well described by $\frac{h}{\sigma} \propto m^{-0.08}$ as the coefficients of m for the solid lines are closely in the ratio of the roughness values. The broken lines are m values 1.5 and 3 times greater than calculated for the rougher and smoother surfaces respectively. It can be seen that under these experimental conditions the effect of increasing the acceptable resistance of a contact is to observe many more contacts.

9.13 Point contact-sliding.

The data of Furey (18) have been redrawn in Figs. 9.20 and 9.21. For values of γ between .04 and .95 there is an excellent correlation between load and m, described by

$$W = 31.5m^{0.692}$$

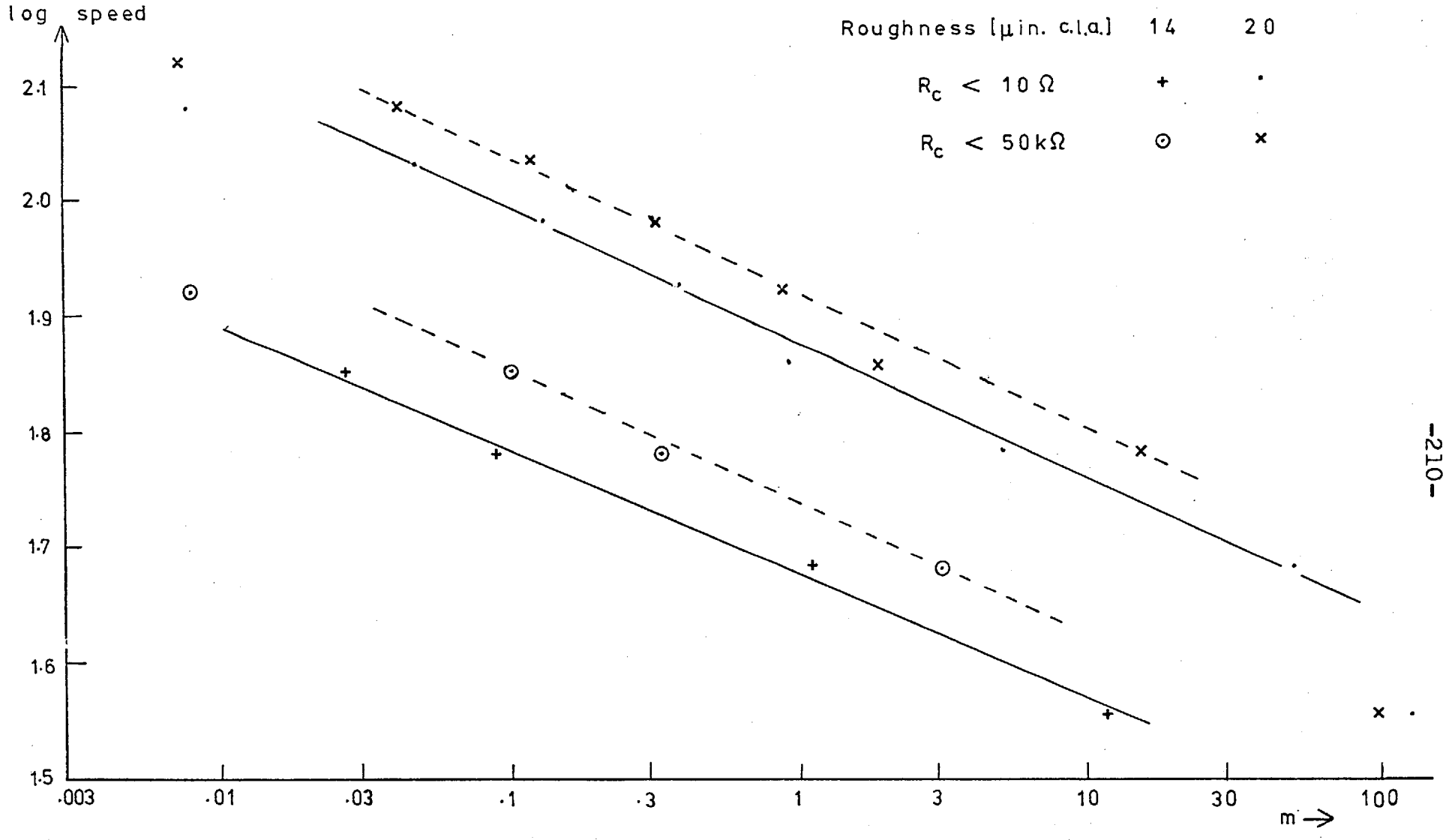


Fig. 9.19 Variation of m with speed for sliding line contact.

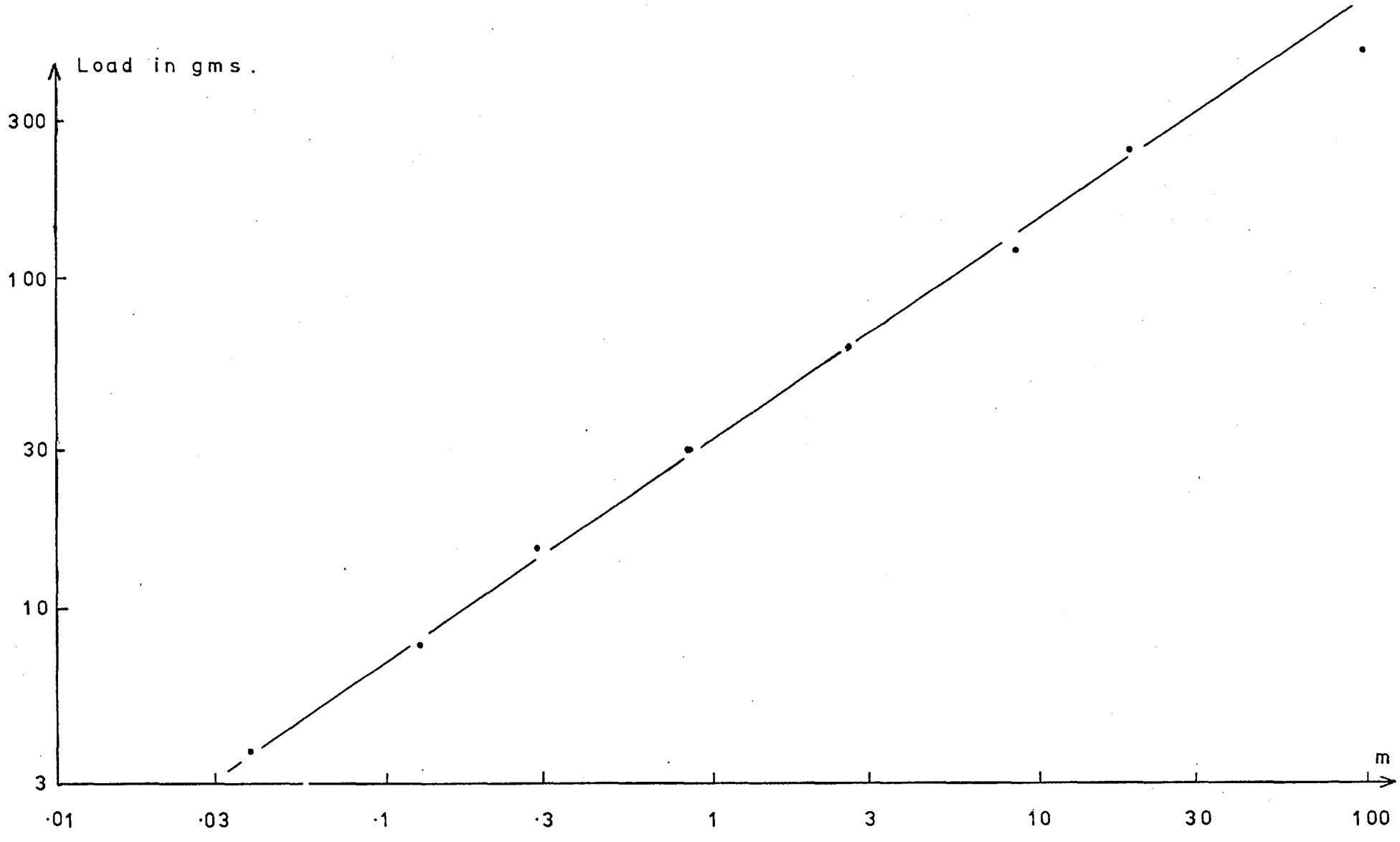


Fig.9.20 Variation of m with load for sliding point contact.

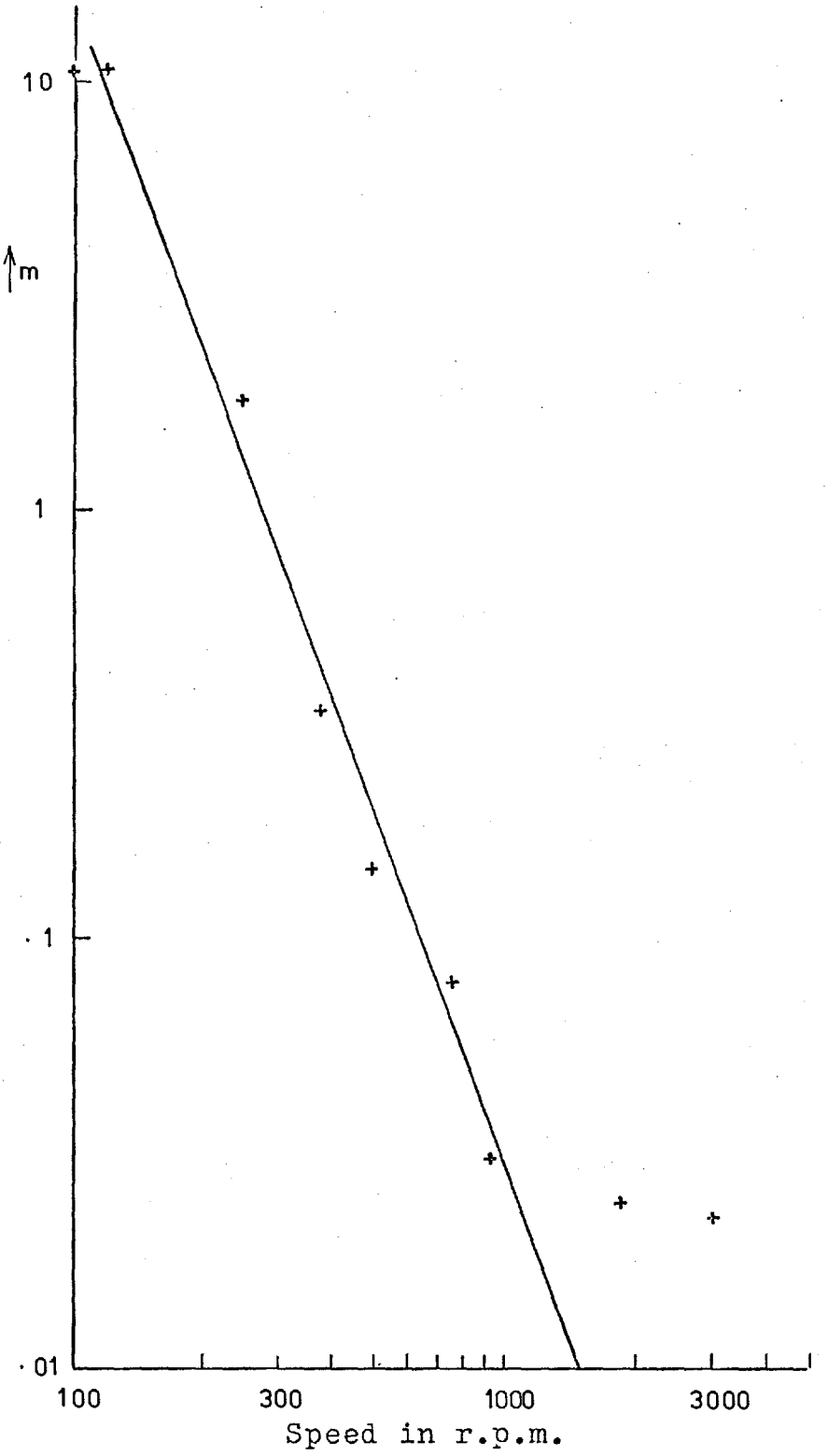


Fig. 9.21 Variation of m with speed for sliding point contact.

