METHODS OF CONTACT ANALYSIS OF NON-CONFORMING ROUGH SURFACES

Martin Alan West

MA

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

of the University of London

and also for the

DIPLOMA OF IMPERIAL COLLEGE

MARCH, 1991

Tribology Section

Mechanical Engineering Department

Imperial College

Exhibition Road

London SW7 2BX

BIBL.
LONDIN.
UNIV.

ABSTRACT

Limiting the vibration and noise generated in rolling and rolling-sliding contacts such as those in gears and bearings is important not only on environmental grounds but also to increase fatigue lives. The dominant source of contact vibration is surface roughness, so to attack the noise and fatigue problem, a thorough understanding of topography and surface contact mechanics is required.

An experimental programme has been undertaken to investigate the links between surface finish and vibration using a mini-disc machine. This entailed the development of a laser device to monitor the vibrations and wear, due to surface interactions, of rough disc specimens rotating in the machine. The results appear to indicate that there is little correlation between overall vibration levels and roughness parameters. However, this could simply reflect the complex nature of the asperity interaction problem.

In an attempt to isolate features of importance a theoretical model of a profile structure function (a function from which roughness parameters are obtained) has been developed. This allows roughness characteristics to be studied on an asperity size basis and is shown to give promising results on the limited tests performed.

From the experimental and theoretical work undertaken it could be seen that to study the interactions of non-

conforming rough surfaces a full understanding of the three-dimensional geometries and deformations of the surfaces involved would be required. To this end a new technique of three-dimensional rough surface contact analysis was developed.

This technique employs a numerical model of rough surface contact capable of simulating the elastic contact of real measured rough surfaces and estimating the effects of plasticity. The real pressure distribution together with a complete picture of local deformation and real contact area on an asperity scale are generated along with the shapes of the recovered surfaces after plastic deformation.

The technique is illustrated by its application to the contact in several typical engineering surfaces and some of the specimens used in the experimental programme. The results indicate that an analytical tool of general application to a wide range of Tribological situations has been created.

CONTENTS

	Page
ABSTRACT	2
CONTENTS	4
ACKNOWLEDGEMENTS	14
NOMENCLATURE	15
CHAPTER 1	
INTRODUCTION	19
1.1 BACKGROUND	19
1.2 LAYOUT OF THE THESIS	20
CHAPTER 2	
THE INITIAL TOPOGRAPHY OF THE TEST SPECIMENS	22
2.1 INTRODUCTION	22
2.2 PREPARATION OF THE TEST SPECIMENS	22
2.3 MEASUREMENT OF THE TEST SPECIMEN TOPOGRAPHIES	24
2.3.1 Two-Dimensional Measurements	24
2.3.2 Three-Dimensional Measurements	35

Table 2.1 Initial surface topography of test specimens (profiles high pass digitally filtered at 200 μm) 34

Figure 2.1 A pair of disc machine specimens 23

Figure 2.2 Representative profiles of test specimen 1 (profiles high pass filtered at 200 μm) 26

Figure 2.3 Representative profiles of test specimen 3 (profiles high pass filtered at 200 μm) 27

Figure 2.4 Representative profiles of test specimen 15 (profiles high pass filtered at 200 μm) 28

Figure 2.5 The exponential autocorrelation function 31

Figure 2.6 Three-dimensional view of test specimen 1 (vertical scale magnified 40 times) 40

Figure 2.7 Three-dimensional view of test specimen 3 (vertical scale magnified 40 times) 41

Figure 2.8 Three-dimensional view of test specimen 15 (axial lay, vertical. scale magnified x40) 42

Figure 2.9 Contour map of test specimen 1 (contours at 0.4 μm height intervals) 43

Figure 2.10 Contour map of test specimen 3 (contours at 0.8 μm height intervals) 44

Figure 2.11 Contour map of test specimen 15 (contours at 0.4 μm height intervals) 45

CHAPTER 3

THE EXPERIMENTAL EQUIPMENT 46

3.1 INTRODUCTION 46

3.2 SIMULATION OF GEAR TOOTH CONTACT CONDITIONS 46

3.3 THE BASIC DISC MACHINE 47

3.3.1 Accurate Disc Speed Indication 54

3.3.2 Torque Measurement 55

3.4 THE VIBRATION MONITOR 57

3.4.1 The Basic Wear and Vibration Monitor 57

3.4.2 Initial Tests on the Laser System 62

3.4.3 Modifications to the Laser System 64

3.4.4 Tests on the Modified Laser System 68

Figure 3.1 The disc specimens and drive shafts 50

Figure 3.2 View of torque monitor and gear pair in drive 52

Figure 3.3 View of disc speed indicating device 53

Figure 3.4 Calibrations of the torque monitor 56

Figure 3.5 The vibration monitor attached to the specimen housing 58

Figure 3.6 Detailed sections of the vibration monitor 59

Figure 3.7 Principle of operation of the vibration monitor 60

Figure 3.8 Cross-section of groove in calibration disc 63

Figure 3.9 Modified oil shield attached to the vibration monitor 67

Figure 3.10 Vibration output for 2 revolutions of the test disc 69

Figure 3.11 Vibration output for 3 revolutions of the test disc 71

Figure 3.12 Detail of the vibration output as the groove comes into contact with the reference disc 73

CHAPTER 4

EXPERIMENTAL RESULTS AND ANALYSIS 74

4.1 INTRODUCTION 74

4.2 TEST CONDITIONS 74

4.3 POSITION AND SPECKLE RESULTS 76

4.4 THE OPTICAL VIBRATION MONITOR RESULTS 79

4.4.1 Vibration Signals at Constant Pressure 79

4.4.2 Vibration Signals at Constant Slide-Roll Ratio 87

4.5 CONCLUSIONS DRAWN FROM EXPERIMENTAL PROGRAMME 87

Table 4.1 Analysis of vibration signals for a contact pressure of 1.02 GN/m^2, (signals high pass digitally filtered at 200 μm) 80

Figure 4.1 The change in the surface roughness of the test specimens during the experiments 77

Figure 4.2 Structure functions of a typical test specimen before and after experiments 78

Figure 4.3 Vibration level against specimen roughness 82

Figure 4.4 Raw laser output against specimen roughness 83

Figure 4,5 Attempt at finding correlation through dimensional analysis 85

Figure 4.6 Noise structure functions for a typical specimen at differing slide-roll ratios 86

CHAPTER 5

SURFACE TOPOGRAPHY CHANGES DURING LUBRICATED WEAR 89
5.1 INTRODUCTION 89
5.2 BACKGROUND 89
5.3 THE STRUCTURE FUNCTION 92
5.4 STRUCTURE FUNCTION CHANGES DURING WEAR 93
5.5 THE STRUCTURE FUNCTION MODEL 95
5.5.1 Basis of the Structure Function Model 95
5.5.2 Structure Function of General Asperity 95
5.5.3 The Complete Structure Function Model 96
5.6 TESTS ON THE MODELLED STRUCTURE FUNCTION 99

Figure 5.1 Structure functions of worn and unworn bearing surfaces 94
Figure 5.2 The general asperity in a profile 95
Figure 5.3 Two examples of the formation of the correction factor 98
Figure 5.4 Comparison of measured and modelled structure functions of a bearing profile 100

CHAPTER 6

6 **THREE-DIMENSIONAL ROUGH SURFACE CONTACT** 102
6.1 INTRODUCTION 102

6.2 THEORETICAL ROUGH SURFACE CONTACT ANALYSIS 103
6.2.1 Elastic Contact 103
6.2.2 Plastic Contact 110
6.2.3 Concluding Remarks on Theoretical Methods 115
6.3 NUMERICAL ROUGH SURFACE CONTACT ANALYSIS 115
6.3.1 Variational Methods 118
6.3.2 Matrix Inversion Method 119

CHAPTER 7

THE THREE-DIMENSIONAL CONTACT MODEL 122
7.1 INTRODUCTION 122
7.2 BASIS OF THE MODEL 123
7.2.1 The Geometry of Contact 123
7.2.2 Constraints of the Model 125
7.3 THE CONTACT MODEL 126
7.3.1 Contact Equations 126
7.3.2 The Plasticity Model 130
7.4 IMPLEMENTATION OF THE MODEL 134
7.4.1 The Contact Program 136
7.5 VALIDATION OF THE MODEL 141
7.5.1 Smooth Plane Tests 141
7.5.2 Geometric Surface Tests 144
7.6 EXAMPLES USING REAL DATA 148
7.6.1 Worn Bearing 148
7.6.2 Abraded Steel Surface 153
7.7 THE EFFECT OF PLASTICITY 158
7.8 CONCLUSIONS 161

Figure 7.1 Two rough surface profiles and their equivalent single profile 124
Figure 7.2 Detail of pressure element used 127
Figure 7.3 Flow diagram of contact program 3DCONT 135
Figure 7.4 Smooth body in contact with a plane:- surface pressure distribution 143
Figure 7.5 Real area of contact of a two-dimensional wavy surface with an elastic half-space 146
Figure 7.6 Changing shape of contact area with load using the contact model 147
Figure 7.7 Worn bearing surface 149
Figure 7.8 Smooth body in contact with worn bearing:- contact patch 150
Figure 7.9 Smooth body in contact with worn bearing:- deformed shape of smooth body 151
Figure 7.10 Smooth body in contact with worn bearing:- surface pressure distribution 152
Figure 7.11 Abraded steel surface 154
Figure 7.12 Smooth body in contact with abraded steel:- contact patch 155
Figure 7.13 Smooth body in contact with abraded steel:- deformed shape of smooth body 156
Figure 7.14 Smooth body in contact with abraded steel:- surface pressure distribution 157
Figure 7.15 Dented surface 159
Figure 7.16 Detail of worn bearing surface 160

CHAPTER 8

3-D CONTACT ANALYSIS OF THE TEST SPECIMENS 162

8.1 INTRODUCTION 162

8.2 THE CONTACT SIMULATIONS 162

8.3 THE CONTACT SIMULATION RESULTS 163

Table 8.1 Summary of results of three-dimensional contact simulations 170

Figure 8.1 Surface pressure distribution between smooth surface and test specimen 1 164

Figure 8.2 Surface pressure distribution between smooth surface and test specimen 3 165

Figure 8.3 Surface pressure distribution between smooth surface and test specimen 15 166

Figure 8.4 Contact patch between smooth surface and test specimen 1 167

Figure 8.5 Contact patch between smooth surface and test specimen 3 168

Figure 8.6 Contact patch between smooth surface and test specimen 15 169

Figure 8.7 Ratio of maximum pressure to Hertzian pressure against average surface roughness for the test specimens 171

CHAPTER 9

CONCLUSIONS 173

9.1 GENERAL CONCLUSIONS 173

9.1.1 The Experimental Programme 174

9.1.2 Theoretical Structure Function Model 174

9.1.3 Three-Dimensional Surface Contact Model 175

9.2 SUGGESTIONS FOR FURTHER WORK 176

APPENDIX A

THE CONSTRUCTION OF EQUATION 5.6c 178

Figure A.1 The possible regimes 179

APPENDIX B

A METHOD OF 3-DIMENSIONAL TOPOGRAPHY MEASUREMENT AND ANALYSIS ON ARCUATE SURFACES 183

- Reproduction of published paper

APPENDIX C

A NEW METHOD FOR ROUGH SURFACE CONTACT ANALYSIS 192

- Reproduction of published paper

APPENDIX D

A 3-DIMENSIONAL METHOD OF STUDYING 3-BODY CONTACT GEOMETRY AND STRESS ON REAL ROUGH SURFACES 204

- Reproduction of published paper

REFERENCES 211

ACKNOWLEDGEMENTS

I would like to give particular thanks to Dr Ritchie Sayles for his encouragement and guidance throughout my period of research. I am also indebted to the late Dr "Mac" Macpherson without whom I would never have embarked on this endeavour.

In addition, I would like to thank Dr Richard Wayte for his assistance with the design and construction of the optical vibration monitor. Thanks are also due to all the members of the Tribology Section, for creating such a congenial academic and social environment.

Finally, I would like to take the opportunity of thanking Westland Helicopters Limited and the Ministry of Defence for funding this research work.

. . . . .

This thesis is dedicated to Emma, my daughter, and to my dear wife Julie without whose assistance and unfailing support this project might never have been completed.

. . . . .

NOMENCLATURE

Symbol	Meaning
a	Hertzian semi-contact width in the x direction; asperity half base width
A	Contact area
b	Hertzian semi-contact width in the y direction
[C]	Matrix of influence coefficients
C_x, C_y	Coordinates of centre of curved body
E	Young's modulus
E^*	Effective elastic modulus $= \frac{1}{\left[\frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2}\right]}$
E{ }	Expectation
E_i	Expectation in region i
F	Applied load
$F_1(R_x/R_y)$	Shape function

$F_2(R_x/R_y)$	Shape function
h_{Ai}, h_{Bi}	Heights of the centre of asperity i relative to the positions of the underlying τ steps
H	Indentation hardness
k	Yield stress in simple shear
L	Profile length
m_2	Second moment of the Power Spectral Density Function
N	Number of contact spots
p	Pressure
(p)	Vector of normal pressures at contacting points
p_0	Maximum Hertzian contact pressure
p_y	Contact pressure to cause plastic yield
p^*	Mean pressure to cause complete contact
P_i	Probability of being in region i
R_a	Centre-line average roughness

R_q	Root-mean-square roughness
R_s	Summit radius
R_x, R_y	Principal radii of curvature
R_e	Equivalent radius of curvature
$R(\tau)$	Autocovariance function (ACVF)
$S(\tau)$	Structure function
u_z	Normal deflection of a surface point
(u_z)	Vector of surface deflections of contacting points
U_E	Total strain energy
V^*	Total complementary energy
x	Horizontal distance
Y	Yield stress in simple tension
$z(x)$	Profile height at position x
β	Correlation length at $\rho(\ )=0.1$
β^*	Correlation length at $\rho(\ )=1/e$

δ	Overall displacement overlap of contacting bodies
Δ	Amplitude of sinusoidal waveform
θ	Asperity base angle
λ	Wavelength
ν	Poisson's ratio
$\rho( )$	Autocorrelation function (ACF)
σ	Root mean square roughness; standard deviation
σ_s	Standard deviation of summit heights
$\sigma_1, \sigma_2, \sigma_3$	Principal stresses
τ	Lag or delay
τ_{max}	Maximum delay length
ψ	Plasticity index

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

The contact of real rough surfaces has always presented a difficult problem in Tribology due to the complex asperity interactions occurring within a contact. The problem is made more difficult by the inaccessibility of the contact region to experimental methods.

The use of simple parameters to define the character of the surface topography is far too crude to indicate the complete contact conditions. The investigations detailed in this thesis therefore concentrate on developing new theoretical and numerical techniques to attempt to isolate topographical features of importance in surface interactions. However, this work was initially driven by an investigation into gear noise problems.

Vibration and noise generated in rolling and sliding Hertzian contacts are important in almost every technological environment. Quieter gear sets with longer fatigue lives are urgently required for applications such as helicopter installations. Rolling fatigue lives are closely related to surface roughness and there is little

doubt that as vibration levels can provide an effective monitor of energy dissipation, through lubricated surface contact interactions, they must also act as a measure of the effective surface roughness. The dominant source of all Hertzian contact vibration is surface roughness, so to attack the fatigue and noise problem with any hope of success, the first priority must lie in a thorough understanding of topography and surface contact mechanics.

1.2 LAYOUT OF THE THESIS

The diverse nature of the work described within this thesis makes the use of a conventional literature review chapter undesirable. Instead relevant literature is dealt with in each chapter as appropriate.

Chapters 2 to 4 detail an experimental programme using a mini-disc machine to investigate the links between surface finish and vibration in rolling and rolling-sliding contacts, such as those found in gears and bearings. This investigation entailed the development of an optical device to monitor the vibrations and wear, due to surface interactions, of rough disc specimens rotating in the mini-disc machine.

Chapter 5 describes the development of a theoretical model of a profile structure function allowing roughness characteristics to be studied on an asperity size basis.

Chapters 6 and 7 cover three-dimensional contact analysis. Chapter 6 is a review of the contact analysis

literature and Chapter 7 a full description of the development of a numerical technique capable of simulating the elastic contact of real measured rough surfaces and estimating the effects of plasticity.

Chapter 8 details the application of the three-dimensional contact analysis technique to some of the specimen surfaces used in the experimental programme.

The remainder of the thesis gives the general conclusions of the investigations and suggestions for further work.

CHAPTER 2

THE INITIAL TOPOGRAPHY OF THE TEST SPECIMENS

2.1 INTRODUCTION

An experimental programme was undertaken to investigate the links between surface finish and vibration in gear sets. Basically this entailed developing a mini-disc machine in which the contact of two disc specimens could be monitored by an optical device (incorporating a laser) to discover the vibration and wear caused by the interaction of the roughness of the specimens as they rotated. Chapter 3 describes the experimental equipment used and the reasons for its choice.

However, the following sections describe the disc specimens used in the mini-disc machine and the measurements made of their topographies before the experiments were carried out. First using two-dimensional profile measurements and then full three-dimensional surface topography maps.

2.2 PREPARATION OF THE TEST SPECIMENS

For the experimental programme using the mini-disc

machine fifteen pairs of disc specimens were produced, see Figure 2.1. Each pair consisted of a reference specimen required by the laser system of the mini-disc machine (see Chapter 3), and a test specimen with a particular form of topography.

The disc specimens were manufactured from EN 39B gear steel and the discs with the wider running tracks (see Figure 2.1) were used as the laser reference specimens.

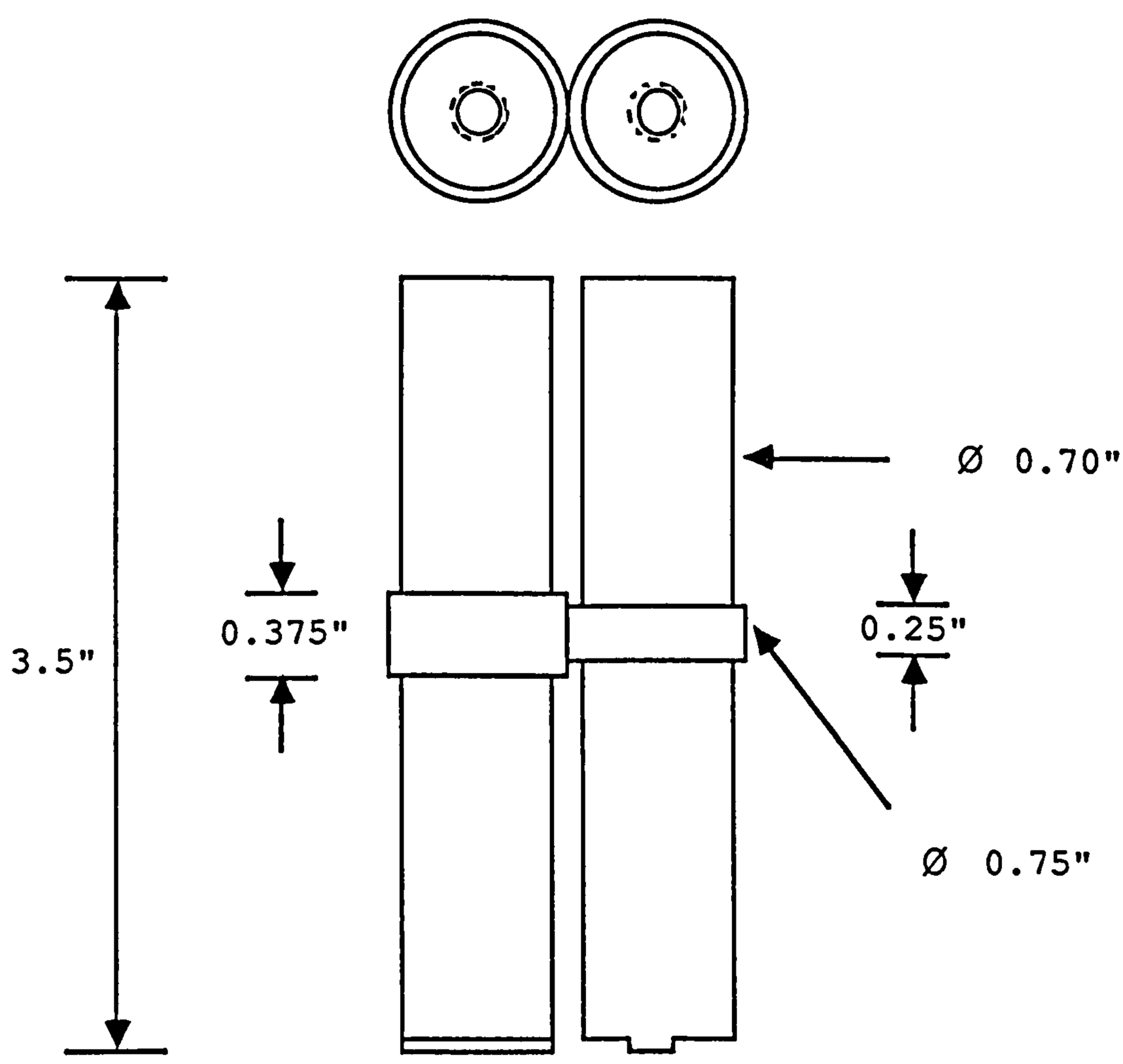


Figure 2.1 A pair of disc machine specimens

The running tracks of the laser reference specimens had all been circumferentially ground at the same time so that their surface topography and laser reflection properties would be as similar as possible. Note that each one was individually calibrated with a static movement of the disc before any tests were run (see Chapter 3 for details of this calibration).

The actual test specimens had their initial circumferentially ground running track surfaces modified to give a range of surface topographies. Ideally the lay of the roughness on these specimens should have been along the axes of the specimens as this would simulate the actual lay of roughness on gear tooth flanks. However, due to difficulties encountered whilst attempting to produce a surface of consistent roughness all the way around the circumference of the specimen with an axial lay direction, only one specimen was produced with its lay in this direction (specimen 15). The other fourteen specimens all had circumferentially ground surfaces but with a range of roughness characteristics.

2.3 MEASUREMENT OF THE TEST SPECIMEN TOPOGRAPHIES

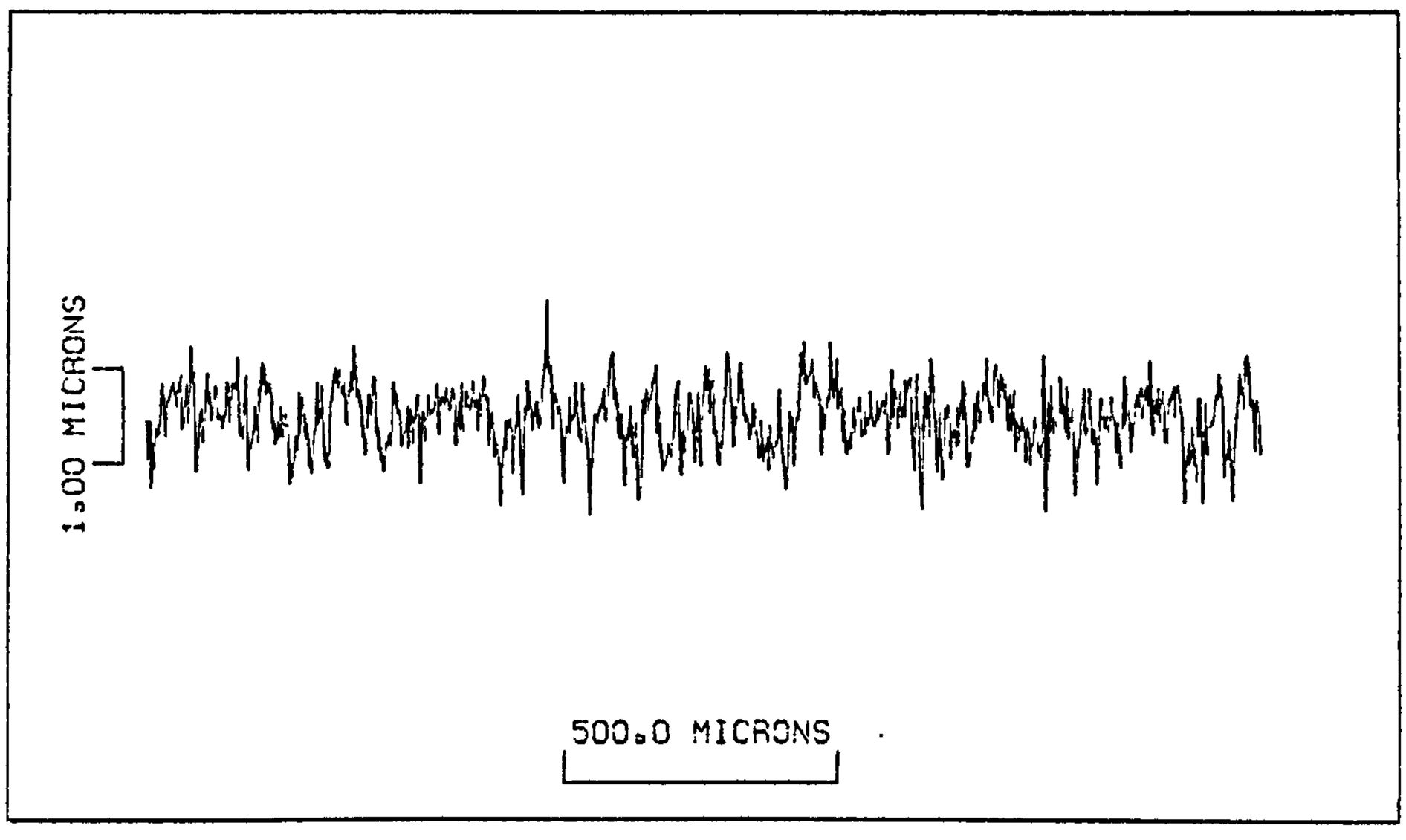
2.3.1 Two-Dimensional Measurements

After production the surface topography of each specimen was characterised in both the axial and circumferential directions. A Talysurf 4 stylus instrument was used to record profiles in the axial

direction and a Rotary Talysurf stylus instrument was used for circumferential profile measurements.

The profiles obtained were high pass digitally filtered using a moving spline filter to remove any waviness from the signals. The filter length chosen was 200 μm as this approximated to the Hertzian contact width for the operating load to be used in the experiments. Figures 2.2, 2.3 and 2.4 show representative axial and circumferential profiles for three of the test specimens.

a) Profile in axial direction



b) Profile in circumferential direction

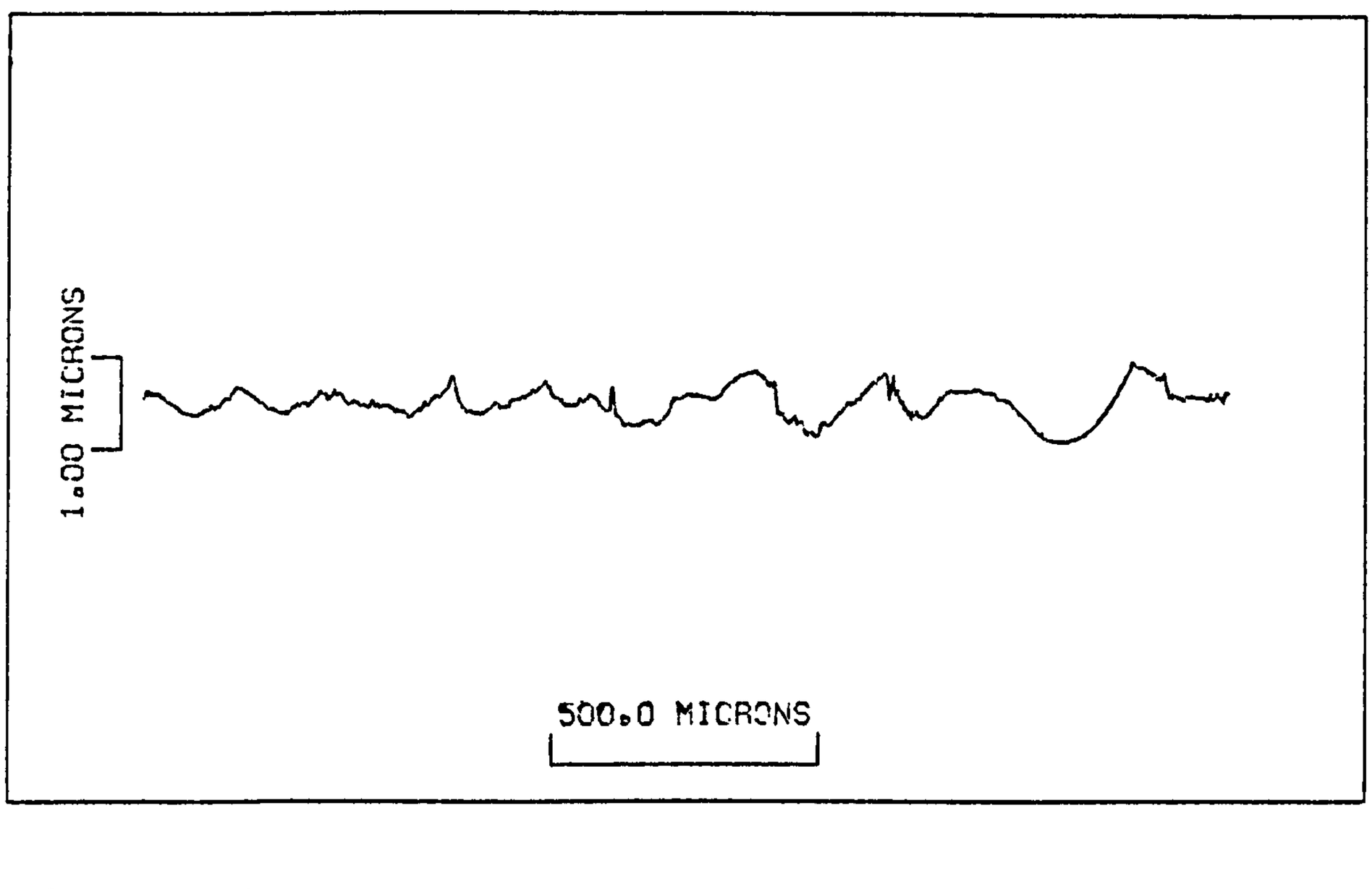
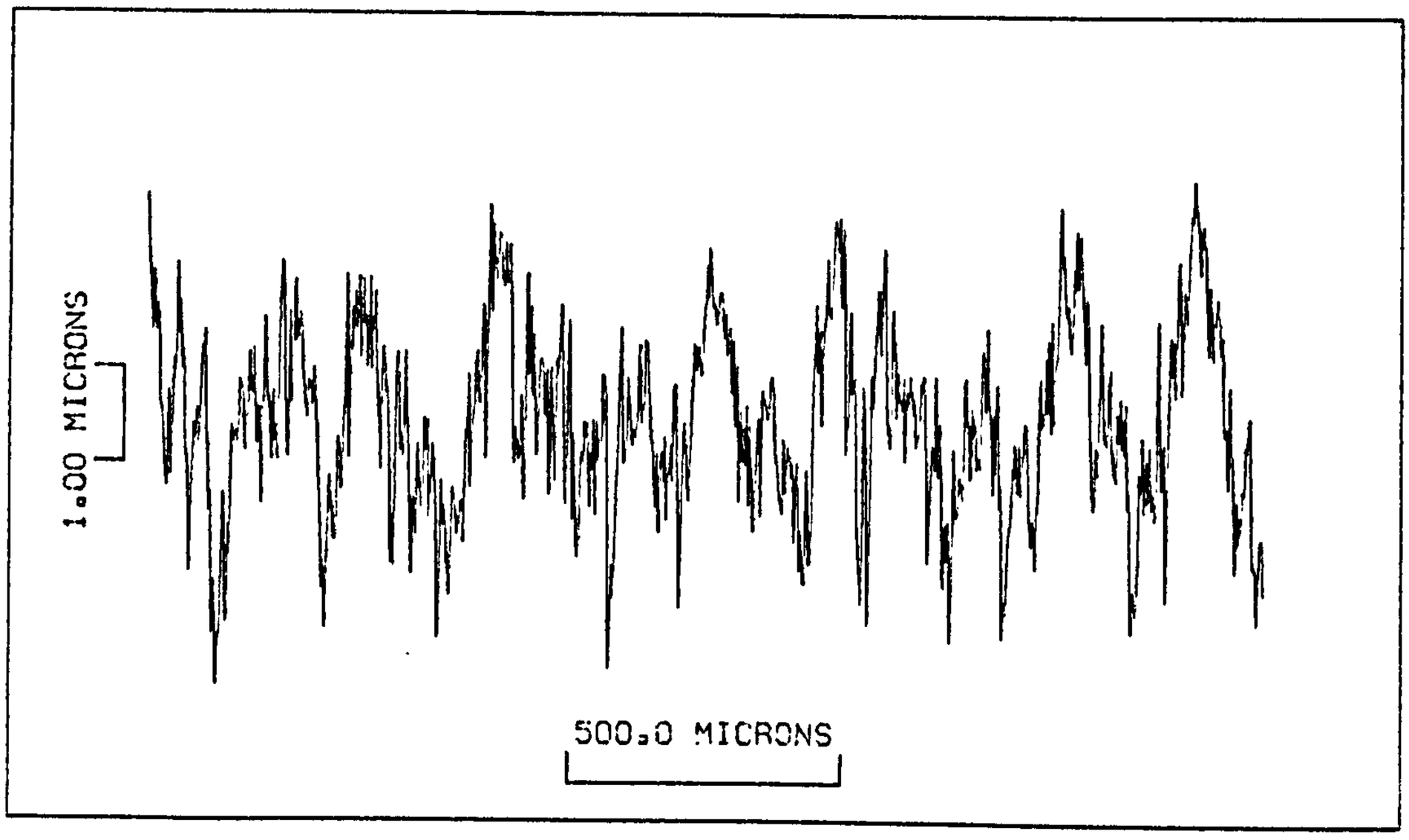