where W is the load in grams. The "S" shaped curve of Fig. 2.2(a) becomes an acceptable straight line.

The transformation of Fig. 2.2(b) into Fig. 9.19 is not so convincing. There is significant deviation at the highest speeds where the values of γ and m may be dominated by a small number of prominent surface features. However, an equation was obtained for the line drawn in Fig. 9.21 and the film thickness is related to m, for fixed load, by the proportionality $h \propto m^{-0.27}$.

9.14 Point contact-rolling with sliding.

Johnson et al. (52) noted that the data of Poon and Haines (27) gave a good fit when plotted on a basis of $m = -\ln(1 - \gamma)$. This is shown in Fig. 9.22. A better fit is obtained using the exponential model where $m = \frac{\gamma}{1 - \gamma}$ (Fig. 9.23). The systematic deviation at high contact is reduced with the new model whereas the low contact points are only slightly increased in value. The effect of the limited frequency response of the electrical circuit used in this experiment has been discussed in chapter three. At the highest film thicknesses γ will be artificially high due to the long time constant. This is consistent with the deviation found at the lowest values of m in Figs. 9.22 and 9.23.

The relation between film thickness and m for Fig. 9.23 is $h \propto m^{-0.118}$. No constant of proportionality is found as the film thickness measurements are in arbitrary units.

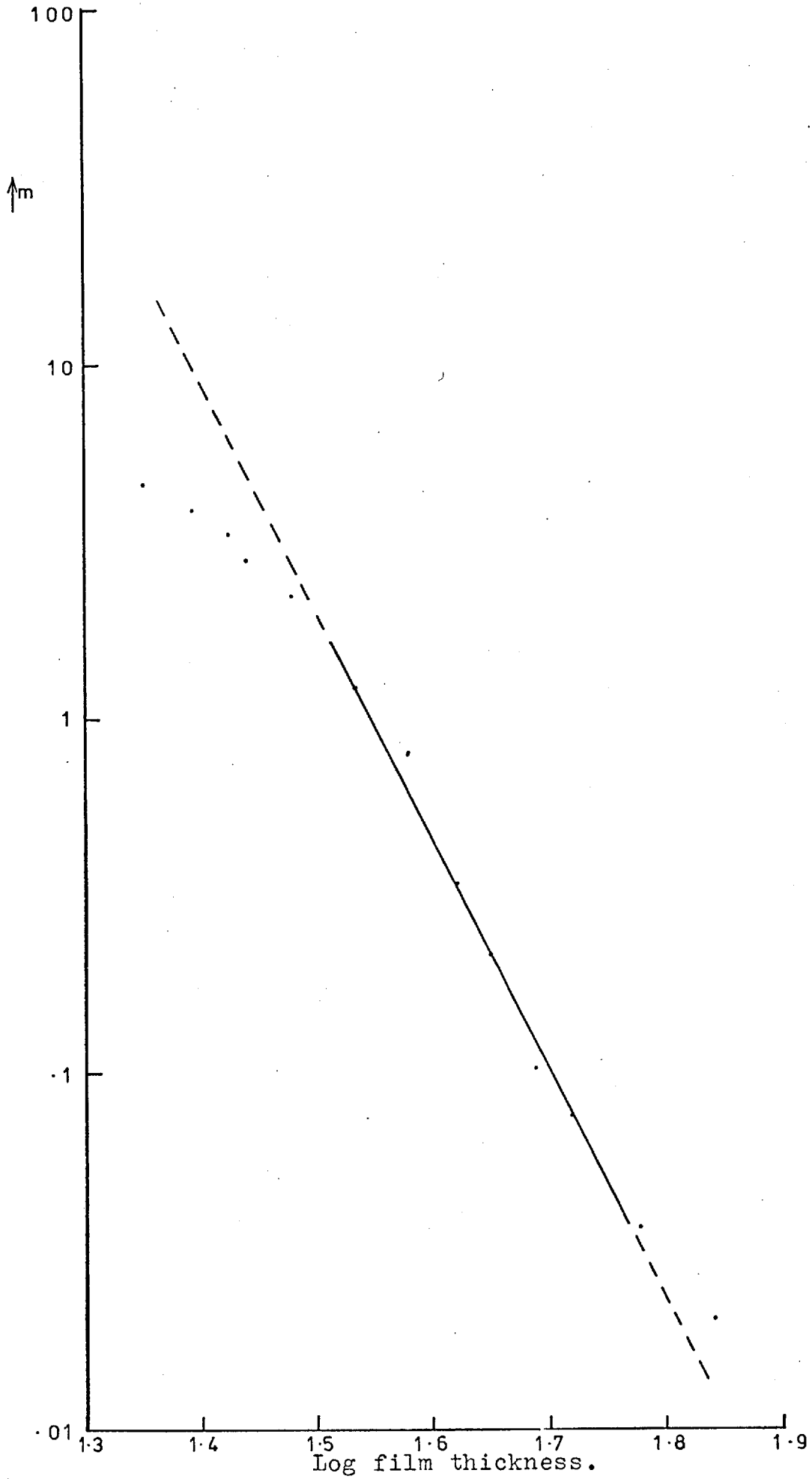


Fig. 9.22 Variation of film thickness with m (constant τ).

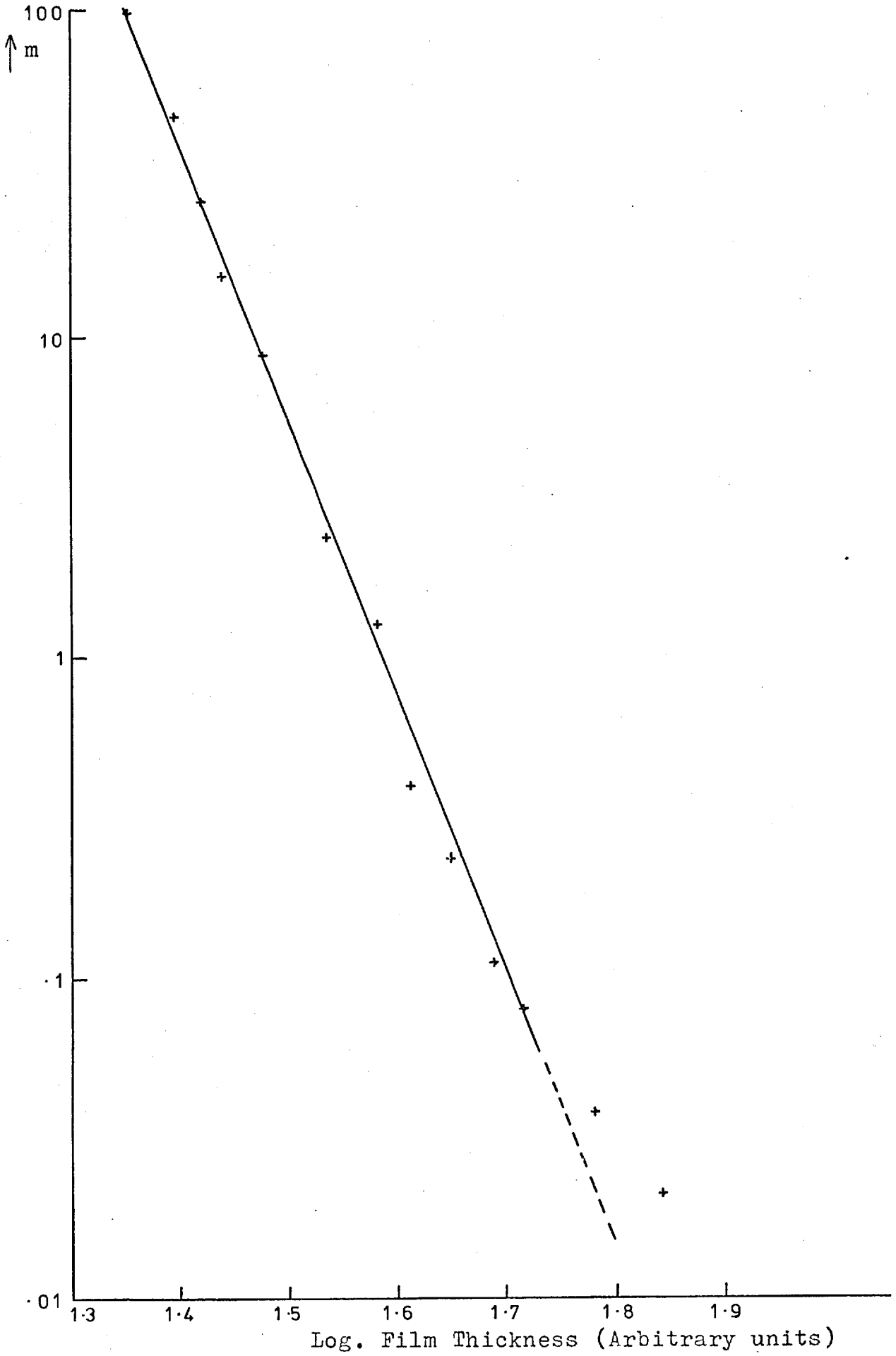


Fig. 9.23 Variation of m with film thickness. (exponential)