Figure 2.2 Representative profiles of test specimen 1 (Profiles high pass filtered at 200μm)

a) Profile in axial direction



b) Profile in circumferential direction

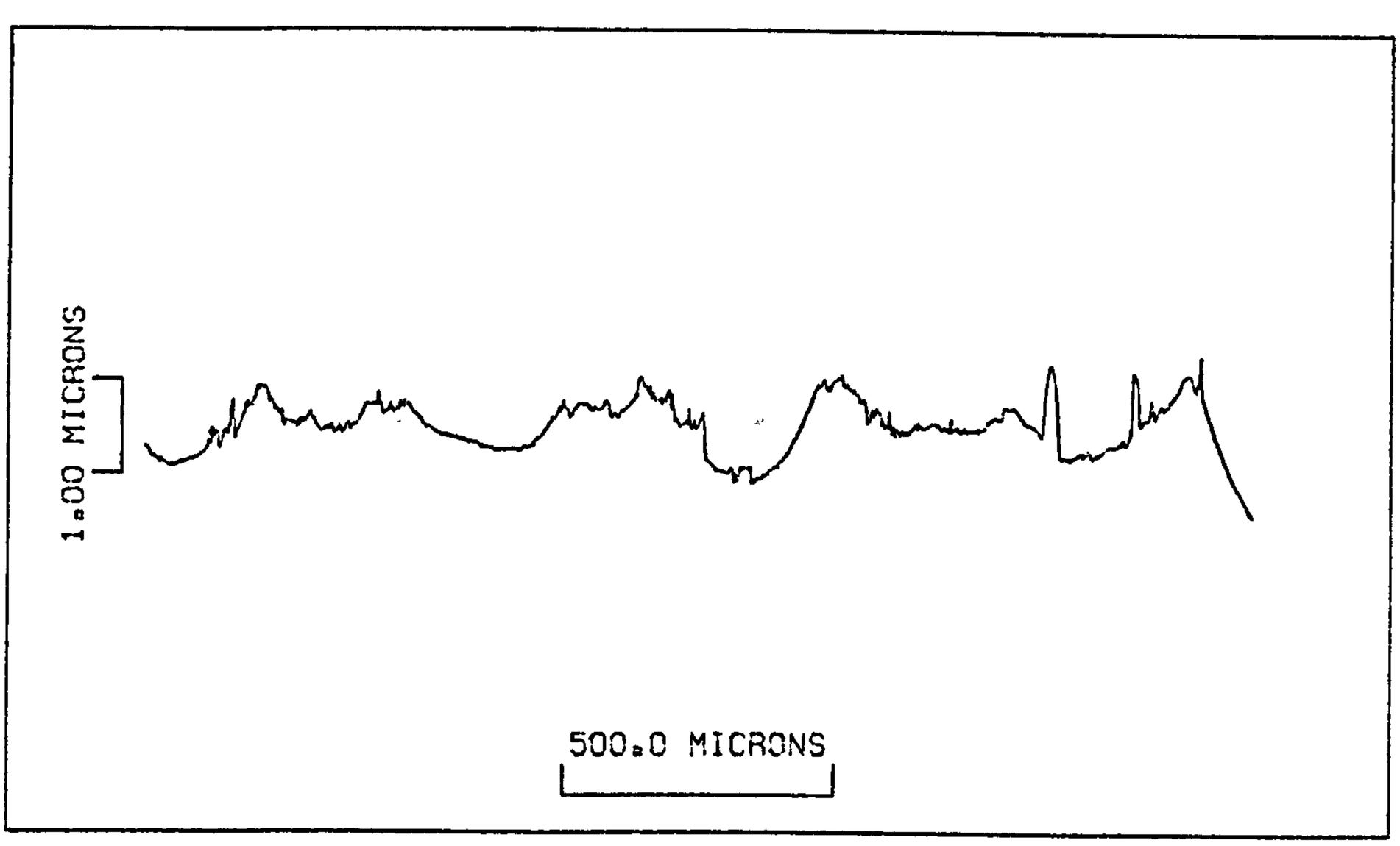
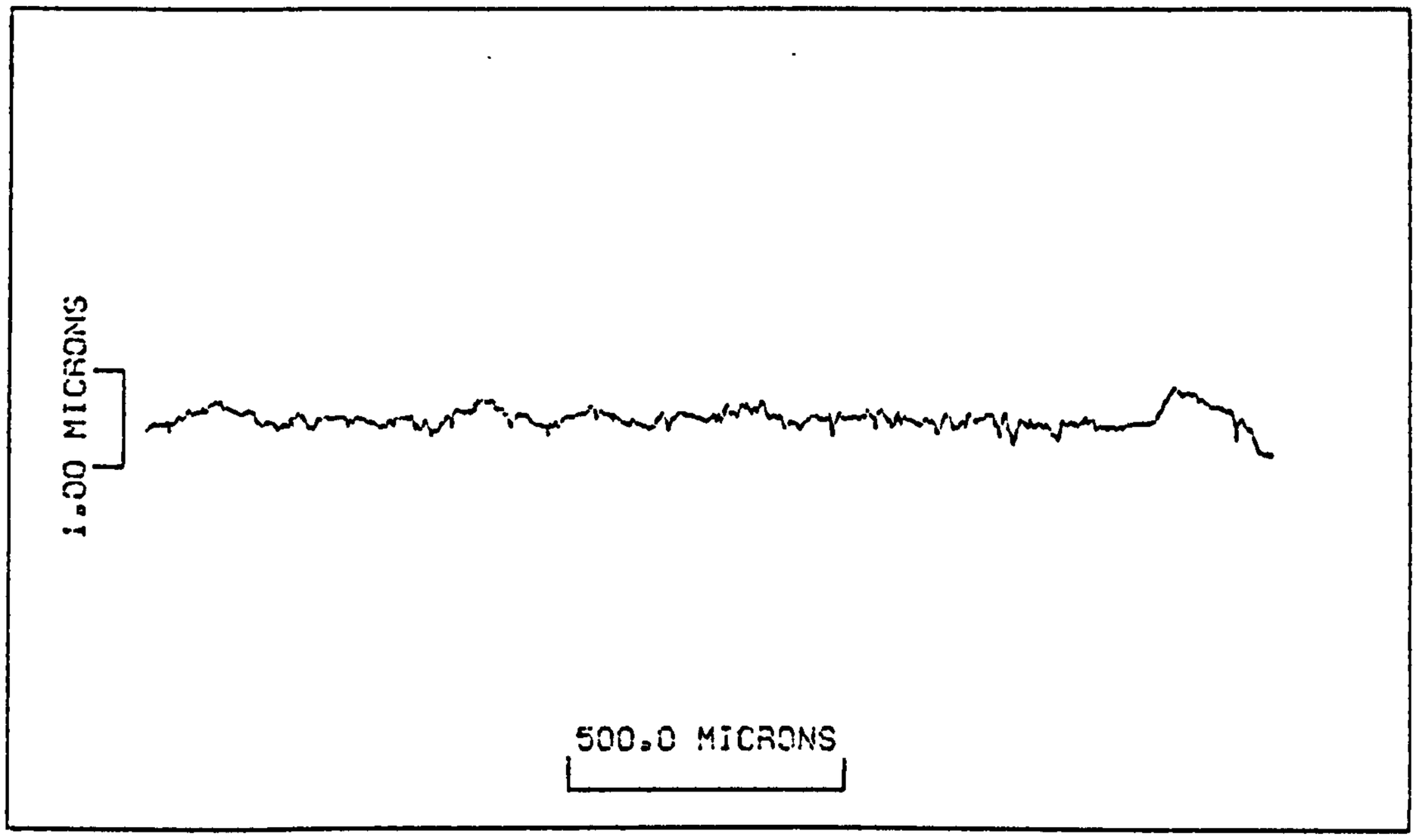


Figure 2.3 Representative profiles of test specimen 3
(Profiles high pass filtered at 200μm)

a) Profile in axial direction



b) Profile in circumferential direction

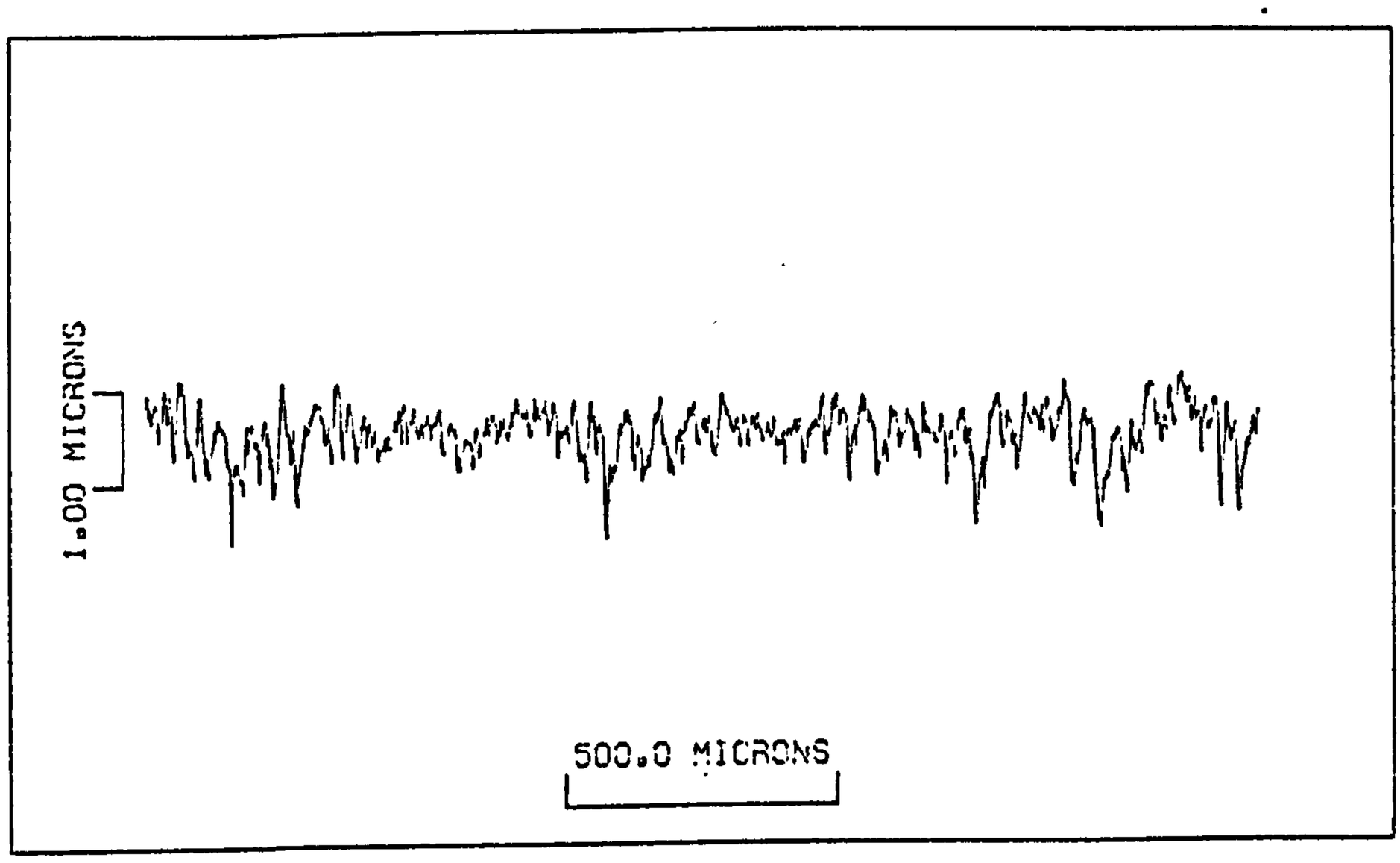


Figure 2.4 Representative profiles of test specimen 15
(Profiles high pass filtered at 200μm)

2.3.1.1 Two-dimensional profile statistics

There are a large number of parameters to choose from when describing the topography of surface profiles, some of general application and others for specific situations. However, in general they can be split into two groups:

2.3.1.1.1 *Height descriptors*

These include average roughness parameters such as the centre-line average roughness R_a and the root-mean-square (RMS) roughness σ or R_q.

The centre-line average roughness is the most widely used roughness parameter because of its ease of measurement; this is defined for a profile of length L as

$$R_a = \frac{1}{L}\int_0^L |z|\,dx \qquad (2.1)$$

where z is the height of the profile measured from the mean line.

Not so widely used for practical measurements but more common in theoretical work is the RMS roughness. This is defined by

$$\sigma = \left\{ \frac{1}{L}\int_0^L z^2\,dx \right\}^{\frac{1}{2}} \qquad (2.2)$$

As it is weighted by the square of the heights it tends to

be more sensitive than R_a to large deviations from the mean line. In some cases this can represent an advantage in the RMS over the R_a parameter. It should also be noted that if redistribution of the topography occurs (during processes such as running in) and the volume of material remains constant there will be no change in the value of R_a but significant changes in σ are possible (Leaver *et al.*, 1974). It seems that where surface deformation is involved the σ parameter is far more sensitive than the R_a roughness.

2.3.1.1.2 *Spatial descriptors*

The autocovariance function (ACVF), or its normalised form the autocorrelation function (ACF), has been the most popular way of representing spatial variation. The ACVF known as $R(\tau)$ is defined as

$$R(\tau) = \lim_{L\to\infty} \frac{1}{L} \int_0^L z(x)\, z(x+\tau)\, dx \qquad (2.3)$$

where L is the sample length, $z(x)$ the height of the profile from the mean line at position x and $z(x+\tau)$ the height at a distance τ later. The ACF, $\rho(\tau)$ is simply the ACVF normalised by the profile sample variance σ^2 (ie the square of the RMS roughness). Thus:

$$\rho(\tau) = \frac{R(\tau)}{\sigma^2} \qquad (2.4)$$

A model found to fit the autocorrelation function of many random surfaces is

$$\rho(\tau) = \exp\left(-\frac{\tau}{\beta^*}\right) \tag{2.5}$$

This simple exponential form (see Figure 2.5) has a correlation length ($\rho(\tau)=0.1$) of $\beta=2.3\beta^*$, where $1/\beta^*$ is the decay rate of the function at the origin. The values of the correlation lengths β and β^* give a measure of the spatial variation of the profile, but in general β^* is to be preferred as it is the more stable parameter of the two.

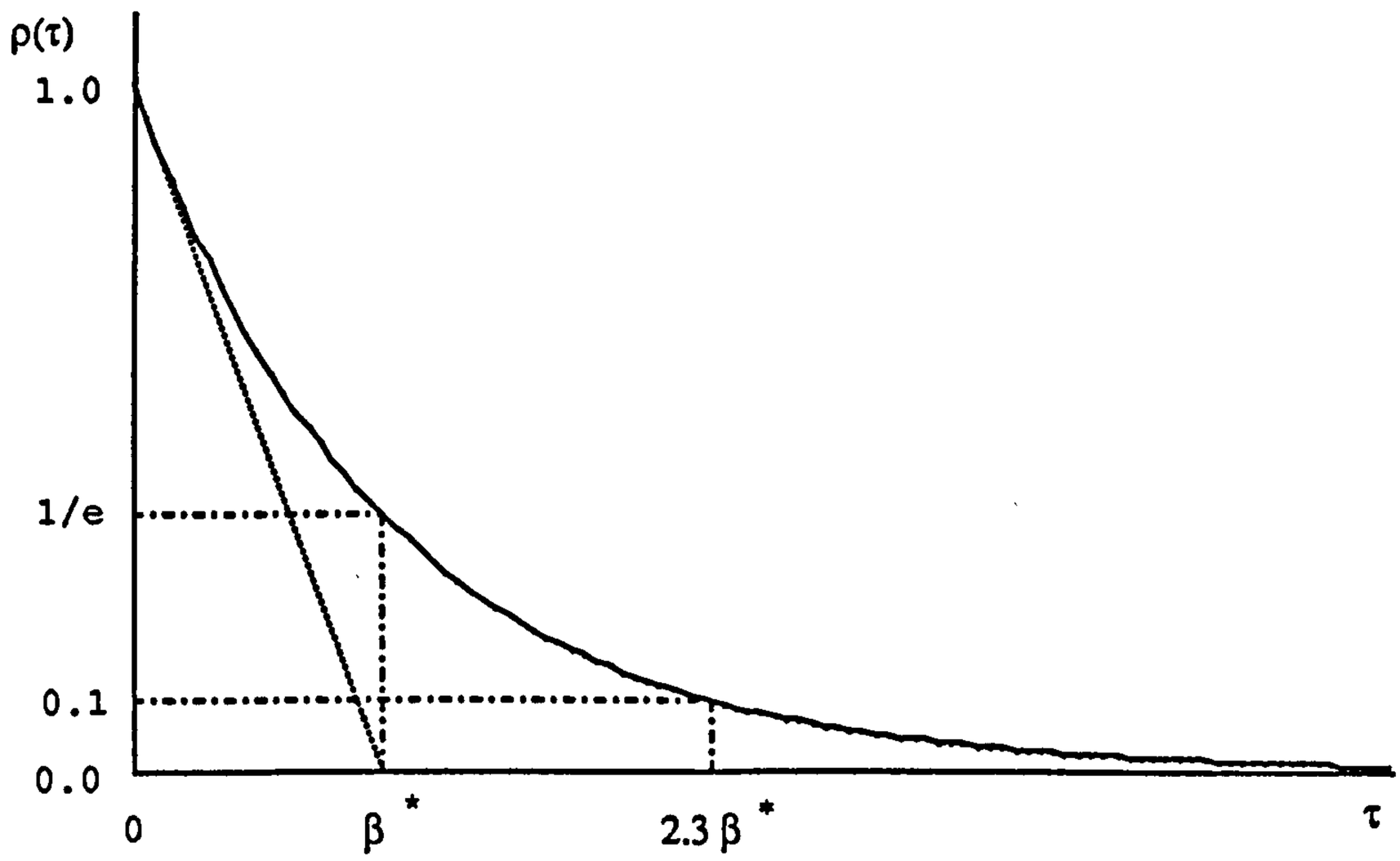


Figure 2.5 The exponential autocorrelation function

Another descriptor of spatial variation is the structure or variance function (SF), (Sayles and Thomas, 1977) defined in integral form as

$$S(\tau) = \lim_{L\to\infty} \frac{1}{L} \int_0^L [z(x) - z(x+\tau)]^2 \, dx \qquad (2.6)$$

The function represents the mean square of the difference in height expected over any spatial distance τ. For stationary profiles it contains the same information as the ACF. The two principal advantages are that its construction is not limited to the stationary case and secondly that it is independent of the mean plane. Thus any profile SF irrespective of its mean line is a section of the surface SF, a property not shared by the ACF. See Chapter 5 for more details of the structure function.

2.3.1.2 Profile statistics of the test specimens

It was decided that from the range of profile statistics available a single height descriptor and a single spatial descriptor should be used for the analyses. The actual height parameter chosen was the RMS roughness σ and the spatial parameter was the correlation length β^* (see Section 2.3.1.1).

The reason for the choice of these particular parameters was twofold. Firstly, previous experimental work by Hirst and Hollander (1974) has shown the transition between safe and unsafe sliding, under boundary lubrication, to be linked to the ratio of σ/β^*. Secondly,

other work being carried out by the author using a range of specimens in which σ and β^* were tightly controlled were giving encouraging results when analysed using a two-dimensional elastic contact model (West *et al.*, 1987 - reproduced as Appendix C). See also Chapters 6 and 7 for more information on the contact model.

Table 2.1 tabulates these profile statistics obtained from the measurements made of the fifteen test specimens and the reference specimens. Note that each reading is an average of three recorded profiles and that the longcrestedness γ^{-1} (longuet-Higgins, 1957) is approximated by the ratio of the axial roughness to the circumferential roughness (Sayles and Thomas, 1976).

SPECIMEN NO.	AXIAL		CIRCUMFERENTIAL		LONG-CRESTEDNESS
	σ (μm)	β^* (μm)	σ (μm)	β^* (μm)	
1	0.297	7	0.151	50	2.0
2	0.657	26	0.191	47	3.4
3	0.970	22	0.289	53	3.4
4	0.486	17	0.202	47	2.4
5	0.669	9	0.273	42	2.5
6	0.136	4	0.080	33	1.7
7	0.476	36	0.159	39	3.0
8	0.281	5	0.159	49	1.8
9	0.748	20	0.160	42	4.7
10	0.648	45	0.141	41	4.6
11	0.306	10	0.118	34	2.6
12	0.519	14	0.258	56	2.0
13	0.623	4	0.315	45	2.0
14	0.510	29	0.124	45	4.1
15	0.096	32	0.276	10	0.35
Reference Specimens	0.281	6	0.144	48	2.0

Table 2.1 Initial surface topography of test specimens, (profiles high pass digitally filtered at 200 μm).

2.3.2 Three-Dimensional Measurements

Previous work by Sayles (1978) has shown that a knowledge of the complete three-dimensional nature of surface asperities is required to explain all aspects of contact vibration. However, Sayles and Thomas (1976) indicate that for ground surfaces the normal profile measurements described in the previous section are sufficient to define the complete surface topography. The results of Table 2.1 were therefore used in the analysis of the experimental results.

As the project progressed facilities for full three-dimensional surface measurement became available, (Webster, 1986). To use these measurements, software was written to produce three-dimensional views and contour maps of the surface topographies. The complete system is described in (Webster *et al.*, 1986 - reproduced as Appendix B) but a brief outline of the viewing software is given here.

2.3.2.1 **Three-dimensional viewing software**

The three-dimensional surface measurement system records a rectangular matrix of surface heights using a Talysurf 4 Stylus instrument for later analysis and viewing. A suite of Fortran routines was written to make this viewing possible.

2.3.2.1.1 *Three-dimensional Filtering*

Once the height data has been digitised and stored in the computer a least-squares mean plane is fitted to remove

any general trends. The option is then available to filter the data on an area basis, if required. If the data is plotted without filtering, the picture can often be confused and features of topographical interest obscured.

The filter routine employed is similar to the moving-average filter commonly used in two-dimensional profilometry, i.e. to obtain a low pass filter, each data point height is replaced by the average height of all the points falling within a filter length centred on that point. To obtain a high pass filter, the low pass result is subtracted from the original height at that point.

When this is extended to three dimensions, the filter length becomes a filter box, centred on the point being considered, with sides equal to the filter length. Any point which lies within this box is included in the average used to find the low pass filtered output at the central point. Again, the high pass filter is obtained by subtracting the low pass output from the original data.

Band pass filtering can also be obtained by successively high and low pass filtering the data with appropriate filter lengths.

2.3.2.1.2 *Three-dimensional viewing with perspective*

A FORTRAN 77 computer program called SFMAP was written to produce three-dimensional views, including the effects of perspective, of raw or filtered data. The view of the surface, on a vdu screen or drum plotter, is defined by two angles and a distance. The two angles are the rotation of the surface about the horizontal and vertical axes through

its central point and the distance is the range from the viewing position to the centre of the surface. This distance is used to enhance or reduce the effects of perspective.

It should be noted that in general because engineering surfaces are relatively flat a true three-dimensional picture would not show any detail at all. To overcome this the vertical scale can be adjusted relative to the horizontal scales. For bearing surfaces a vertical scale of approximately 75 times greater than the horizontal usually gives good results. The view obtained is therefore a true three-dimensional picture of a surface 75 times rougher than the actual surface.

The picture is formed by joining the data points as profiles in either or both of the grid directions, with any hidden detail removed. To save time and aid clarity not all the profiles have to be drawn.

Once a picture has been produced on the vdu screen several options are available:-

1) Another viewing position can be selected and this new view drawn on the vdu or plotter.

2) Areas of the surface can be isolated interactively so that important topographical features can be enlarged for more detailed analysis.

3) Individual profiles (in either of the grid directions)

can be isolated to be stored as separate data files for more detailed study.

2.3.2.1.3 *Contour mapping*

A separate FORTRAN 77 computer program called CTMAP was written to draw contour maps of raw or filtered surface data. This program can draw up to 20 constant height contours at height intervals supplied by the user.

The contours are generated by searching the height data for adjacent points straddling a particular contour height and then using linear interpolation to find the actual position that the contour passes between the data points. Once a contour has been defined by its positions between data points, the actual contour path is drawn by fitting a smooth cubic spline through the contour points.

The resultant contour maps can be output to a vdu screen or drum plotter.

2.3.2.2 **Test specimen surface maps**

Three-dimensional measurements were made of three of the test specimens: specimens 1,3 and 15, as these covered the range of the surface topographies to be used (specimen 15 being the specimen with its lay in the axial direction).

Figures 2.6 to 2.11 show 3-dimensional views and contour maps of these three specimens. These measurements were not to be used in a quantitative way but to give a feel for the differences between the test specimens before they were used in the experiments.

However, later in the project quantative contact analysis was attempted using these measurements (see Chapter 8).

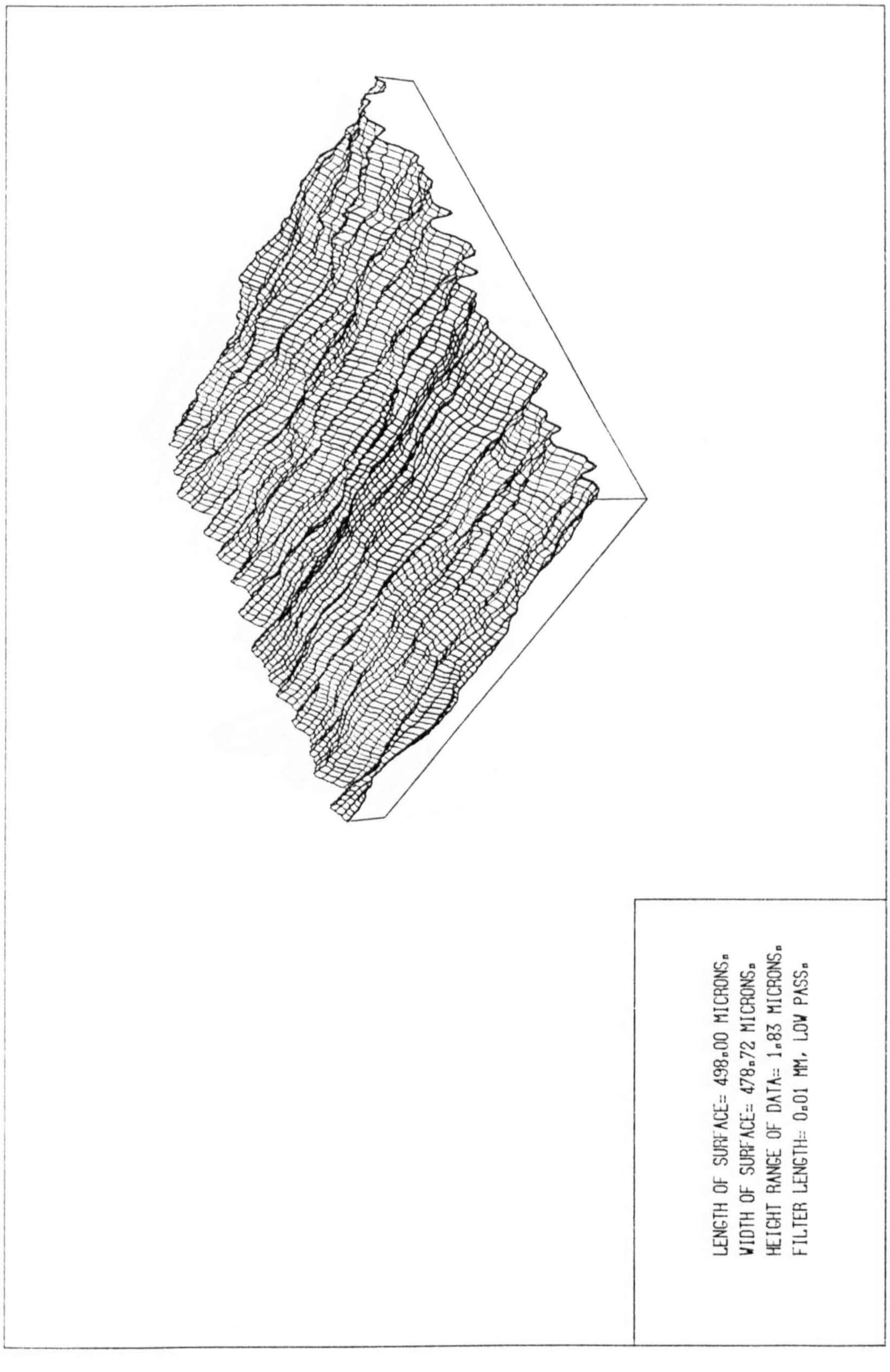


Figure 2.6 Three-dimensional view of test specimen 1
(Vertical scale magnified 40 times)