9.15 Point contact-rolling.

The experimental results of Tallian (Fig. 2.13) for a 100lb load show a peak in count rate at $\gamma \simeq 0.63$. The results are not sufficiently accurate to indicate which distribution is the most suitable for analysis. Both the circular distribution and the constant τ models have a gentle peak and small errors in measurement can shift the measured peak significantly. This could be the explanation of the earlier peak encountered in the case of the 50lb load (Fig. 2.12). From Fig. 9.9, the count rate for the circular distribution is about 1.3 times greater than that for the constant τ model, for γ from 0 to 0.5. The deviation from linearity is less than 2% over this region. Thus by choosing $\tau' = \frac{2r}{1.3} \simeq 1.54r$ the simpler model will describe this part of the distribution quite well. It is not surprising that this value of τ' is close to the mean duration for the circular distribution which is $\frac{\pi r}{2} \simeq 1.57r$. The difference between count rates increases gradually with γ, reaching about 1.5 for $\gamma = 0.95$. After this point the difference is more marked. However, for a large range of γ the distributions are very similar and for simplicity of working the results that follow have been treated according to the constant τ model where $m = -\ln(1 - \gamma)$. Fig. 9.24 shows Tallian's experimental points for 100lb load. The equation $\frac{h}{\sigma} = \frac{5}{3}m^{-0.242}$ yields values of film thickness within 3% of the experimental points. Fig. 9.25 is a replot of data taken from Tallian et al. (42) for 182lb load. The equation of the line in this case is $\frac{h}{\sigma} = 1.76m^{-0.245}$ Tallian observed that in some cases γ varied with applied

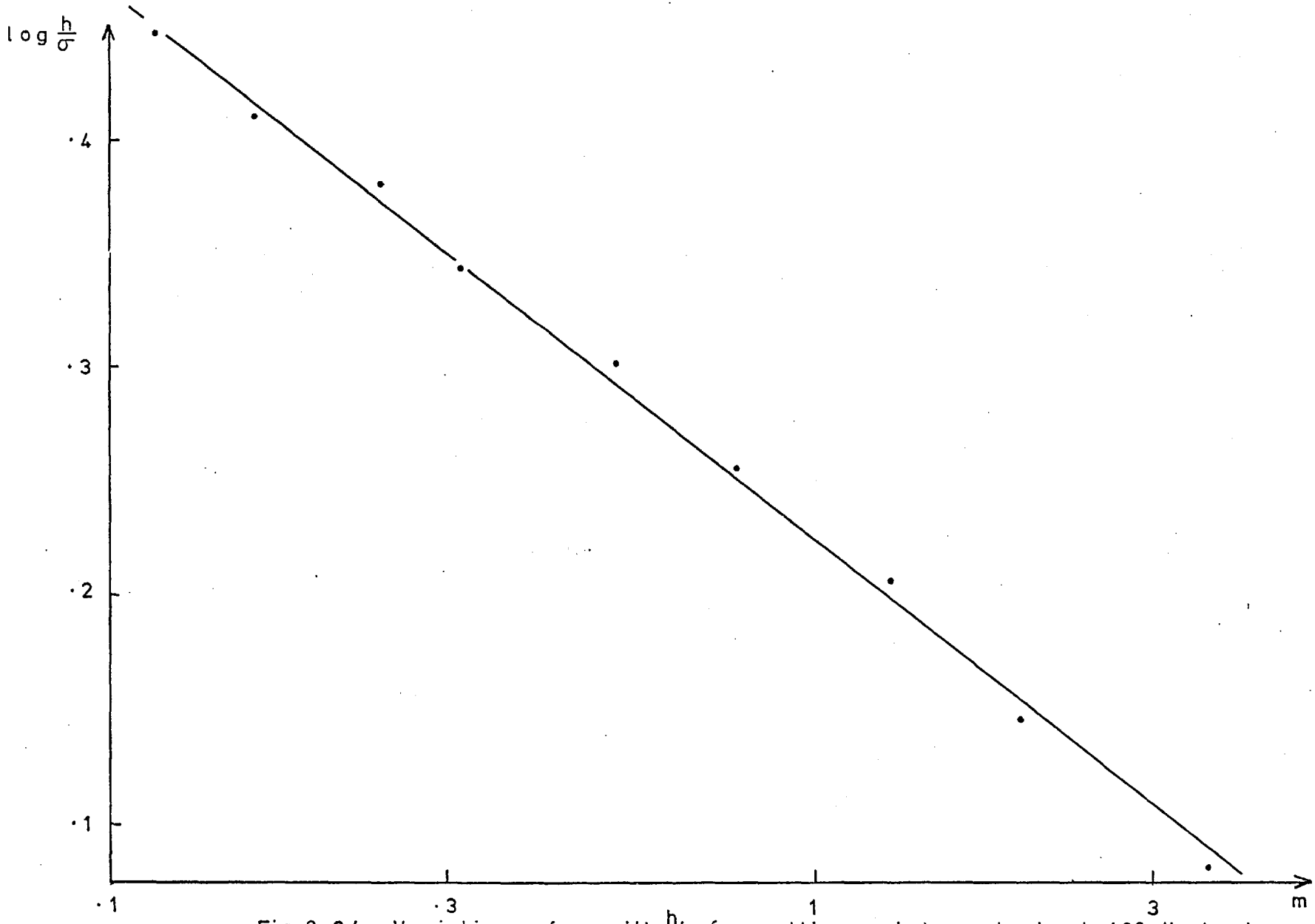


Fig. 9.24 Variation of m with h/σ for rolling point contact at 100 lb. load.

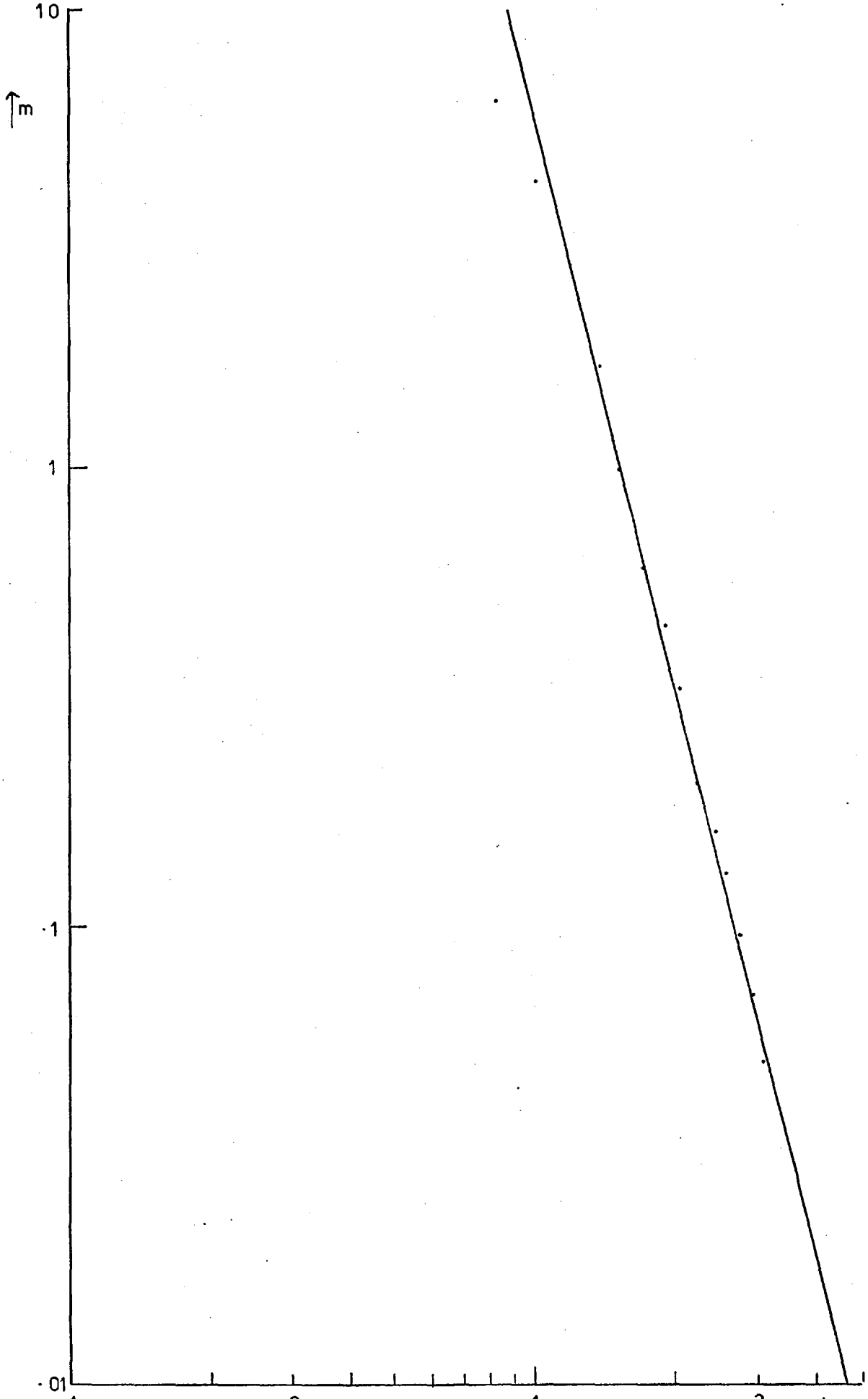


Fig. 9.25 Variation of m with h/σ for rolling point contact at 1821b load.