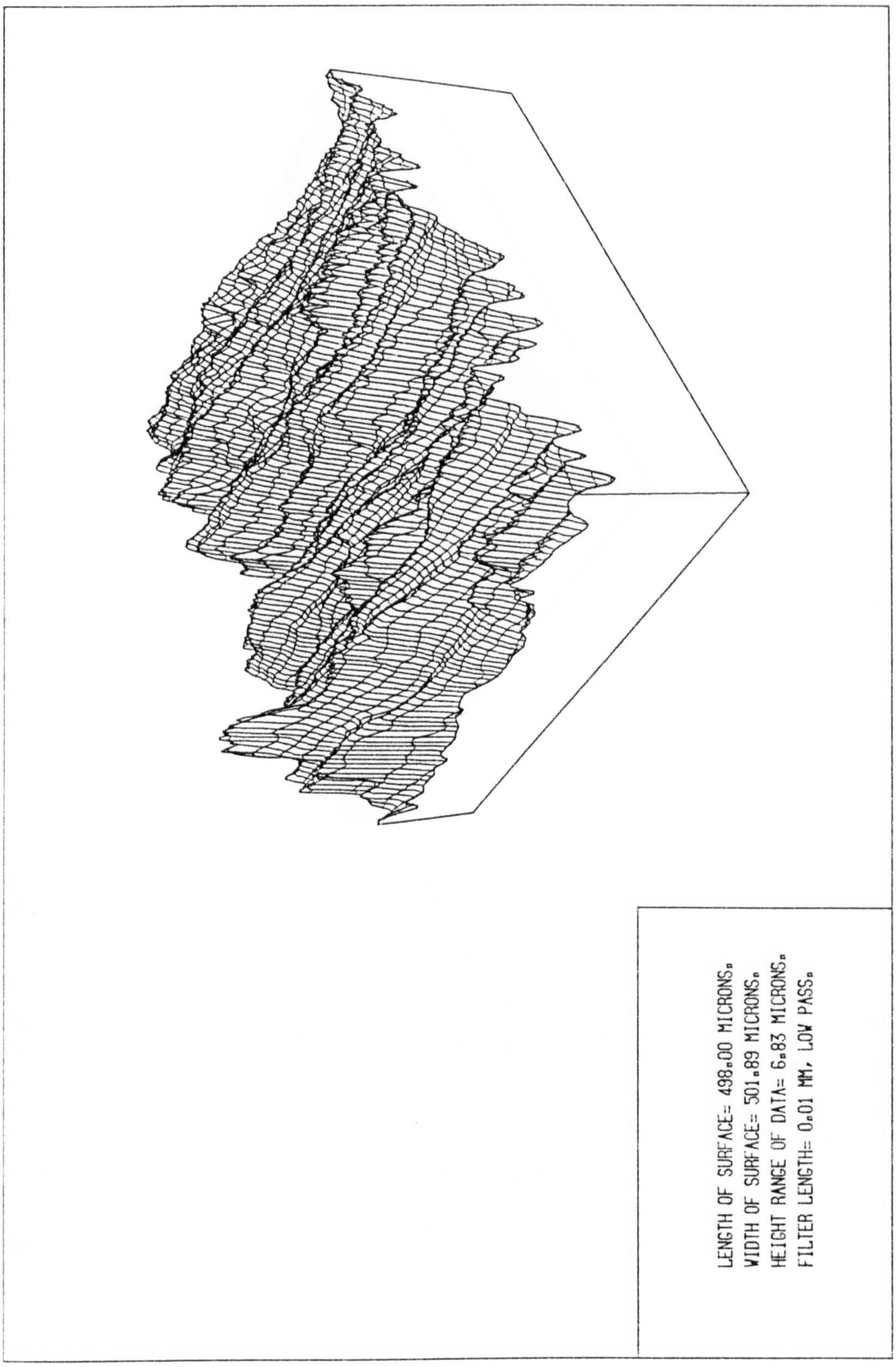


Figure 2.7 Three-dimensional view of test specimen 3
(Vertical scale magnified 40 times)

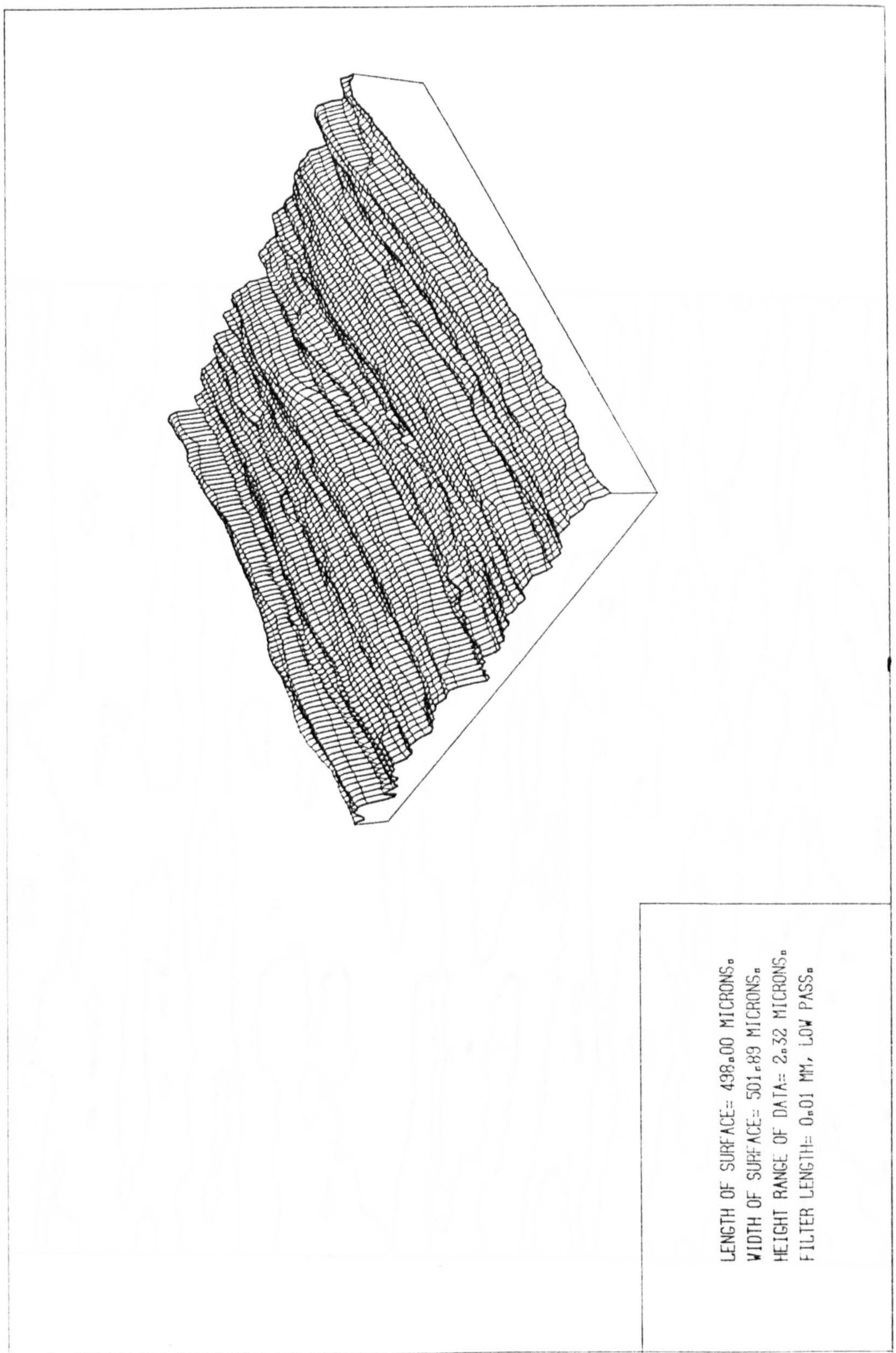


Figure 2.8 Three-dimensional view of test specimen 15 (Axial lay, vertical scale magnified x40)

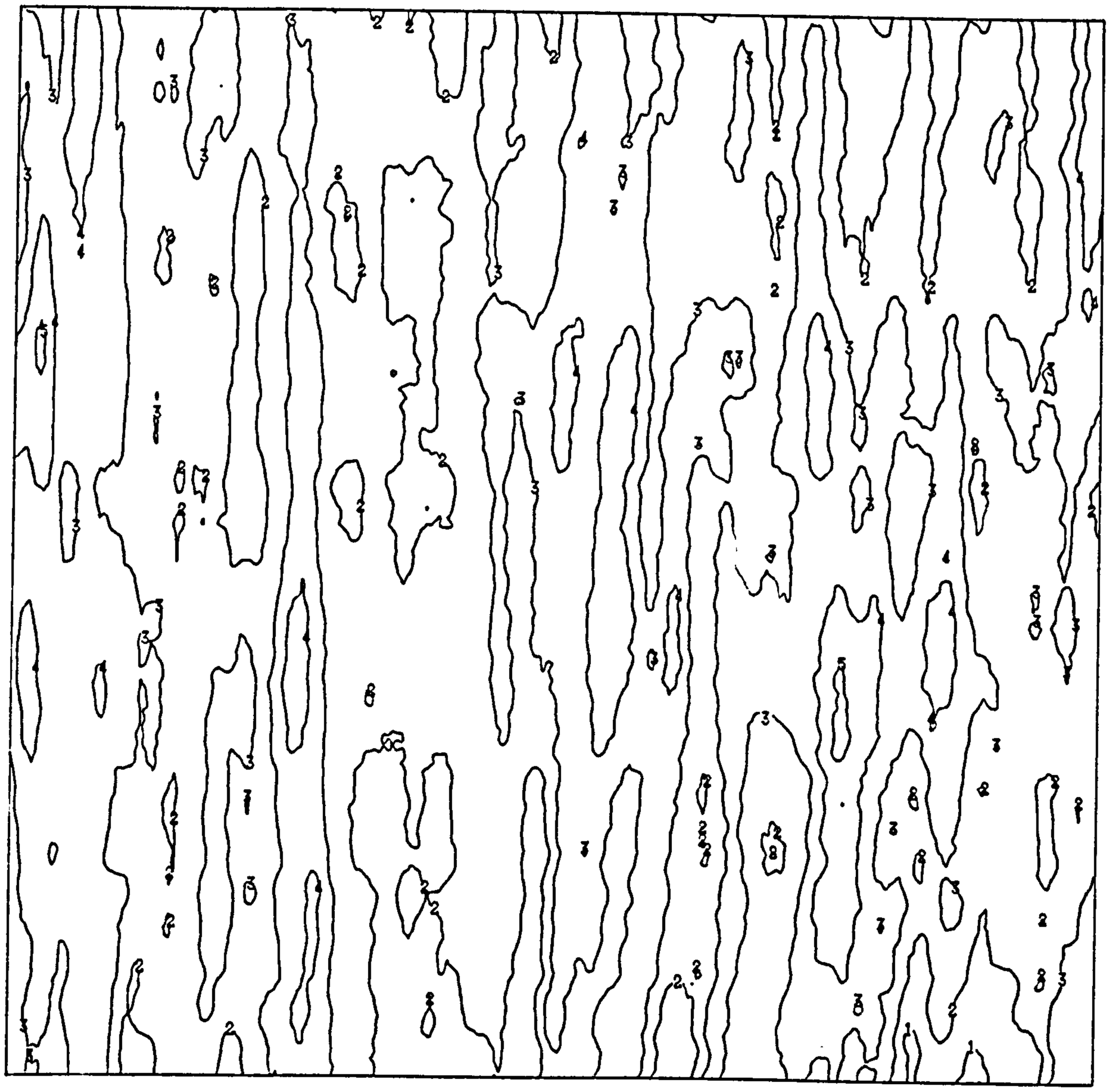


Figure 2.9 Contour map of test specimen 1
(Contours at 0.4 μm height intervals)

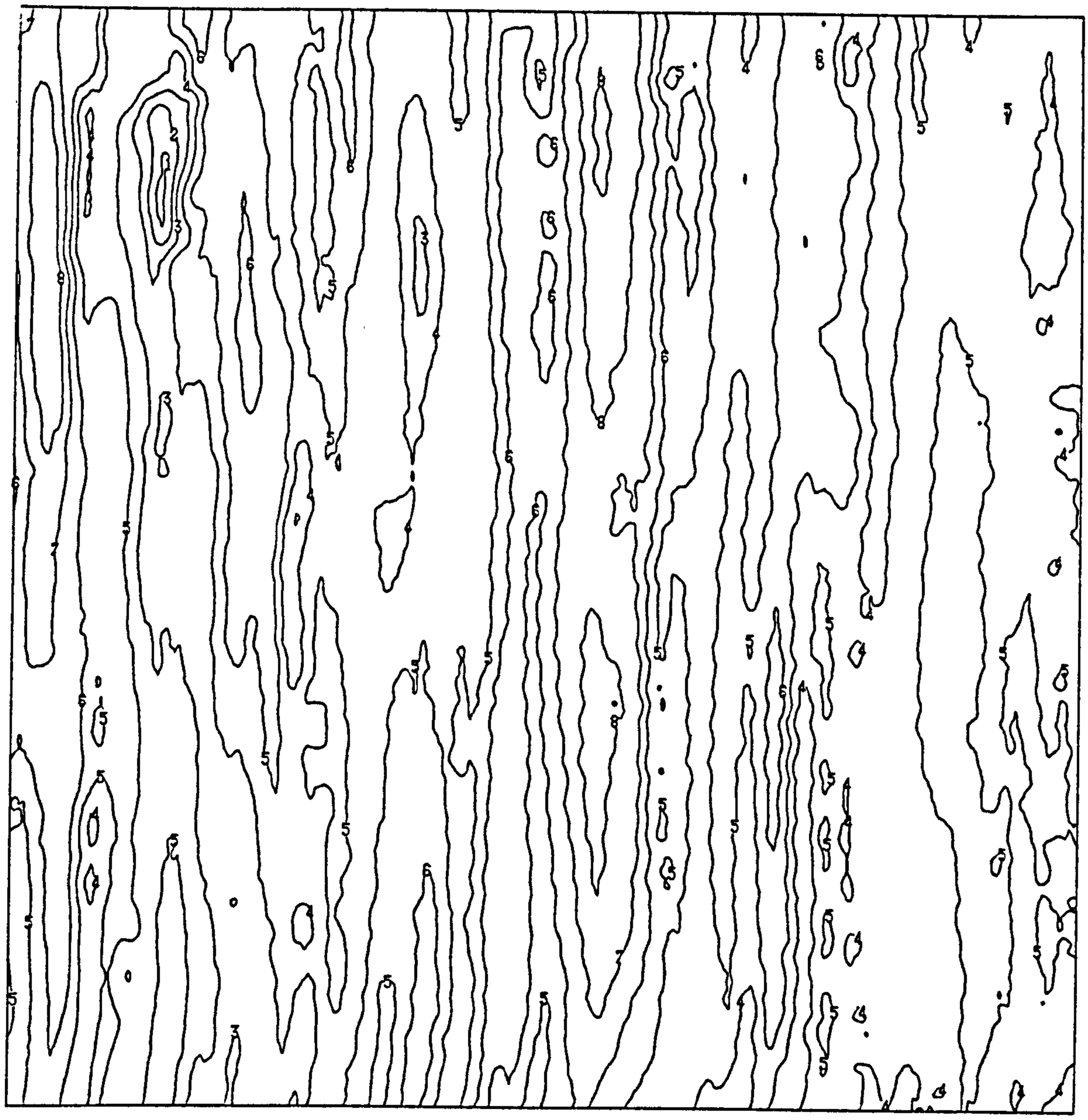


Figure 2.10 Contour map of test specimen 3
(Contours at 0.8 μm height intervals)