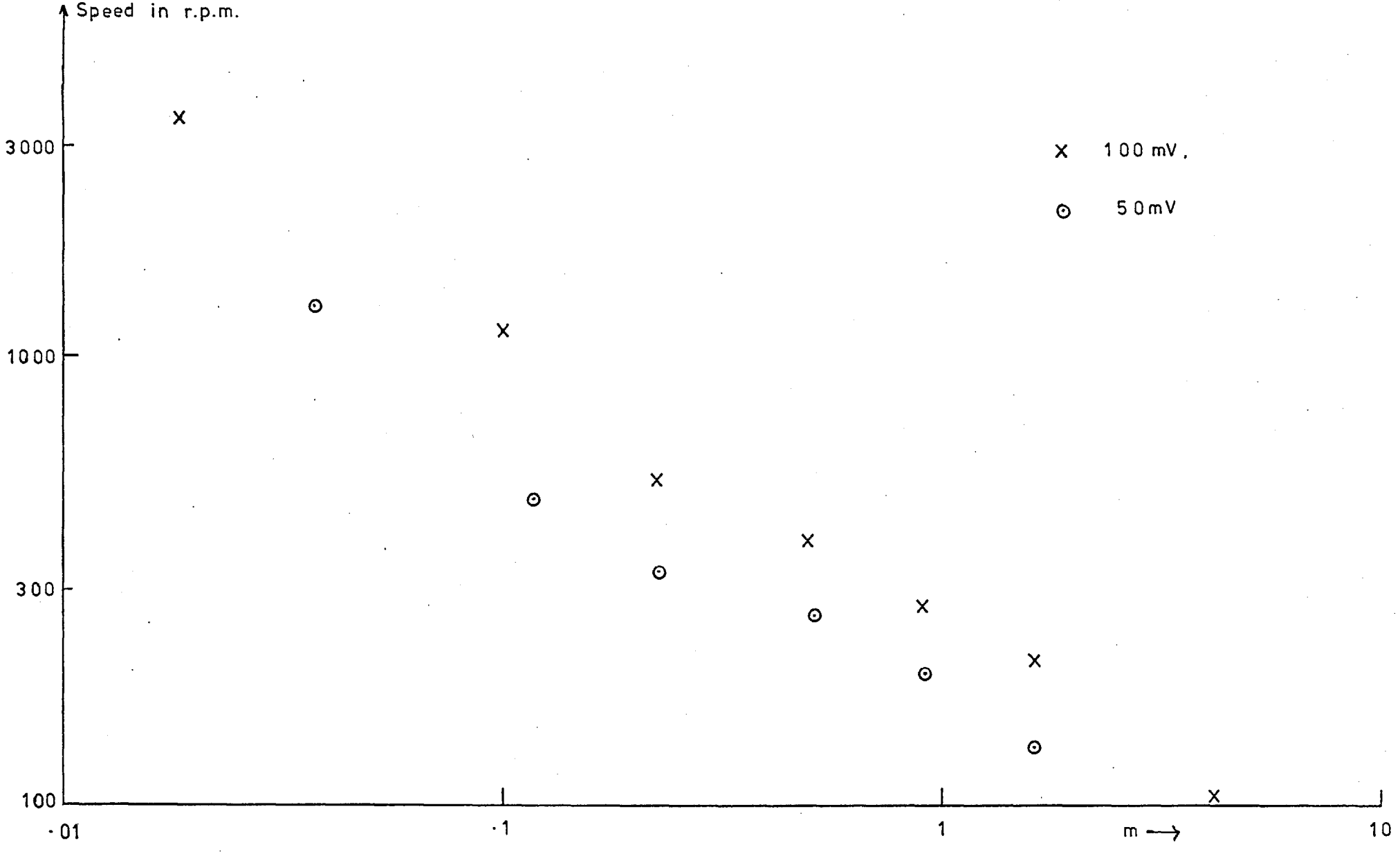


Fig.9.26 The effect of applied potential on m.

potential. Fig. 9.26 shows the effect of halving the normal applied potential to 50mV. Although linearity is not apparent in this case, the m values are consistently two to three times lower with the lower potential. The load in this experiment was 50lb.

9.16 Interpretation of results.

The results of section 9.9 to 9.15 are now colla ted. The variation of m with film thickness is shown in Fig. 9.27.

Experimental configuration		Exponent of m	Comment
1. Roller bearing		- 0.17	Electrical
		- 0.25	measurements across inner and outer races.
2. Tapered roller bearing	to	- 0.42	γ corrected for one
		- 0.74	race and rollers contacting. Load dependent.
3. Rolling line contact		- 0.53	Exponential model.
		- 0.55	Constant τ model.
4. Sliding line contact		- 0.08	
5. Sliding point contact		- 0.27	
6. Rolling/sliding point contact		- 0.12	
7. Rolling point contact		- 0.24	100lb. load.
		- 0.27	182lb. load.

Fig. 9.27 Dependence of m on film thickness.

The data on load dependence are not explicit in most cases discussed above. From Garnell's data a variation of $m \propto W^{\frac{1}{2}}$ may be inferred from the constancy of $R(W)^{\frac{1}{2}}$. From Furey's data $m \propto W^{1.45}$. Other experimental work does not indicate a simple relationship between m and load.

The exponents of m in Fig. 9.27 vary by an order of magnitude. The rolling and sliding results for point contact show similar variation although there are major differences in the experimental variables and the similarity may be due to the interaction of several factors. One of the surfaces used by Furey was ground to a finish of about 10μins. whereas Tallian used superfinished specimens. The load in these two experiments differed by two orders of magnitude.

The data of Czichos for sliding line contact and of Poon and Haines for rolling with sliding in a point contact show a stronger m dependence on film thickness than any of the other experimental arrangements considered.

There is no obvious connection between the exponent of m and the roughness of the surfaces used by different workers. For instance, Christensen, Tallian and Poon used very smooth surfaces while those of Furey and Czichos were considerably rougher. Leaver et al. have stressed the significance of filtering data obtained from surface profiles in order to include only those frequencies relevant to the appropriate contact width. Another important point in boundary lubrication must be the effect of part of the nominally contacting surfaces being altogether

free of any hydrodynamic pressure. In this case the deformation of the surface should perhaps follow the laws for dry contact. With very rough surfaces the curvature of the dominant features may be sufficiently high for the independent generation of lubricant films. In this "mini-EHL" process the electrical behaviour would be dominated by the finer surface features or "micro-asperities". This process may qualitatively explain the results of Furey in Fig. 2.3.

Electrical behaviour is dominated by the peaks on the contacting surfaces. Where the surface is random the peak height distribution is simply related to the height distribution. Ground surfaces are not Gaussian in nature and the heights of the major peaks may well have less scatter than one would expect from the c.l.a. roughness value. A small reduction in film thickness would then bring many more points into contact than with a random surface. This could explain the low exponent of m found in Czichos' results, which can be rewritten $m \propto \left(\frac{h}{\sigma}\right)^{-12.5}$. At low film thicknesses most of the major peaks would be in contact so that a further reduction in thickness would cause m to vary more slowly than previously. This may explain the deviation of the rough surface behaviour found in Fig. 9.19 at low film thicknesses and that of the highest load in Fig. 9.20 although part of this could well be caused by experimental error. There is no reason to expect a simple power law dependence of m on oil film thickness but the linearity of Figs. 9.20, 9.23, 9.24 and 9.25 is remarkable.

At the beginning of this chapter it is argued that for rolling contact $m \propto Area$. For sliding contact the duration of a contact is determined by the time taken for asperities to slip past each other. From section 4.9 this time appears to be a function of film thickness and surface topography. If the surface roughness is unchanged by load and the film thickness virtually unchanged then this time, τ_m, will be constant with load. The number of contact occurrences C', should be proportional to area. Then $C'\tau_m \propto W^n$ where $n = \frac{2}{3}$ for point contact and $\frac{1}{2}$ for line contact. The value of n found from Furey's experiments is about twice that expected. The conclusion from this must be that the available number of contact points is increased by elastic deformation of the surfaces.

Clearly there are important differences between all the experiments discussed. The wide variation of electrical behaviour reflects this although, at this stage, defying definitive interpretation. The direct use of the approach discussed in this chapter is in monitoring conditions in the test machines that show the best correlations. For instance, the equation $\frac{h}{\sigma} = \frac{5}{3}m^{-0.25}$ gives a measure of film thickness to within 18% for all of Tallian's published results discussed above. In most cases the agreement is within 10%.

Another possibility is the monitoring of surface topography changes during running by measuring γ and C'_m and using the equations in Fig. 9.10 to find τ_m.

The passage of wear particles and any other

conducting material through the contact would be reflected by anomalously high values of γ and low values of C'_m. If a wear factor is interpreted as the probability that an asperity collision produces a wear particle, then an estimate of the rate of collisions occurring throughout the contact given by $\frac{C'_m}{1 - \gamma}$ will no doubt be of use.

Chapter 10. Conclusions.

10.1 Introduction.

A representative review is given in chapter two of the research that has been effected into electrical resistance measurements applied to lubricated contacts. From this it is clear that such measurements are extremely sensitive to small changes inside the contact. Tallian's work (42) best exemplifies the critical dependence of contact resistance on separation. For $\gamma = \frac{1}{2}$ in Fig. 2.13, a 10% change in γ corresponds to a film thickness change of about 50Å. The experimental work in chapter five indicates that chemical effects may also be detected by measurement of γ. A definitive interpretation of electrical measurements in a lubricated contact must await better understanding of conduction mechanisms through oils and the electrical characteristics of particular oil and additive combinations. This does not, however, exclude the possibility of a relative quantitative approach to the interpretation of electrical data. The drop to the low resistance state may perhaps be interpreted as corresponding to a critical separation between the two nominally contacting surfaces. The fact that the number of these contacts varies with external circuitry may not be too important. If a change in the series resistor is reflected by an alteration in the critical separation, one might expect to observe the same relative changes in the electrical signal, properly interpreted, caused by a changing environment. Prior to this work no basis for the direct interpretation of the electrical resistance signal has been formulated, even on a relative scale.

The model proposed by Johnson et al. (52) has been discussed at length in chapter four. It is instructive to examine some of the physical assumptions made in this paper. They are briefly:

1. flat parallel Hertzian area.
2. no change in surface topography except at contacting asperities, i.e. no interaction between adjacent asperities.
3. Gaussian distribution of surface heights. From this it follows that peak heights are distributed normally and that both distributions are related.
4. a law relating the contacts observed to those actually occurring taking into consideration multiple and overlapping contacts. The relation $\gamma = 1 - e^{-m}$ was assumed.

Of these 1 above must be suspect particularly at high loads. The constriction region in both point and line contacts must be expected to play an important part in electrical conduction as it represents areas of closest approach of the two surfaces.

Little is known about the elastic behaviour of rough surfaces in lubricated contact and the supposition of constant surface topography except at micro-contact points is as good a guess as any. The investigation of behaviour different from 2 above is a daunting proposition.

Assumption 3 is reasonable for surfaces finished by processes involving random action, for example lapping or

bead blasting. It is not reasonable for other finishing processes such as grinding or milling. It is less restricting to assume that the peak height distribution is Gaussian, although this may not be simply related to the distribution of all surface heights.

10.2 Basis of new approach.

The main premise for the new approach described in chapter nine is that asperity collisions occur at random. This may be unreasonable for a rolling surface finished by grinding transverse to the direction of rolling, but the situation becomes more complicated when two such surfaces interact. A certain amount of support for this idea is given by the experimental work in chapter eight. Once this assumption is made and a suitable distribution identified, the result of equation 9.4 is immediately obtained.

No explicit behaviour of surfaces in contact is assumed. The distributions of contact duration that are examined in sections 9.4, 9.5 and 9.6 must however be justified in physical terms. Thus the two rolling models stem from similar ideas to those expressed in 1 and 2 of the preceding section. The sliding model distribution of dwell times was obtained in section 4.7 for nominally flat rough surfaces. An exponential distribution fits Christensen's measurements (Fig. 9.11) of dwell time extremely well, close to the origin.