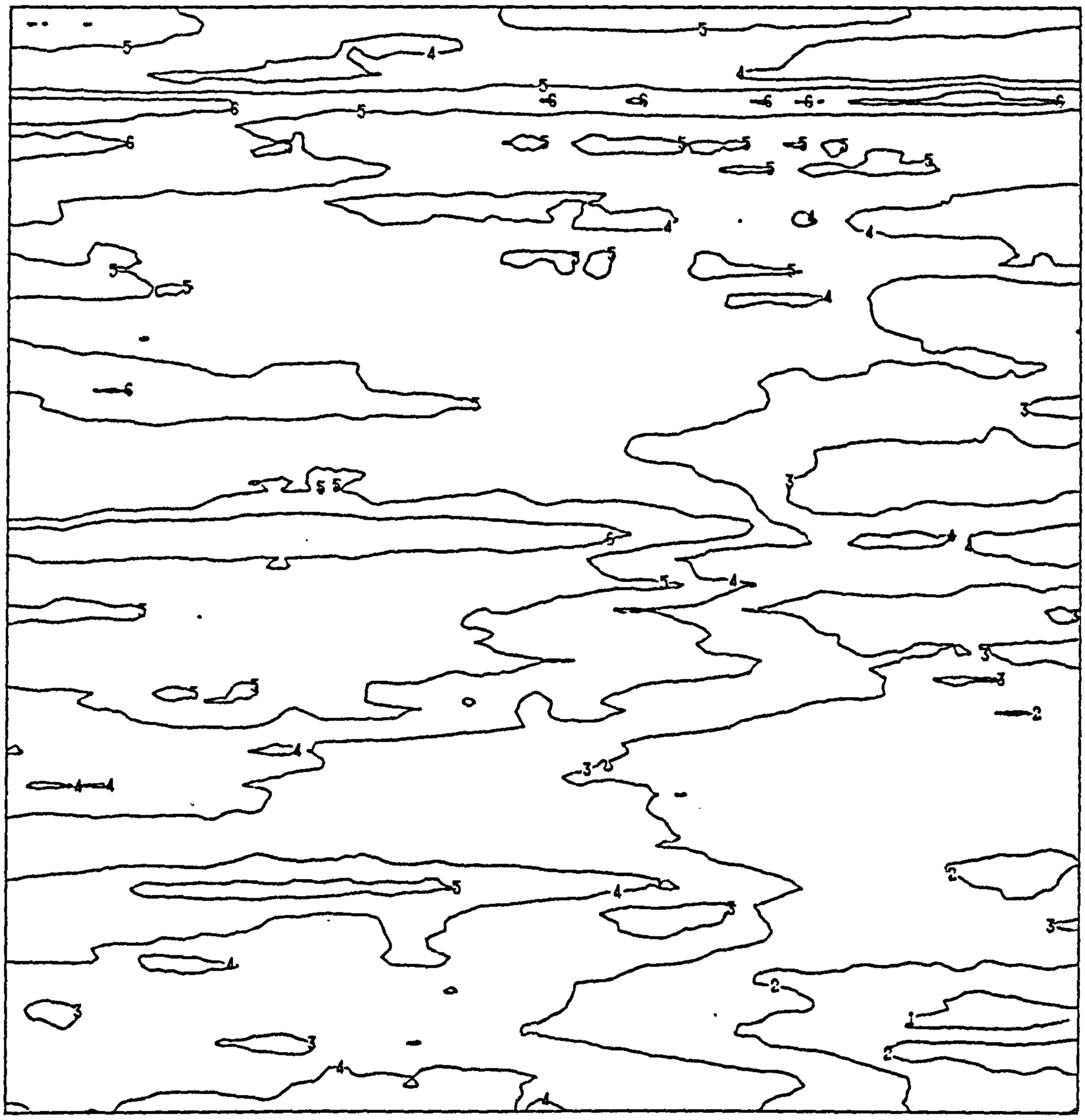


Figure 2.11 Contour map of test specimen 15
(Contours at 0.4 μm height intervals)

CHAPTER 3

THE EXPERIMENTAL EQUIPMENT

3.1 INTRODUCTION

The following sections describe the basic experimental apparatus used to carry out the investigations into the links between surface roughness and vibration in gear contacts, with reasons for its choice. Modifications made at an early stage are also described.

The development of a laser device for monitoring the vibrations of roughened disc specimens rotating in a mini-disc machine is described. First the basic laser system with its principles of operation are detailed. Then, modifications to the system in the light of a vibrational analysis carried out by Westland Helicopters Limited are described. Finally, further tests of the operation of the system, after the modifications had been made, are detailed.

3.2 SIMULATION OF GEAR TOOTH CONTACT CONDITIONS

The general method used to simplify and simulate gear tooth contact, is to replace the gear teeth with two discs

running against each other. This is because the conditions of relative surface motion at any point of contact of any chosen gear tooth combination can be reproduced by a pair of discs of any arbitrarily chosen diameters, by giving them the appropriate respective rolling velocities. The Hertzian stress conditions can then be reproduced by applying the appropriate load between the discs, when the materials of the discs correspond to those of the gears.

However, it should be noted that to correctly reproduce the conditions of oil-film formation, the radii of the discs should correspond to those of the gear tooth profiles at the point of contact being investigated. Even then the simulation is incomplete, because if the discs rotate at constant speed, the rate of change of relative surface motion between actual gear teeth is not reproduced. This gives rise to a significant difference between, in particular, the contact conditions between a pair of discs in pure rolling and the contact of actual gear teeth at the pitch point (Smith, 1983).

Notwithstanding these points, the complexities and added complications of using actual gear teeth for experimental work makes the use of discs inevitable, and therefore a disc machine was employed for the research programme.

3.3 THE BASIC DISC MACHINE

There are many types of disc machine in use, some with more than two discs. That described by Graham (1979)

having as many as four. However, for simplicity a two disc machine was thought to be appropriate for this research.

Earlier experiments with disc machines studying the vibration of disc specimens, by Carson and Johnson (1971), utilised a disc machine where the discs and their bearings were mounted on springs so that accelerometers attached to the bearing housings could monitor the vibration. This approach relies on the disc bearings being infinitely stiff so that any vibration is transmitted to the housing. It also has the disadvantage that the large mass of the whole system will damp out any short wavelength vibration (the vibration expected to come from asperity contact) just leaving the vibration due to the waviness of the disc specimens. This type was therefore not suited to the experiments to be carried out.

Due to the high cost of disc machine specimens, the tendency in recent years has been towards the use of smaller discs. Typically the older type of disc of 75 mm diameter or more is being replaced by so-called mini-discs of diameter less than 25 mm. For vibrational experiments these newer discs have the advantage that as they are of low inertia, their response at higher frequencies is an improvement over the older type. The discs used were therefore of this type. See Chapter 2 for a description of the discs used and their surface topographies.

To be able to monitor the contact vibrations due to the surface topographies, it was important to isolate the

specimens from external vibration sources. To achieve this the discs were supported in hydrostatic bearings, and nylon flexible couplings with driving dogs were used to transfer power from the drive shafts (Figure 3.1). Note that the contact regions of the two specimens were of different widths. This was so that an unworn region would be left on each side of the actual running track for wear measurements to be made.

BIBL.
LONDIN.
UNIV.

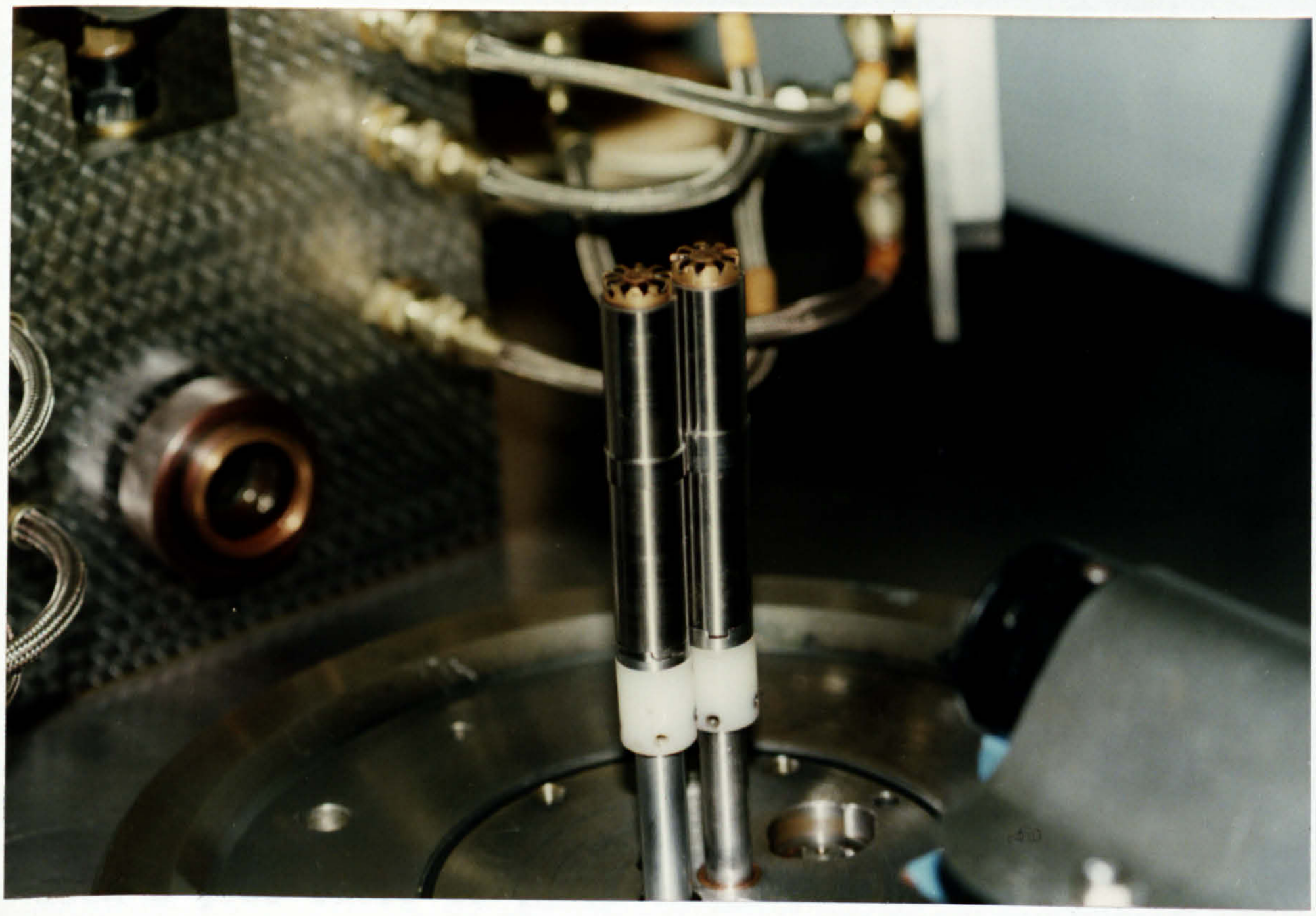


Figure 3.1 The disc specimens and drive shafts

The normal method used to achieve a series of slide-roll ratios to simulate different parts of a gear tooth is by the use of a single motor driving both disc specimens through belts and pulleys. By changing the pulley ratios several relative speeds are obtained. However, this limits the number of different slide-roll ratios that can be obtained from the machine to the number of pulley ratios available. The mini-disc machine that was used for the experiments overcame this problem by having separate variable speed motors driving each disc. This allowed an infinite range of slide-roll ratios to be selected.

The advantage of the pulley system is that the slide-roll ratio once selected is constant even if the motor speed varies. To achieve this with the mini-disc machine used, the D.C. motors were controlled using very stable Thyristor Converters with the input to one being a specified percentage of the input to the other. They were capable of holding the drive shaft speed to within 0.1% for a 100% load change. It should be noted that as the drive to the (potentially) slower disc passed through a gear pair with a gear ratio of 12:26 (Figure 3.2), then either disc could be made to go faster.

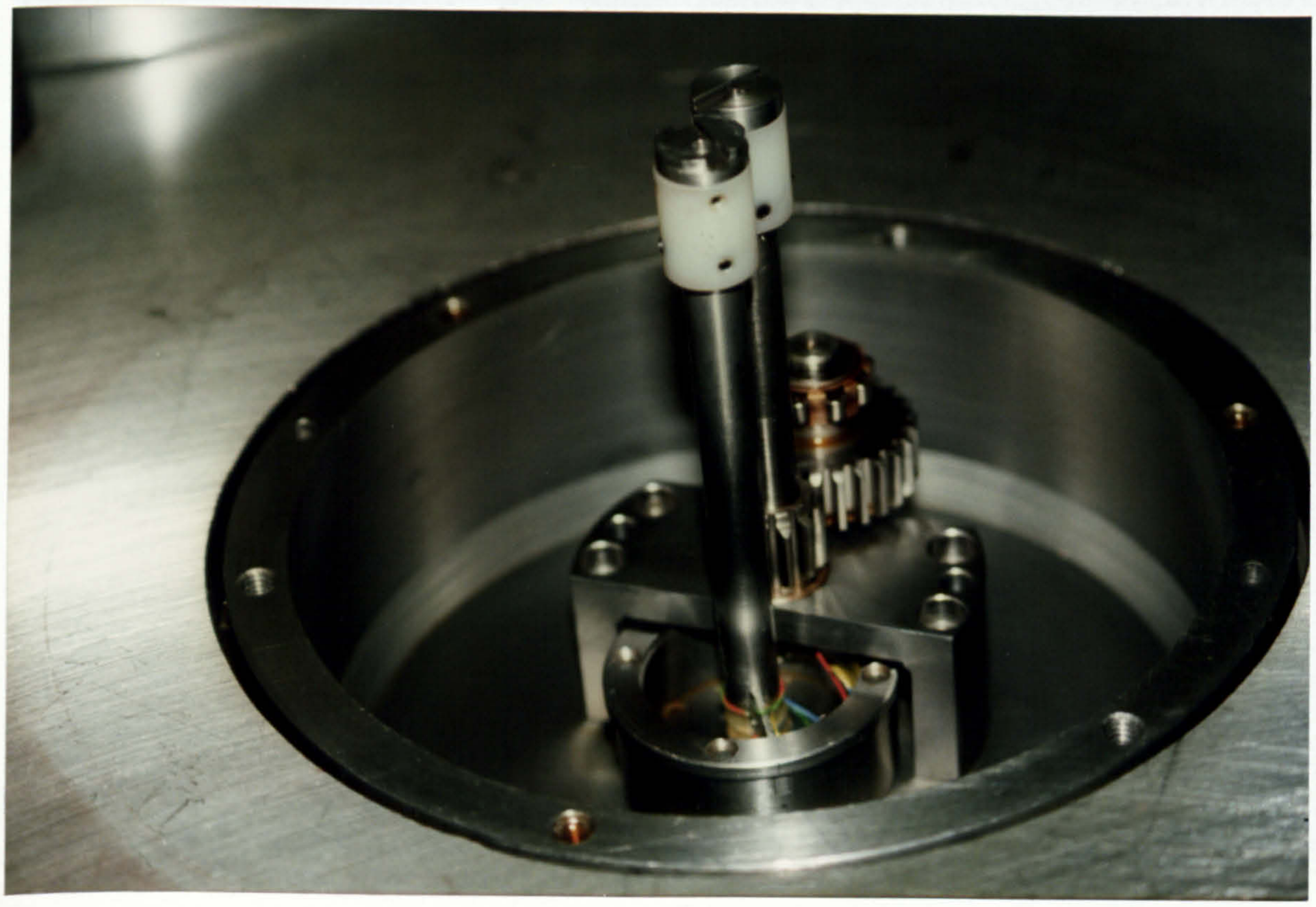


Figure 3.2 View of torque monitor and gear pair in drive

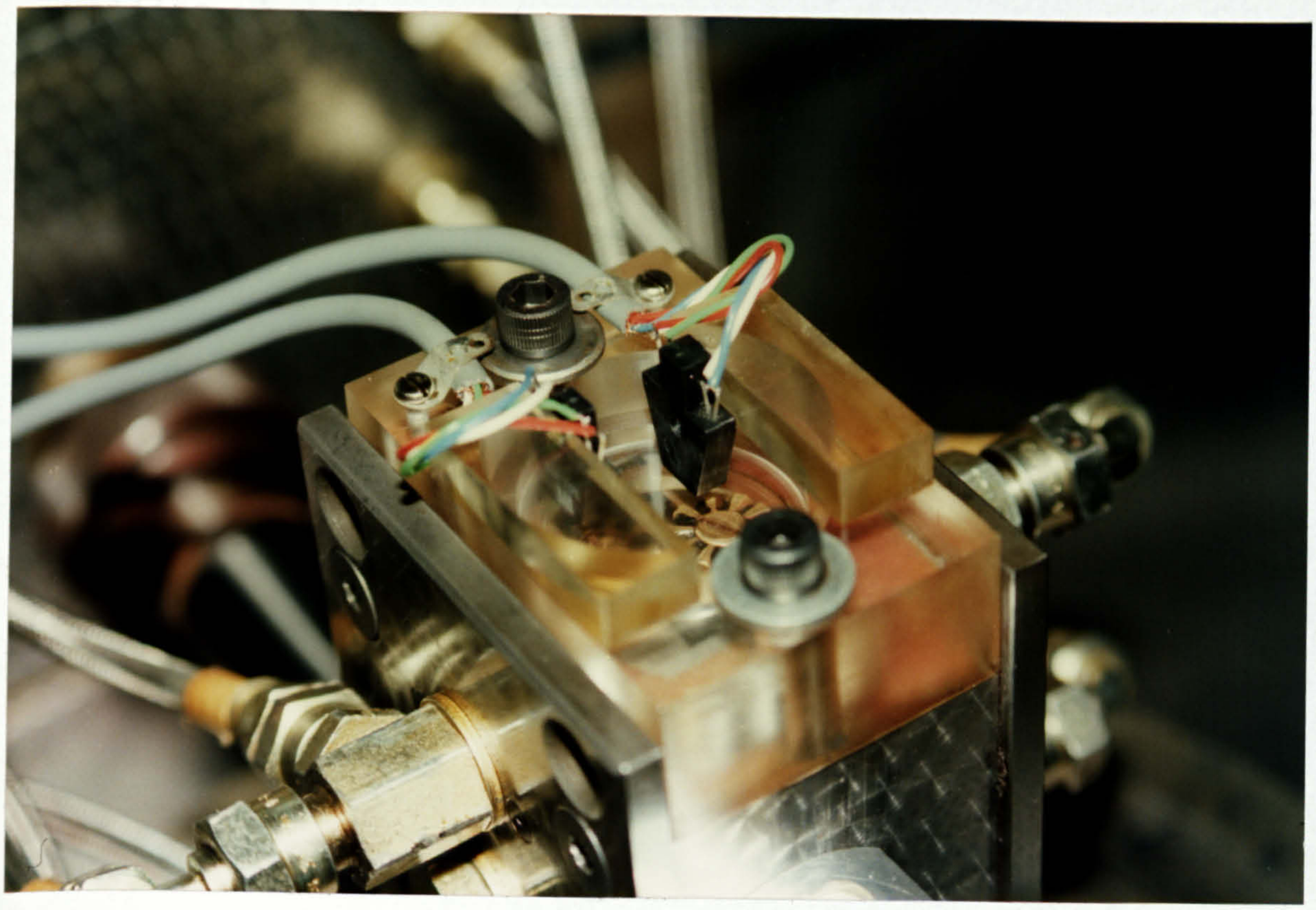


Figure 3.3 View of disc speed indicating device

3.3.1 Accurate Disc Speed Indication

It was realised at an early stage that the speed indicators supplied on the disc machine (two revolution counters showing 0 to 3000 RPM) would not be adequate to give an accurate assessment of the running conditions. This was particularly so when operating near to pure rolling when it was necessary to ascertain the exact slide-roll ratio. Thus it was decided to attempt to monitor the disc speeds directly by modifying the top of the specimen housing.

Brass discs with ten equally spaced grooves cut into them and filled with black paint were screwed to each disc specimen (Figure 3.1). The top of the specimen housing was replaced by a perspex window, through which two photo diodes could distinguish when the blackened grooves passed beneath as the discs rotated (Figure 3.3). The output signal from each diode was thus in the form of a rough square wave with its wavelength equal to ten times the rotational speed of that particular disc.

To actually measure this frequency, a variable frequency signal was combined with the diode output and the phase difference between the two signals recorded. If the two signals were of different frequencies, then the phase difference between them constantly varied. However, if the two frequencies were identical, then the phase difference was constant and this could be deduced from a stationary phase lag indicating pointer. The frequency of the reference signal and hence the disc was then found

using a simple frequency counter.

It should be noted that, using this method, the actual speed could be any multiple of the frequency indicated by the counter. However, as the original revolution counters gave an approximate indication of the speed, this problem was easily avoided.

3.3.2 Torque Measurement

As supplied the mini-disc machine was fitted with a strain gauged bar torque monitoring device in the final drive to one of the specimens (Figure 3.2). This had been previously calibrated by Westland Helicopters Limited (Jessop, 1984). On running the disc machine this device quickly failed due to the oil being used attacking the strain gauges. From the original calibration of the device it was seen to be non-linear.

An attempt was made to improve the linearity of the torque monitor by modifying the end mountings of the strain-gauged bar to produce point contacts. After new strain gauges had been attached the whole of the cup containing the bar was filled with silicone rubber. This seemed to prevent oil attack and did not have an adverse effect on the unit's performance. A new calibration was then carried out which showed that the unit was more linear than before (Figure 3.4).

However, a short time into the experimental programme the unit failed again, and as it did not seem to be giving

any useful data, other than confirming the pure rolling condition, its use was not continued.

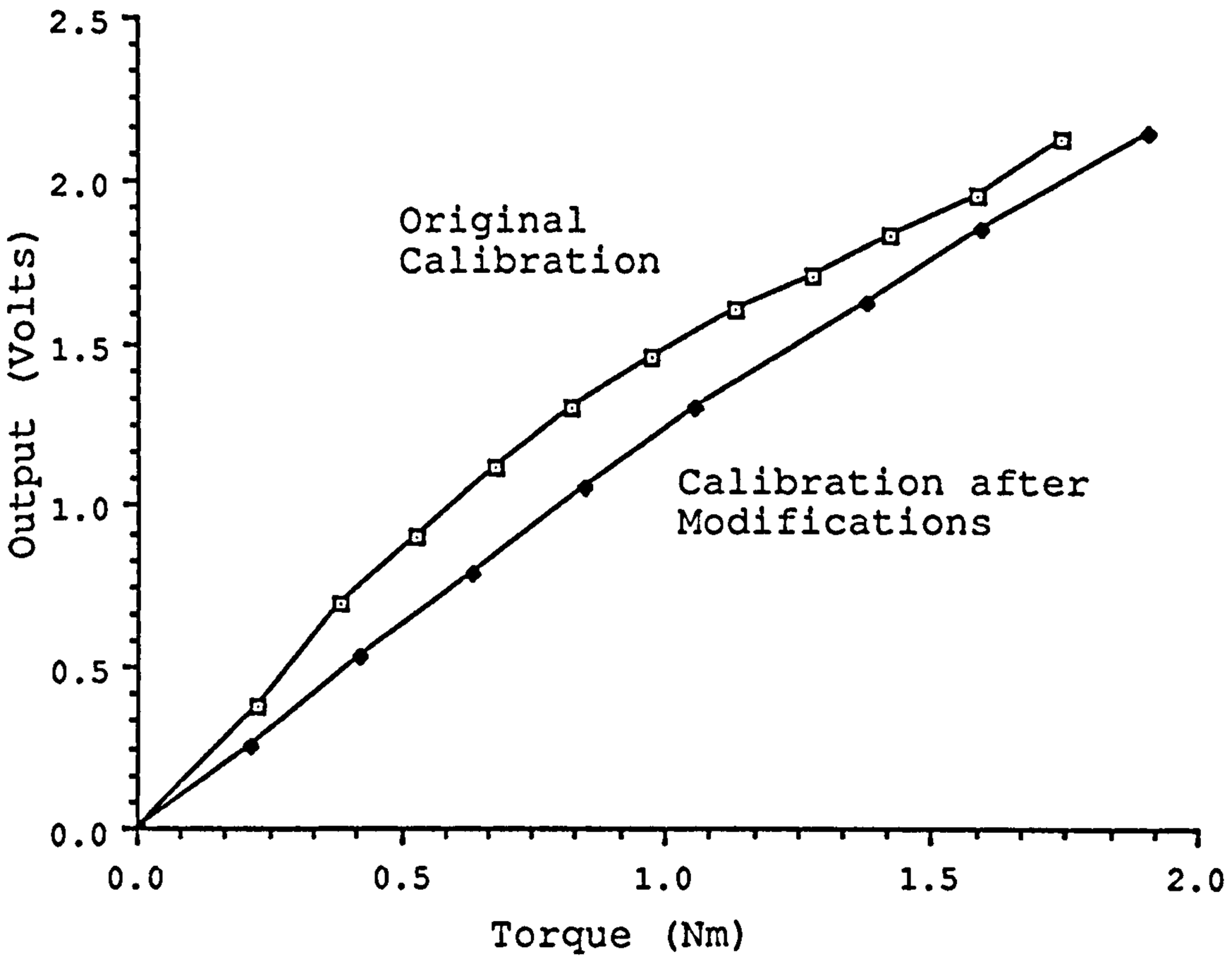


Figure 3.4 Calibrations of the torque monitor

3.4 THE VIBRATION MONITOR

The following sections describe the development of a laser device for vibration monitoring on the mini-disc machine.

3.4.1 The Basic Wear and Vibration Monitor

The simplest form of vibration monitor for most applications is the accelerometer. Unfortunately the use of accelerometers for monitoring the contact vibrations between the two disc specimens was not possible. The principle of using hydrostatic bearings for the support and loading of the discs was to isolate them from external vibration sources, hence also isolating any accelerometers mounted on the casing from the contact vibration. To overcome this problem an optical technique was employed to monitor the contact vibration.

Figure 3.5 gives a view of the device attached to the specimen housing of the disc machine, while Figure 3.6 shows detailed sections of the device as described below.