No mechanism of physical contact is assumed and the model is valid for as long as the distribution of dwell times is exponential, even if the parameters of the distribution should vary.

10.3 Applications of the new model.

When the most suitable of the three special cases described in chapter nine has been chosen, it is a simple matter to calculate the average number of contact points, m, from a measurement of the contact time fraction, γ. The examples in chapter nine show the results of this for some published research. It is remarkable that many of the examples show simple power law dependence of m on film thickness, speed or load. Perhaps even more remarkable is the range of values which has been obtained for the exponent of film thickness, x, in the equation $m \propto h^x$. Values of x lie between 1.3 and 12.5. It is likely that this large discrepancy must reflect important differences in the manner in which the various types of surface behave under increasingly severe conditions. A qualitative argument is put forward in section 9.16 for the high value of x obtained from Czichos' work (49), but it is clear that a great deal of work in this direction is needed.

Apart from such fundamental work on the nature of the contact of lubricated surfaces a much more practical application is evident. The excellent correlation between values of m and some simple power law dependence on film thickness evinced in chapter nine suggests that this approach could be used with considerable accuracy to monitor film thickness "on-line". In this respect it is extremely satisfying that the results for changes in applied potential difference or acceptable contact resistance shown in Figs. 9.19 and 9.26 reveal very similar dependence on speed. It would, of course, be necessary to calibrate the system

beforehand for a range of loads and probably for every oil and additive combination.

The load dependence of m that one would expect from simple premises ignoring oil film thickness changes is that $m \propto$ (nominal area of contact). The data of Furey (9.20) show that $m \propto W^{1.45}$. The difference between this and the W dependence on nominal contact area must be ascribed to changes in film thickness or surface topography or both. Indeed, if the exponent x has a high value, the variation of m with load may be more dependent on changes of film thickness rather than nominal contact area.

Another application for this new approach is in the field of wear. Archard (117) has interpreted the wear coefficient as the probability of an asperity collision forming a wear particle. In support of this, values of the wear coefficient greater than unity are not found (118). Equation 9.4 allows, for the first time, an estimate to be made of the number of asperity collisions occurring throughout the contact area. If a correlation can be found between actual contact rate and wear rate, an electrical method of estimating wear emerges.

Reasonable agreement has been found in chapter nine between the measured mean duration of electrical contacts and that predicted from simple model surfaces undergoing sliding. This suggests the possibility that, under otherwise steady operating conditions, modification of surface topography may be detectable while running tests. This could be done in conjunction with the proposed wear estimation technique.

10.4 Suggestions for further work.

There are three main limitations to the use of electrical resistance techniques in lubrication studies. First, the lack of understanding of conduction mechanisms inside typical lubricated contacts makes direct physical interpretation of resistance measurements impossible. Second, the passage of wear particles and other debris through the Hertzian zone causes spurious contacts to register. Third, the limited frequency response of the system as a whole reduces the accuracy of measurement of count rate and contact time, particularly when sliding behaviour dominates. These are three areas for further study. Basic conduction mechanisms may best be studied by slow speed work similar to that of Tallian (46). A certain amount of information may be gained on a more practical time scale by the study of a system with variable discriminator levels and controlled contact current. This is discussed in chapter three (p.38).

Careful design of electronic equipment will extend the frequency response of measuring equipment into the 10MHz region without difficulty, although the electrical properties of the test elements must also be considered. Ultimately, a theoretical correction will have to be applied to the known response of any measuring equipment in order to estimate the frequency of events occurring outside the measurable range.

The problem of spurious contacts caused by extraneous matter in the test oil is more difficult to circumvent. That such events occur seems likely from

observations of dwell times much longer than the Hertzian transit time (46 and 48) even under conditions of rare electrical contact. At the moment it is not known how seriously this problem restricts the accuracy of measurement. If found to be important, a possible solution might be an electronic device which rejects any dwell times that are statistically unlikely to be caused by multiple contacts.

The four applications described in the preceding section are all worthy of further study. They are again, briefly:

1. Study of the behaviour of the number of contact points, m, under a wide range of load and speed for different surface finishes.
2. Examination of the accuracy of film thickness estimation by the methods outlined in chapter nine.
3. Wear rate estimation.
4. Study of mean contact duration and surface topography under sliding conditions.

In connection with the sputtering work described in chapter seven, a line of development was briefly pursued that is of interest. The aim was to produce model rough surfaces with controlled wavelength and height. This was performed by sputtering chromic oxide through readily available grids on to a very flat glass substrate. The random component of movement of depositing species ensured that the features produced would not be sharp sided, but have gentle slopes. Fig. 10.1 shows Talysurf traces of

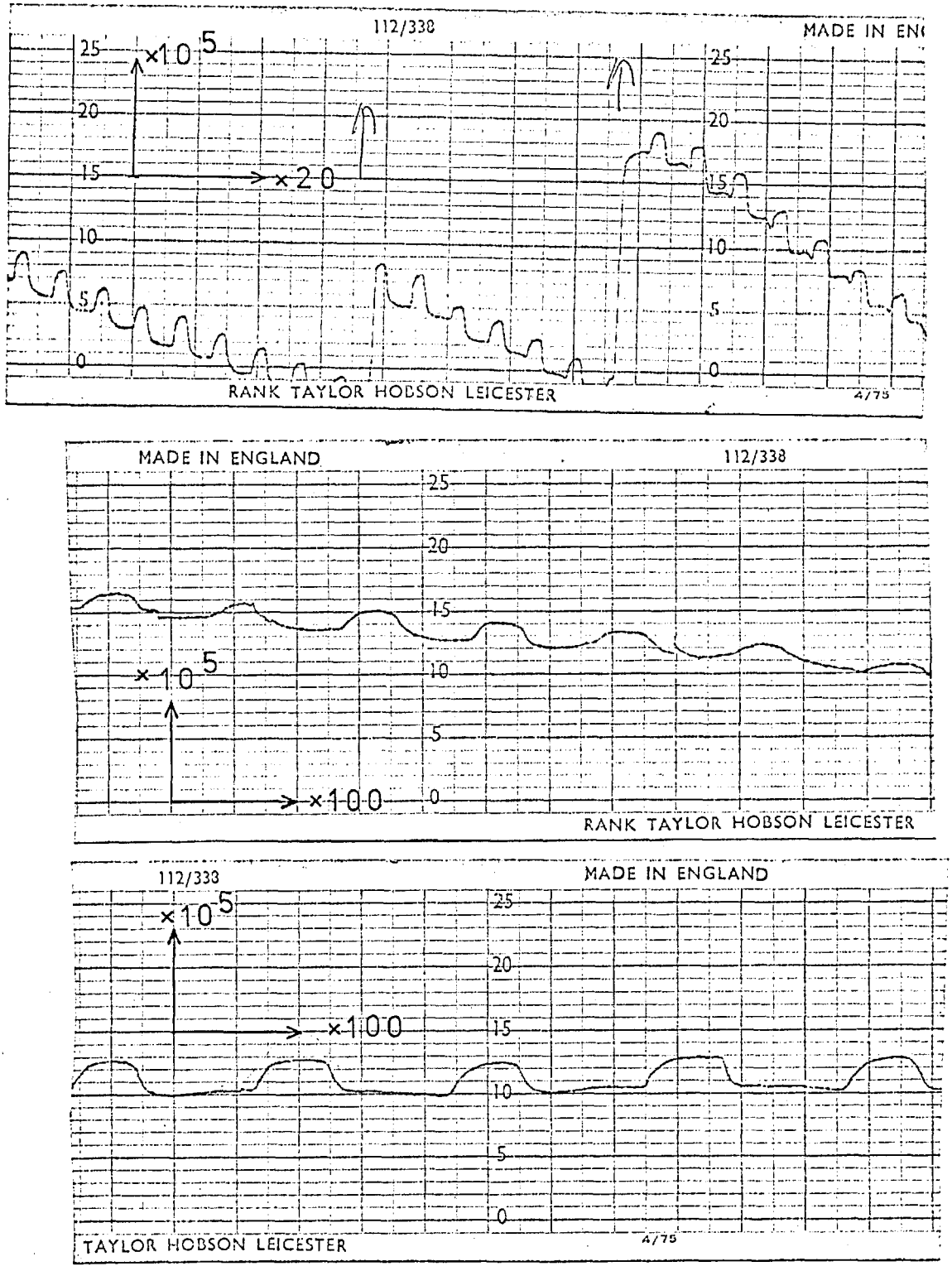


Fig. 10.1 Talysurf traces of model surfaces.

surfaces prepared using coarse grids and finer surfaces with a peak separation of about 60μm. were also produced.

These surfaces could be used to study basic rough surface phenomena, for instance, the effect of traction on surface roughness or the effect of asperity height or spacing separately on oil film thickness. These model surfaces are most suitable for use in conjunction with optical interferometry as their regular spacing allows interference fringes to be identified. This is not possible with practically prepared rough surfaces which have a random structure. These surfaces might also be used to examine the effect of surface roughness on adhesion as reported by Fuller and Tabor (119).

10.5 Main contribution of this work.

A basis has been proposed and developed for the interpretation of electrical resistance measurements in lubricated contacts. No attempt has previously been made to interpret such an electrical signal directly. The results of applying this theory to published results show great practical significance, and it may be used to study the nature of the contact of lubricated rough surfaces on a scale that has not previously been possible.

Many of the suggested areas for future research are now being pursued.

Appendix A. The distribution of intervals between C random events occurring in time t.

This distribution was analysed by a numerical method using the random number generator of a digital computer. In each trial a hundred numbers between zero and unity were evolved and then placed in ascending order. An interval was obtained by subtracting each of these numbers from the succeeding one. For a total of four thousand trials the probability p(N) that an interval lies in the Nth band of width 0.01 was obtained. To test for an exponential distribution N was plotted against logp(N) (Fig. A.1). Clearly this is a good approximation.

A somewhat more rough and ready method was used to check the variation of C_m with τ'. A series of 250 four digit pseudo-random numbers was obtained from logarithms of gravimetric constants of the elements. These were arranged in ascending order and the number of non overlapping events (C_m) was tabulated as τ' was increased. In this example $t = 10{,}000$, $C = 250$ and τ' was increased in units of four. These results are shown in Fig. A.2. The line represents $C_m = 250\exp(-\frac{\tau'}{40})$.

It is clear from these examples that an exponential distribution of intervals is a good approximation if C is sufficiently large and events occur at random.

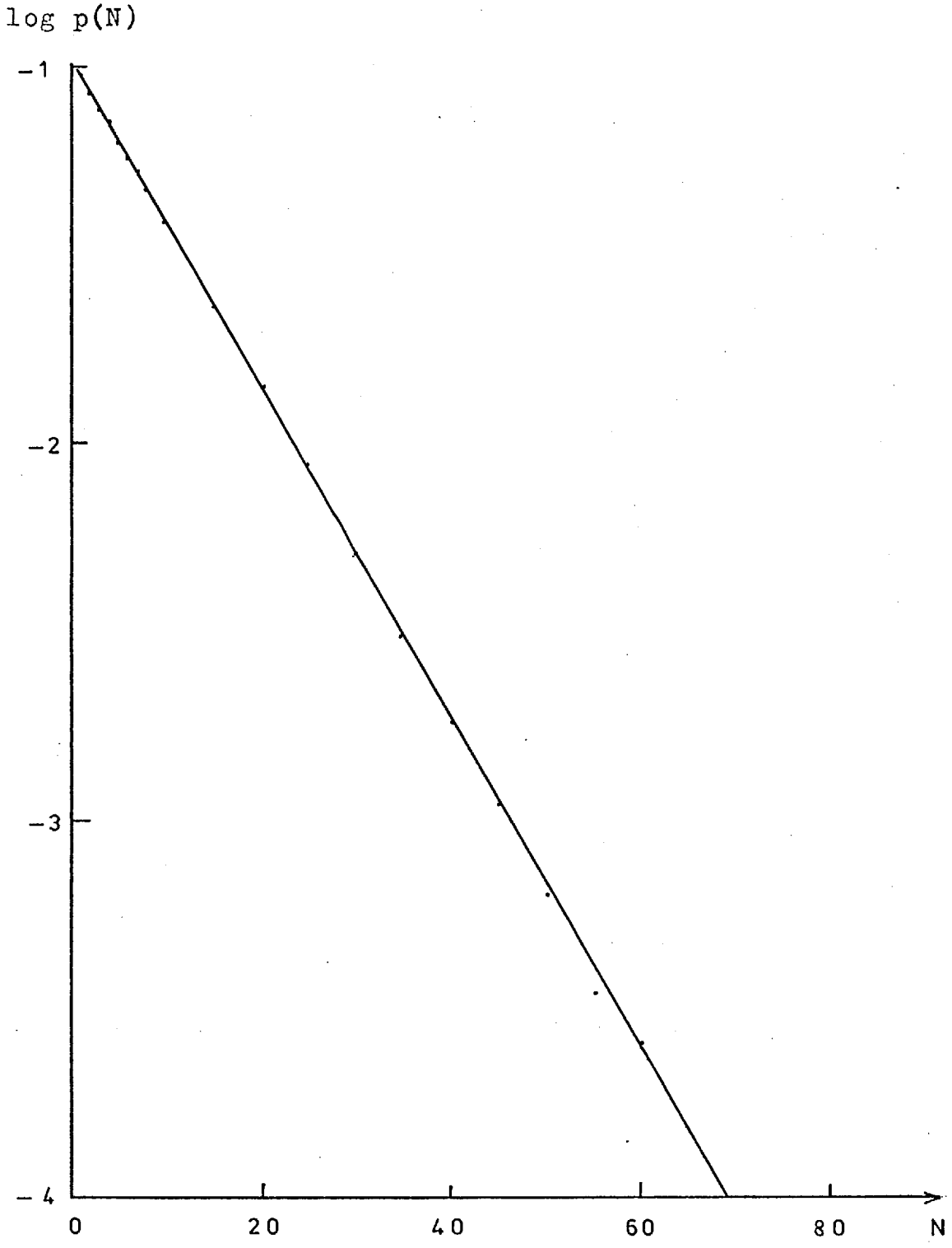


Fig. A.1 Distribution of intervals between random events.

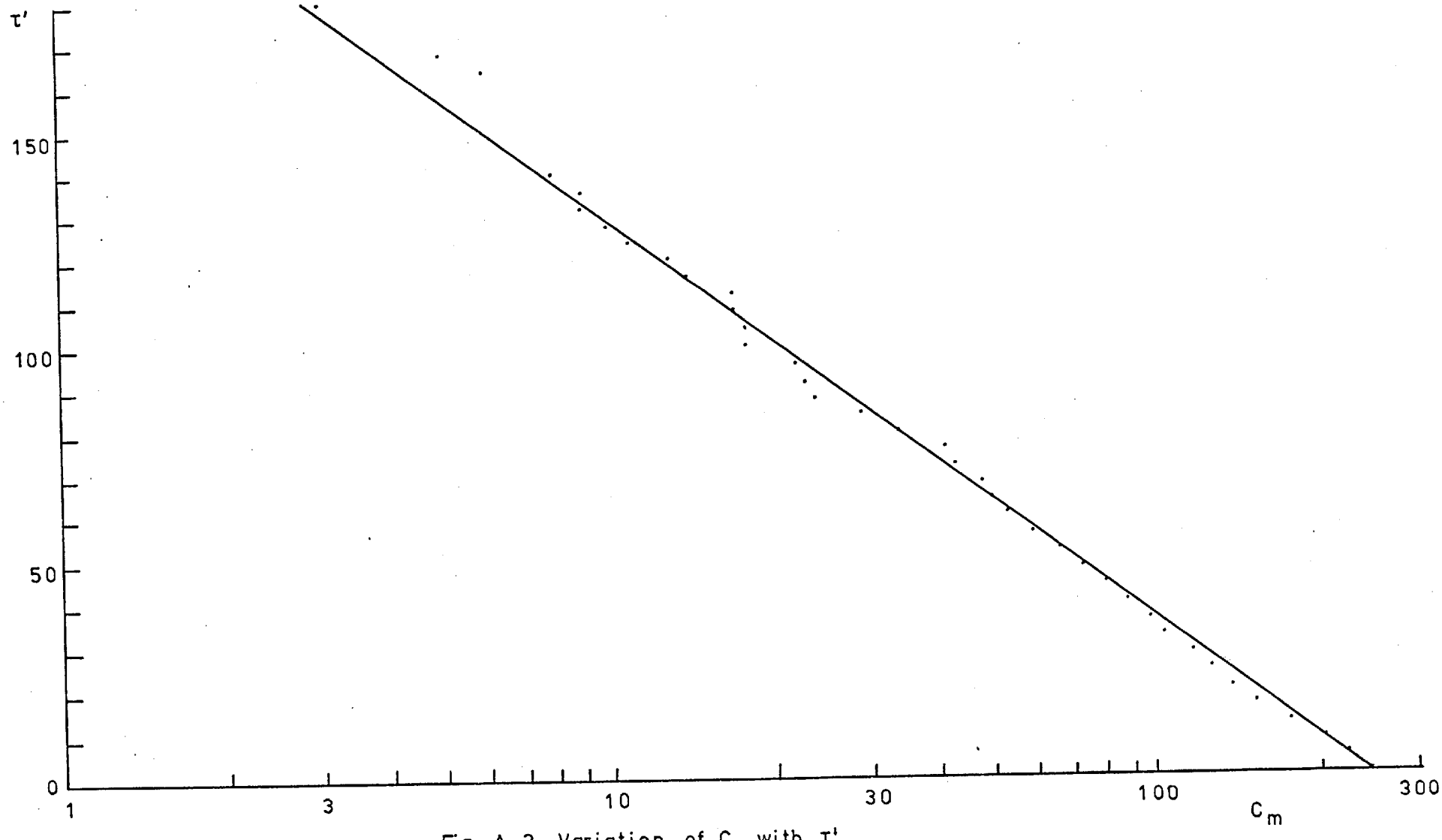


Fig. A.2 Variation of C_m with τ'.

Appendix B. Evaluation of integral $\int_o^1 \exp(-Am)(1-m^2)^{\frac{1}{2}}dm$.

A series solution is adopted for this integral as simpler methods do not suffice. If the above integral is denoted by I(A) and the exponential term is expanded,

then $$I(A) = \int_o^1 \left(1 - Am + \frac{(Am)^2}{2!} - \frac{(Am)^3}{3!} + \ldots\right)(1-m^2)^{\frac{1}{2}}dm.$$

If these integrations are performed separately the (n + 1)th term is given by $I_n(A)$ where

$$I_n(A) = \int_o^1 \frac{(-Am)^n}{n!}(1-m^2)^{\frac{1}{2}}dm.$$

Now put $m = \sin u$, then $dm = \cos u.\, du$ and

$$I_n(A) = \frac{(-A)^n}{n!}\int_o^{\pi/2} \sin^n u \cos^2 u\, du$$

$$= \frac{(-A)^n}{n!}\int_o^{\pi/2} (\sin^n u - \sin^{n+2} u)\, du \qquad B.1$$

The integral $\int_o^{\pi/2} \sin^n u\, du$ is a standard form given by

$$\frac{(\pi)^{\frac{1}{2}}\Gamma\left(\frac{n+1}{2}\right)}{2\Gamma\left(\frac{n+2}{2}\right)} \text{ for } n>-1 \qquad B.2$$

The Gamma function $\Gamma(n)$ is defined as

$\Gamma(n) = \int_o^\infty \exp(-x)x^{n-1}dx$ and has the property that

$$\Gamma(n+1) = n\Gamma(n) \qquad B.3$$

Thus if n is a positive integer then

$$\Gamma(n) = (n-1)!.$$

All other relevant values of the Gamma function may be found using the recurrence relation defined above and the fact that

$$\Gamma(1.5) = \frac{\pi^{\frac{1}{2}}}{2}.$$

From equations B.1 and B.2,

$$I_n(A) = \frac{(-A)^n}{n!}\left(\frac{\pi^{\frac{1}{2}}}{2}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)} - \frac{\pi^{\frac{1}{2}}}{2}\frac{\Gamma\left(\frac{n+3}{2}\right)}{\Gamma\left(\frac{n+4}{2}\right)}\right)(n>-1)$$

$$= \frac{(-A)^n}{n!}\frac{\pi^{\frac{1}{2}}}{2}\left(\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)} - \frac{\frac{n+1}{2}\Gamma\left(\frac{n+1}{2}\right)}{\frac{n+2}{2}\Gamma\left(\frac{n+2}{2}\right)}\right).$$

Applying B.3 to the second term,

$$I_n(A) = \frac{(-A)^n}{n!}\frac{\pi^{\frac{1}{2}}}{2}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)}\cdot\frac{1}{n+2}.$$

If both denominator and numerator are now multiplied by a factor of $\frac{n+1}{2}$, then

$$I_n(A) = \frac{\pi^{\frac{1}{2}}}{2}\frac{(-A)^n}{(n+1)!}\cdot\frac{\Gamma\left(\frac{n+3}{2}\right)}{\Gamma\left(\frac{n+4}{2}\right)} \qquad \text{B.4}$$

If the coefficient of A^n is written as C_n, then

$$I_n(A) = C_n A^n.$$

Finally summing all the I_n

$$I(A) = \sum_0^\infty I_n = \sum_0^\infty C_n A^n.$$

The first forty-two values of C_n are tabulated in Fig. B.1.

From equation B.4, $C_o = \frac{\pi^{\frac{1}{2}}}{2} \frac{\Gamma(1.5)}{\Gamma(2)}$

Now $\Gamma(1.5) = \frac{\pi^{\frac{1}{2}}}{2}$ and $\Gamma(2) = 1$, thus $C_o = \frac{\pi}{4}$.

Similarly $C_1 = -\frac{\pi^{\frac{1}{2}}}{2x2!} \frac{\Gamma(2)}{\Gamma(2.5)}$

$$= \frac{\pi^{\frac{1}{2}}}{4x1.5\,\Gamma(1.5)} = -\frac{1}{3}.$$

It is easily shown that $C_{n+2} = \frac{C_n}{(n+2)(n+4)}$ and subsequent values are generated, using this relation, from C_o and C_1.

To obtain the values of the integral I(A) it is now only necessary to find the sum $\sum_{o}^{\infty} C_n A^n$. The coefficients of A_n have been found to nine figure accuracy and it would lead to error to use these values for A greater than about ten. Once the maximum value of $C_n A^n$ has been reached the terms decay fairly rapidly. The table (Fig. B.2) shows the value of I(A) for some chosen values of A. As A tends to zero, I(A) tends to $\frac{\pi}{4}$.

$C_0 = 7.85398164 \times 10^{-1}$ $C_1 = -3.33333333 \times 10^{-1}$

$C_2 = 9.81747704 \times 10^{-2}$ $C_3 = -2.22222222 \times 10^{-2}$

$C_4 = 4.09061543 \times 10^{-3}$ $C_5 = -6.34920635 \times 10^{-4}$

$C_6 = 8.52211549 \times 10^{-5}$ $C_7 = -1.00781053 \times 10^{-5}$

$C_8 = 1.06526444 \times 10^{-6}$ $C_9 = -1.01799043 \times 10^{-7}$

$C_{10} = 8.87720363 \times 10^{-9}$ $C_{11} = -7.11881422 \times 10^{-10}$

$C_{12} = 5.28404978 \times 10^{-11}$ $C_{13} = -3.65067400 \times 10^{-12}$

$C_{14} = 2.35895079 \times 10^{-13}$ $C_{15} = -1.43163685 \times 10^{-14}$

$C_{16} = 8.19080136 \times 10^{-16}$ $C_{17} = -4.43231222 \times 10^{-17}$

$C_{18} = 2.27522260 \times 10^{-18}$ $C_{19} = -1.11085519 \times 10^{-19}$

$C_{20} = 5.17096045 \times 10^{-21}$ $C_{21} = -2.29990723 \times 10^{-22}$

$C_{22} = 9.79348571 \times 10^{-24}$ $C_{23} = -3.99983865 \times 10^{-25}$

$C_{24} = 1.56946886 \times 10^{-26}$ $C_{25} = -5.92568690 \times 10^{-28}$

$C_{26} = 2.15586382 \times 10^{-29}$ $C_{27} = -7.56792707 \times 10^{-31}$

$C_{28} = 2.56650455 \times 10^{-32}$ $C_{29} = -8.41816137 \times 10^{-34}$

$C_{30} = 2.67344224 \times 10^{-35}$ $C_{31} = -8.22889674 \times 10^{-37}$

$C_{32} = 2.45720794 \times 10^{-38}$ $C_{33} = -7.12458592 \times 10^{-40}$

$C_{34} = 2.00752283 \times 10^{-41}$ $C_{35} = -5.50161075 \times 10^{-43}$

$C_{36} = 1.46748745 \times 10^{-44}$ $C_{37} = -3.81262006 \times 10^{-46}$

$C_{38} = 9.65452266 \times 10^{-48}$ $C_{39} = -2.38437777 \times 10^{-49}$

$C_{40} = 5.74673968 \times 10^{-51}$ $C_{41} = -1.35245478 \times 10^{-52}$

Fig. B.1 Coefficients in integrals I_n.

A	I(A)
.01	.7821
.10	.7530
.20	.7225
.30	.6937
.50	.6407
.80	.5718
1.00	.5315
1.35	.4708
1.40	.4630
1.50	.4480
2.00	.3832
5.00	.1909
10.00	.09896

Fig. B.2 The integral I(A).

References.

1. Needs, S.J. "Boundary film investigation." Trans. ASME 62 1940 331.
2. Beeck, O., Givens, J.W. and Smith, A.E. "On the mechanism of boundary lubrication." Proc. Roy. Soc. A177 1940 90.
3. Courtney-Pratt, J.S. and Tudor, G.K. "An analysis of the lubrication between the piston rings and cylinder wall of a running engine." Proc. I. Mech. E. 155 1946 293.
4. Tudor, G.K. "An electrical method of investigating the lubrication in a journal bearing." J. Council for Scientific and Industrial Research (Australia) 21 1948.
5. Lewicki, W. "Some physical aspects of lubrication in rolling bearings and gears." Engineer 200 1955 176-178 and 212-215.
6. Brix, V.H. "An electrical study of boundary lubrication." Aircraft Engineering 19 1947 294.
7. Cameron, A. "Surface failure in gears." J. Inst. Pet. 40 1954 191.
8. Siripongse, C., Rogers, P.R. and Cameron, A. "Thin film lubrication." Engineering 186 1958 146.
9. Ibrahim, M. and Cameron, A. "Oil film thickness and the mechanism of scuffing in gear teeth." Conference on lubrication and wear, I. Mech. E. 1963 Paper 20.
10. O'Donoghue, J.P. and Cameron, A. "Temperature at scuffing." Proc. I. Mech. E. 180 (3B) 1965-66 85.
11. Chu, P.S.Y. and Cameron, A. "Flow of electrical current through lubricated contacts." Trans. ASLE 10 1967 226.

12. Bathgate, J. and Yates, F. "The application of film thickness flash temperature and surface fatigue criteria to work gears." Trans. ASLE 13 (1) 1970 21.

13. Dyson, A. Thornton Research Centre, Shell Report R1215 1962.

14. Dyson, A. "Investigation of the discharge-voltage method of measuring the thickness of oil films formed in a disc machine under conditions of EHL." Proc. I. Mech. E. 181 Pt 1 (23) 1966-67 633.

15. Twisleton-Wykeham-Fiennes, W.G. "A study of the electrical conductivity of lubricating oil films." Ph.D. Thesis, University of London 1971.

16. Fiennes, W.G. and Anderson, J.C. "An analysis of voltage discharge measurements in lubrication research." I. Mech. E. Tribology Convention 1972 Paper 112/72 109.

17. Brix, V.H. Private communication.

18. Furey, M.J. "Metallic contact and friction between sliding surfaces." Trans. ASLE 4 1961 1.

19. Furey, M.J. "Surface roughness effects on metallic contact and friction." Trans. ASLE 6 1963 49.

20. Furey, M.J. and Appeldoorn, J.K. "The effect of lubricant viscosity on metallic contact and friction in a sliding system." Trans. ASLE 5 1962 149.

21. Christensen, H. "Some experiments on 'running-in' and scuffing of discs." Acta. Polytech. Scand. Mech. Eng. Ser. 25 1966.

22. Millns, J.E.M. Private communication.

23. Garnell, P. and Higginson, G.R. "The mechanics of roller bearings." Proc. I. Mech. E. 180 (3B) 1965-66 197.

24. Garnell, P. "Further investigations of the mechanics of roller bearings." Proc. I. Mech. E. 181 (1) 1966-67 16.

25. Higginson, G.R. and Leaver, R.H. "Fluid lubrication of tapered roller bearings." Proc. I. Mech. E. 184 (3L) 1969-70 18.

26. Leaver, R.H., Sayles, R.S. and Thomas, T.R. "Mixed lubrication and surface topography of rolling contacts." Proc. I. Mech. E. 188 44/74 1974 461.

27. Poon, S.Y. and Haines, D.J. "Frictional behaviour of lubricated rolling-contact elements." Proc. I. Mech. E. 181 (1) 1966-67 363.

28. Weibull, W. "A statistical distribution function of wide applicability." J. App. Mech. Trans. ASME 73 1951 293.

29. Crook, A.W. "The lubrication of rollers." Phil. Trans. Roy. Soc. A250 1957-58 387.

30. Archard, J.F. and Kirk, M.T. "Lubrication at point contacts." Proc. Roy. Soc. A261 1961 532.

31. Crook, A.W. "Elastohydrodynamic lubrication of rollers." Nature 190 1961 1182.

32. Crook, A.W. "Developments in elastohydrodynamic lubrication." J. Inst. Pet. 49 1963 295.

33. Orcutt, F.K. "Experimental study of elastohydrodynamic lubrication." Trans. ASLE 8 1965 381.

34. Dyson, A., Naylor, H. and Wilson, A.R. "The measurement of oil-film thickness in elastohydrodynamic contacts." Proc. I. Mech. E. 180 (3B) 1965-66 119.

35. Dowson, D. and Higginson, G.R. "Elastohydrodynamic lubrication." Pergamon Press, New York 1966.

36. Hamilton, G.M. and Moore, S.L. "A modified gauge for investigating elastohydrodynamic contact." Proc. I. Mech. E. 182 (3A) 1967-68 251.

37. Vichard, J.P. and Godet, M.R. "Simultaneous measurement of load, friction and film thickness in a cam-and-tappet system." Proc. I. Mech. E. 182 (3G) 1967-68 109.

38. Astridge, D.G. and Longfield, M.D. "Capacitance measurements and oil film thickness in a large radius disc and ring machine." Proc. I. Mech. E. 182 (3N) 1967-68 89.

39. Snidle, R. "Theoretical and experimental investigations of lubrication at point contacts." Ph.D. Thesis, University of Leicester 1970.

40. Wymer, D.G. "Elastohydrodynamic lubrication of a rolling line contact." Ph.D. Thesis, University of London 1972.

41. Napel, W.E. and Bosma, R. "The influence of surface roughness on the capacitance measurement of film thickness in elastohydrodynamic contacts." Proc. I. Mech. E. 185 1970-71 635.

42. Tallian, T.E., Chin, Y.P., Huttenlocher, D.F., Kamenshine, J.A., Sibley, L.B. and Sindlinger, N.E. "Lubricant films in rolling contact of rough surfaces." ASLE Trans. 7 1964 109.

43. Bendat, J.S. "Principles and application of random noise theory." Wiley, New York 1958.

44. Sidik, S.M. "Analysis of electrical contact occurrences between rolling surfaces with application to elastohydrodynamic lubrication." NASA Technical Note D-7134.

45. Tallian, T.E., McCool, J.I. and Sibley, L.B. "Partial elastohydrodynamic lubrication in rolling contact." Proc. I. Mech. E. 180 (3B) 1965-66 169.

46. Tallian, T.E. and McCool, J.I. "The observation of individual asperity interactions in lubricated point contact." Trans. ASLE 11 1968 176.

47. Tallian, T.E. "Elastohydrodynamic Hertzian contacts - Part II." Mech. Engineering Dec. 1971.

48. Christensen, H. "Nature of metallic contact in mixed lubrication." Proc. I. Mech. E. 180 (3B) 1965-66 147.

49. Czichos, H. "Untersuchungen uber die Verteilung metallischer und nichtmetallischer Kontaktanteide bei Mischreibung." Wear 17 1971 209.

50. Czichos, H., Grimmer, W. and Mittman, H.-U. "Rapid measuring techniques for electrical contact resistance applied to lubricant additive studies." Wear 40 1976 265.

51. Dowd, J.R. and Barwell, F.T. "Tribological interaction between piston and cylinder of a model high pressure hydraulic pump." Trans. ASLE 18 1975 21.

52. Johnson, K.L., Greenwood, J.A. and Poon, S.Y. "A simple theory of asperity contact in EHL." Wear 19 1972 91.

53. Dowson, D. and Whomes, T.L. "Effect of surface quality upon the traction characteristics of lubricated cylindrical contacts." Proc. I. Mech. E. 182 1967-68 292.

54. Nathoo, M.H. "Preparation of Langmuir films for electrical studies." Thin Solid Films 16 1973 215.

55. Nathoo, M.H. and Jonscher, A.K. "High field and a.c. properties of stearic acid films." J. Phys. C., Solid State Phys. 4 1971 L301.

56. Jonscher, A.K. and Nathoo, M.H. "High field transient behaviour of stearic acid films." Thin Solid Films 12 1972 S15.

57. Rennell, R.W. and Anderson, J.C. "Dielectic and conductivity measurements on thin liquid films of n-heptanol." Radio and Electronic Engineer 45 1975 401.

58. Holm, R. and Holm, E. "Electrical contacts handbook." 3rd edition. Springer, Berlin 1958.

59. Fiennes, W.G. and Anderson, J.C. "Electrical conduction in thin lubricating oil films." Radio and Electronic Engineer 44 1974 141.

60. Smit, J. and Wijn, H.P.J. "Ferrites." Chapter 12, Philips' Technical Library 1959.

61. Simmons, J.G. "DC conduction in thin films." Monograph EE/5 Mills and Boon, London 1971.

62. O'Dwyer, J.J. "The theory of electrical conduction and breakdown in solid dielectrics." Clarendon Press, Oxford 1973.

63. Duke, C.B. "Tunnelling in solids." Academic Press 1969.

64. Bowden, F.P. and Tabor, D. "Friction and lubrication of solids." Vol. 1. Clarendon Press, Oxford 1954.

65. Hertz, H. "On the contact of elastic bodies." J. Reine. Agnew. Math. 1881 92 156.

66. Lincoln, B.F. "Friction and elastic properties of high polymeric materials." Brit. J. Appl. Phys. 3 1952 260.

67. Zhuravlev, V.A. "On the physical basis of the Amontons-Coulomb law of friction." J. Tech. Phys. (USSR) 10 1940 1447.

68. Archard, J.F. "Elastic deformation and the laws of friction." Proc. Roy. Soc. A 243 1957 190.

69. Christensen, H. "Stochastic models for hydrodynamic lubrication of rough surfaces." Proc. I. Mech. E. 184 (1) 1969-70 1013.

70. Greenwood, J.A. and Williamson, J.B.P. "The contact of nominally flat surfaces." Proc. Roy. Soc. A 295 1966 300.

71. Greenwood, J.A. "On the area of contact between a rough surface and a flat." J. Lub. Tech. Trans. ASME 1 1967 81.

72. Greenwood, J.A. and Tripp, J.H. "The elastic contact of rough spheres." J. Lub. Tech. Trans. ASME 89 1967 81.

73. Greenwood, J.A. and Tripp, J.H. "The contact of two nominally flat rough surfaces." Proc. I. Mech. E. 185 1970-71.

74. Thompson, R.A. and Bocchi, W. "A model for asperity load sharing in lubricated contacts." ASLE-ASME Lub. Conf. Pittsburgh, Pa. Oct. 5-7 1971.

75. Whitehouse, D. and Archard, J.F. "The properties of random surfaces, of significance in their contact." Proc. Roy. Soc. A 316 1970 97.

76. Christensen, H. "Stochastic models for hydrodynamic lubrication of rough surfaces." Proc. I. Mech. E. 184 (1) 1969-70 1013.

77. Peklenik, J. "New developments in surface characterisation and measurements by means of random process analysis." Proc. I. Mech. E. 182 (3K) 1967-68 108.

78. Hirst, W. and Hollander, A.E. "Surface finish and damage in sliding." Proc. Roy. Soc. A 337 1974 379.

79. Onions, R.A. and Archard, J.F. "The contact of surfaces having a random structure." J. Phys. D. App Phys. 36 1973 667.

80. Nayak, P.R. "Random process model of rough surfaces." Trans. ASME J. Lub. Tech. 93 1971 398.

81. Longuet-Higgins, M.S. "Statistical properties of an isotropic random surface." Phil. Trans. Roy. Soc. A 250 1957 157.

82. Longuet-Higgins, M.S. "The statistical analysis of a random moving surface." Phil. Trans. Roy. Soc. A 249 1957 321.

83. Longuet-Higgins, M.S. "The statistical geometry of random surfaces." Hydrodynamic stability. Proc. 13th Symposium App. Math. Am. Math. Soc. 1962.

84. Tallian, T.E. "The theory of partial elastohydrodynamic contacts." Wear 21 1972 49.

85. Fowles, P.E. "A thermal elastohydrodynamic theory for individual asperity-asperity collisions." Trans. ASME J. Lub. Tech. 93 (F) 1971 383.

86. Williamson, J.B.P. "Microtopography of surfaces." Proc. I. Mech. E. 1967-68 Conf. Properties and metrology of surfaces, Oxford 182 (3K) 21.

87. Hisakado, T. and Tsukizoe, T. "Effects of distribution of surface slopes and flow pressures of contacting asperities on contact between solid surfaces." Wear 30 1974 213.

88. Hisakado, T. "Effect of surface roughness on contact between solid surfaces." Wear 28 1974 217.

89. Hisakado, T. "Surface roughness and deformation of contacting asperities between a rough and a flat surface." Wear 35 1975 53.

90. Cameron, A., Lewis, J.C., and Macpherson, P.B. "Hydrostatic disc machine." I. Mech. E. Conf. Nov. 1971.

91. Macpherson, P.B. Imperial College of Science and Technology. Private communication.

92. Leverson, R.C. "Sputtered thin film solid lubricants." Ph.D. Thesis, University of London 1973.

93. Maissel, L.I. and Glang, R. "Handbook of thin film technology." McGraw-Hill.

94. Anderson, J.C. and Chapman, B.N. "Science and technology of surface coatings." Academic Press 1974.

95. Hartley, N.E.W. "Ion implantation and surface modification in tribology." Wear 34 1975 427.

96. Dearnaley, G. "New uses of ion acceleration." Ed. Ziegler, J.F. Plenum Publishing Co., N.Y. 1975 Chapter 5.

97. Anderson, G.S., Mayer, W.N. and Wehner, G.K. "Sputtering of dielectrics by high frequency fields." J. App. Phys. 33 1962 2991.

98. Kirk, M.T. "Hydrodynamic lubrication of perspex." Nature 194 1962 965.

99. Archard, J.F. and Kirk, M.T. "Influence of elastic modulus on the lubrication of point contacts." Proc. Lubrication and Wear Convention (Bournemouth) 1973 Paper 15 181. (Instn. Mech. Engrs., London).

100. Cameron, A. and Gohar, R. "Theoretical and experimental studies of the oil film in lubricated point contact." Proc. Roy. Soc. A291 1966 520.

101. Gohar, R. and Cameron, A. "Optical measurement of oil film thickness under elastohydrodynamic lubrication." Nature 200 1963 458.

102. Gohar, R. "Oil films under elastohydrodynamic conditions." Ph.D. Thesis, University of London 1965.

103. Gohar, R. and Cameron, A. "The mapping of elastohydrodynamic contacts." Trans. ASLE 10 1967 215.

104. Gohar, R. "A ball-and-plate machine for measuring elastohydrodynamic oil films." Proc. I. Mech. E. 182 (3G) 1967-68 43.

105. Foord, C.A., Hamann, W.C. and Cameron, A. "Evaluation of lubricants using optical elastohydrodynamic lubrication." Trans. ASLE 11 (1) 1968 31.

106. Foord, C.A. "Pitting and film thickness in rolling contact." Ph.D. Thesis, University of London 1967.

107. Foord, C.A., Wedeven, L.D., Westlake, F.J. and Cameron, A. "Optical elastohydrodynamics." Proc. I. Mech. E. 184 (1) 1969-70 487.

108. Dowson, D. and Jones, D.A. "An optical-interference method of measurement of time-dependent elastohydrodynamic film profiles." Proc. I. Mech. E. 182 (3G) 1967-68 49.

109. Westlake, F.J. and Cameron, A. "Fluid film interferometry in lubrication studies." Nature 214 1967 633.

110. Westlake, F.J. and Cameron, A. "A study of ultra-thin lubricant films using an optical technique." Proc. I. Mech. E. 182 (3G) 1967-68 75.

111. Westlake, F.J. and Cameron, A. "High speed photographic study of lubricated contacts using optical interferometry." J. Phot. Sci. 17 (4) 1969 137.

112. Westlake, F.J. "A study of ultra-thin films." Ph.D. Thesis, University of London 1970.

113. Allchurch, G. Ph.D. Thesis, University of London. To be published.

114. Pemberton, J.C. "An optical investigation into the lubrication of cylindrical rollers." Ph.D. Thesis, University of London 1976.

115. Hedley, C. Ph.D. Thesis, University of London. To be published.

116. Hertz, H. "On the contact of elastic solids." Reprinted in "Miscellaneous Papers", translated by Jones, D.E. and Schott, G.A. Macmillan, London 1896.

117. Archard, J.F. "Contact and rubbing of flat surfaces." J. App. Phys. 24 1953 981.

118. Engel, P.A. "Impact wear of materials." Elsevier Scientific Publishing Co. 1976.

119. Fuller, K.N.G. and Tabor, D. "The effect of surface roughness on the adhesion of elastic solids." Proc. Roy. Soc. A345 1975 327.