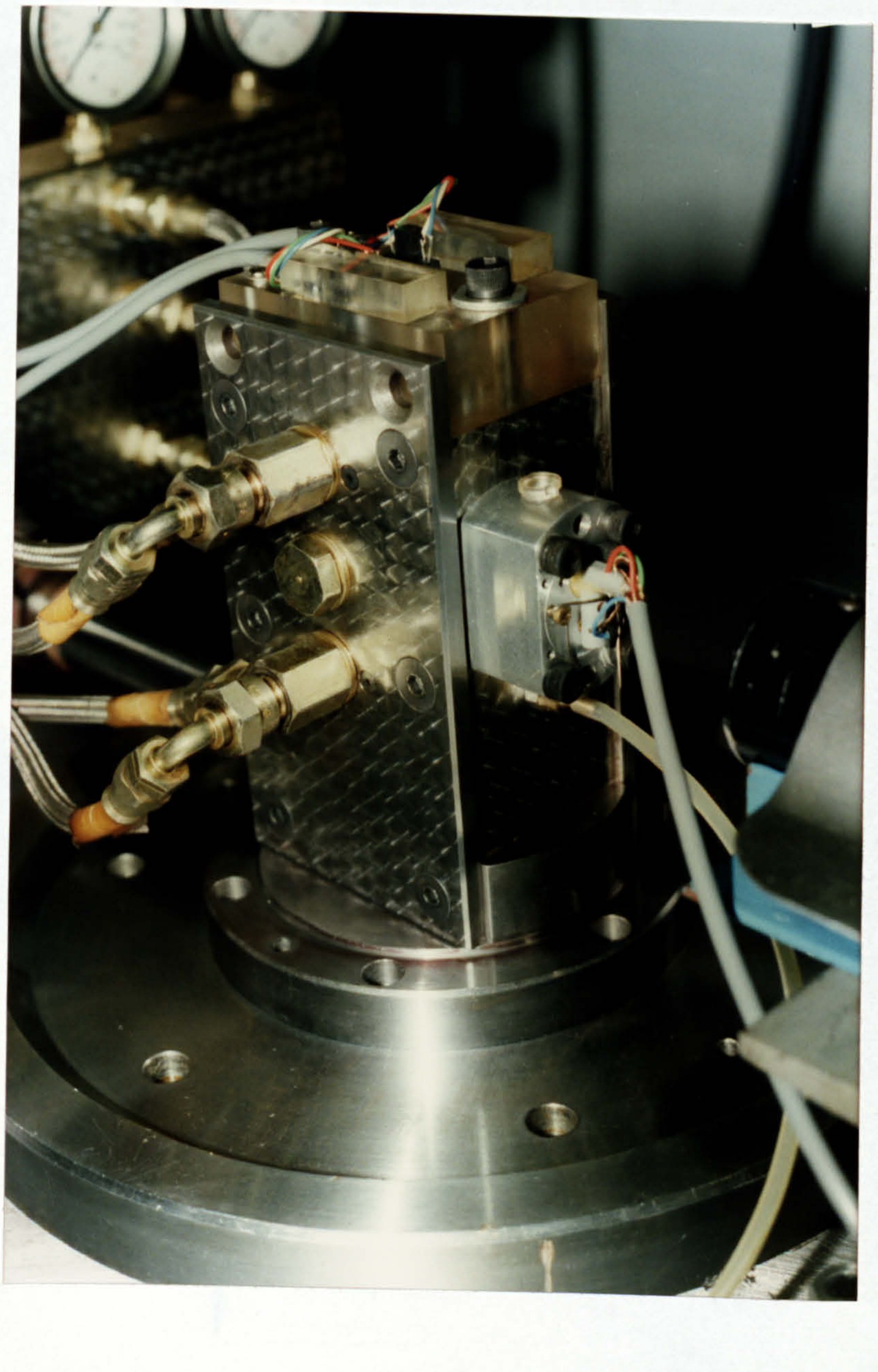


Figure 3.5 The vibration monitor attached to the specimen housing

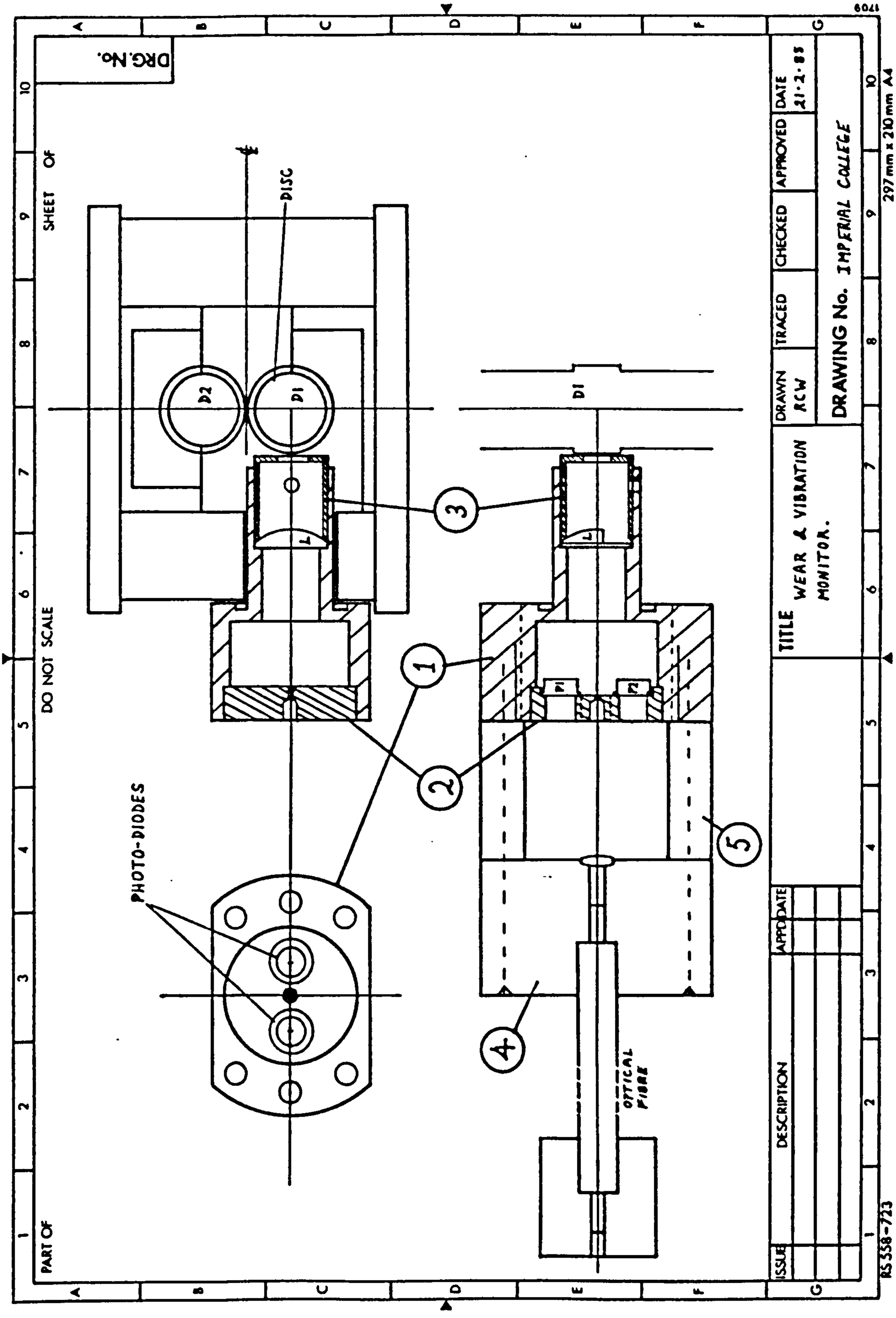


Figure 3.6 Detailed Sections of vibration monitor

The principle of operation of the device was that light shone onto one of the disc specimens at right angles to their line of centres would be reflected back in different directions, dependent on the position of the centre of that disc in the direction of the line of centres (see Figure 3.7). Hence the movement of the disc towards and away from the other disc, due to the contact between them, could be monitored by detecting where the light beam was reflected back onto the detector.

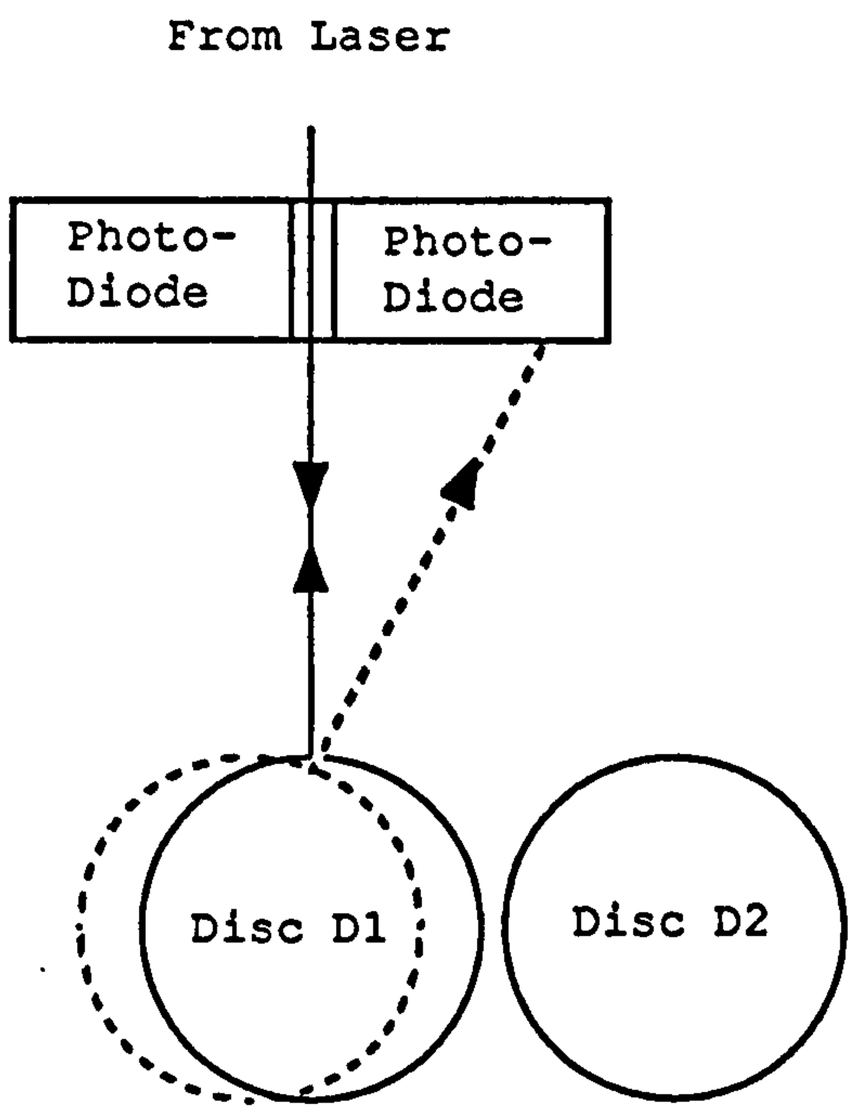


Figure 3.7 Principle of operation of the vibration monitor

The light source used was a HeNe Laser, the beam from which was originally intended to pass through an optical fibre into the device. This would have meant that any

vibration of the laser itself relative to the specimen housing would be eliminated. However, internal speckling caused by the vibrating optical fibre made its use undesirable, so the beam was shone directly into the device, with the laser being firmly clamped to the disc machine frame, as close to the specimen as possible.

Referring to Figure 3.6, the laser passed into the lens L then onto the disc specimen D1. The light reflected off the disc then passed back through the lens and was focussed into a line (due to the cylindrical disc surface) on the detector P1. This detector consisted of a double diode silicon chip, so that when the laser line was correctly aligned the two side by side diodes gave equal output signals. If the disc D1 moved towards the other disc D2, the laser line moved across the diodes and caused one diode to give a greater output than the other, the opposite being true if the disc D1 moved away from disc D2. This differential output from the two diodes being roughly proportional to the disc displacement, allowing movements of less than 2 μm to be detected. The overall range of movement that the device could cope with was approximately 200 μm.

The output from the two diodes was used in two ways:-

1) The A.C. output gave the vibration of the disc D1 about its average position

2) The difference between the A.C. and D.C. outputs gave the average position of disc D1 at any time. The

change in this average position could be used to monitor the wear of the discs, since as wear took place the centres of the two discs became closer together.

The lens L was cut away at the bottom to allow light from the disc to return directly to the PIN diode P2, which measured the speckle pattern caused by the topography of disc D1. The output of the diode P2 could therefore be used as a monitor of the disc surface roughness during the vibration experiments.

The output signals from the vibration and roughness monitors were recorded via an analogue to digital converter into an LSI 11/03 micro-computer. The maximum recording speed was 15,000 data points per second allowing one data point to be recorded every 40 μm around the disc's circumference, at an operating disc speed of 10 Hz. The output signals could then be analysed using a suite of signal analysis programs available on a LSI 11/23 micro-computer.

3.4.2 Initial Tests on the Laser System

To test the vibration monitoring system a test specimen was produced. This consisted of a normal disc with an axial groove scratched into it with a hardened lathe tool. Figure 3.8 shows a cross-section of the indentation at its deepest point. The depth was approximately 20 μm and width 300 μm. This specimen was then used as disc D1.

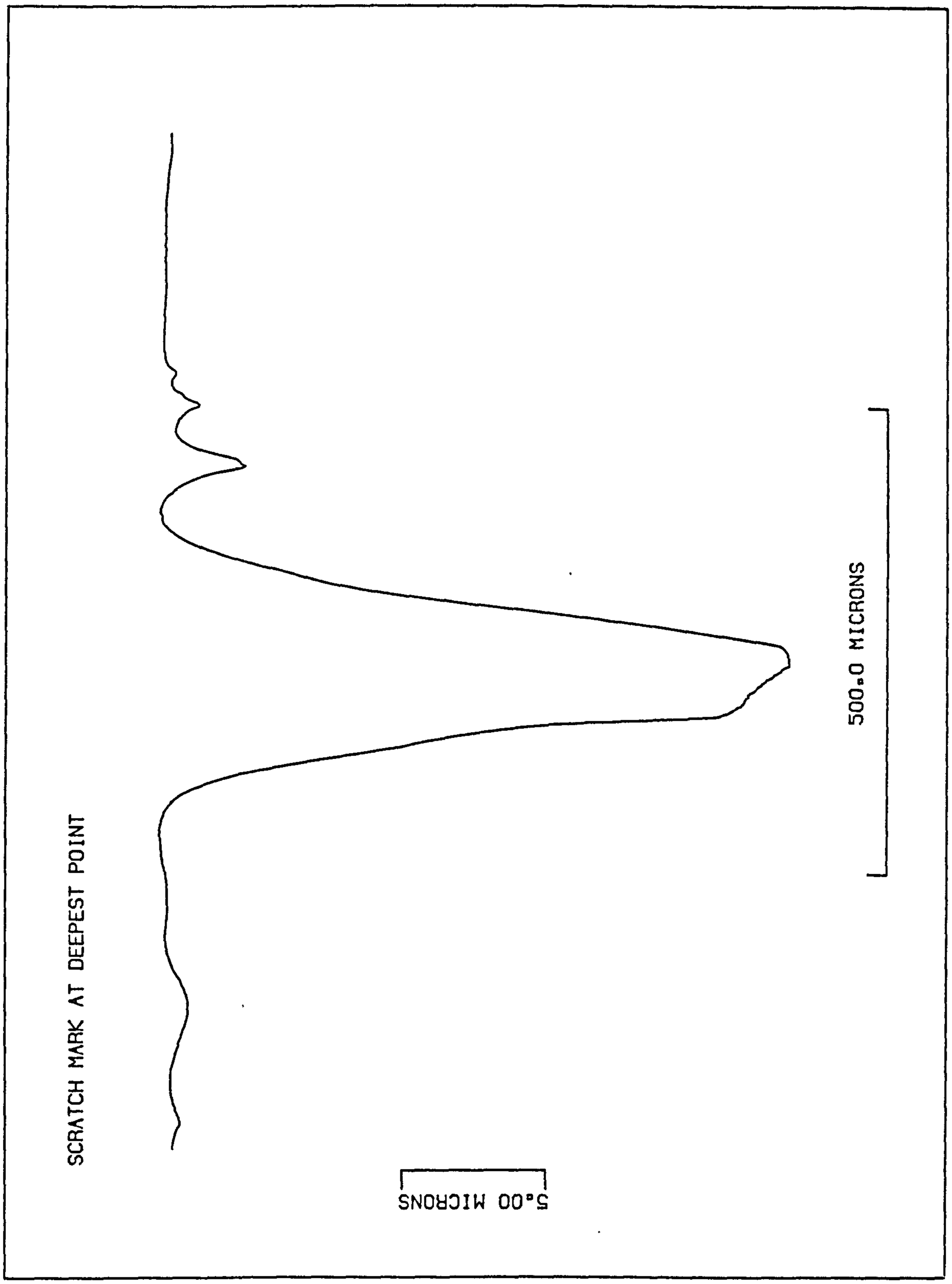


Figure 3.8 Cross-section of groove in calibration disc

When the disc machine was operating with the test specimen, both the vibration signal and surface finish signal showed a once per revolution feature. However, as these two signals were exactly in phase it was thought that the light beam being used for the vibration signal was being deflected as the laser light fell on the indentation, rather than when the other disc came into contact with the indentation.

To check whether this was the case, Westland Helicopters Limited were invited to carry out a vibration analysis of the disc machine using accelerometers. Their report (Stucky, 1985), concluded that accelerometers could be used to detect the vibration caused by the indentation above the noise generated by the machine, and that the optical vibration monitor was only detecting the deflections of the laser beam as the indentation came into view. The system had therefore to be modified before any experiments could commence.

3.4.3 Modifications to the Laser System

The failure of the laser system was thought to be due to four main causes:-

1) Vibration of the lasing tube within its mountings and of the mountings themselves causing a spurious vibration signal.

2) The hydrostatic bearing system was too stiff, i.e. the

gaps between the hydrostatic pads and the specimens were too small. Thus any vibration of the specimen casing was being transferred to the discs, and vice versa. The casing was also having the effect of damping out the contact vibrations from the specimens.

3) Oil from the hydrostatic bearings was passing across the laser beam and an oil film forming on the disc itself. This had the effect of distorting the laser beam, thus giving spurious results.

4) The surface topography of the specimen onto which the laser light fell was affecting the vibration signal by deflecting the beam.

To overcome the first problem the laser being used was replaced by a new sealed 2 mW HeNe laser, where the lasing tube was solid within its casing. The casing of this laser was of stronger construction than the previous laser, so that it was possible to clamp it tighter to the disc machine frame. It was also mounted as close as possible to the specimen housing, so that the effect of any adverse vibrations would be minimised.

The second problem was dealt with by inserting steel shims in the specimen housing to increase the clearance between the hydrostatic pads and the specimens. After adding the shims the no load clearance between each pad and the specimens was approximately 25 μm. This had the effect of decreasing the bearing stiffness.

The third problem, an oil film forming on the disc in

the path of the laser beam, was dealt with by modifying the existing oil shield (shown as (3) in Figure 3.6). This shield consisted of a cylinder with a hole in it for the laser light to pass through, a clean air supply being fed into it to maintain a stream of air passing through the hole to clear the disc surface.

The shield was modified by the addition of a nylon guard, attached to the front of the cylinder, shaped to conform to the disc surface, but kept approximately 1 mm away from it (see Figure 3.9). This allowed the air supply to be more effective in keeping the disc surface clear of oil.

The fourth problem was more difficult to overcome. The ideal solution would have been to use a perfectly smooth polished specimen for the laser light to shine on. However, trials with such a specimen revealed that this would not work as the light reflected from the specimen had to fall on detectors above and below the centre line on which the light reached the specimen (Figure 3.6). With the polished specimen, the light was reflected directly back and no light fell onto the detectors.

The solution to this problem was to use normally ground discs all produced at the same time so that their topographies would be as similar as possible. The circumferential grinding marks would reflect the laser light towards the detectors, whilst the similarity between these reference discs would ensure that their use did not adversely effect the vibration signals. Each reference

specimen would also be calibrated with a static disc movement to eliminate any effects of differing surface topography or reflectivity.

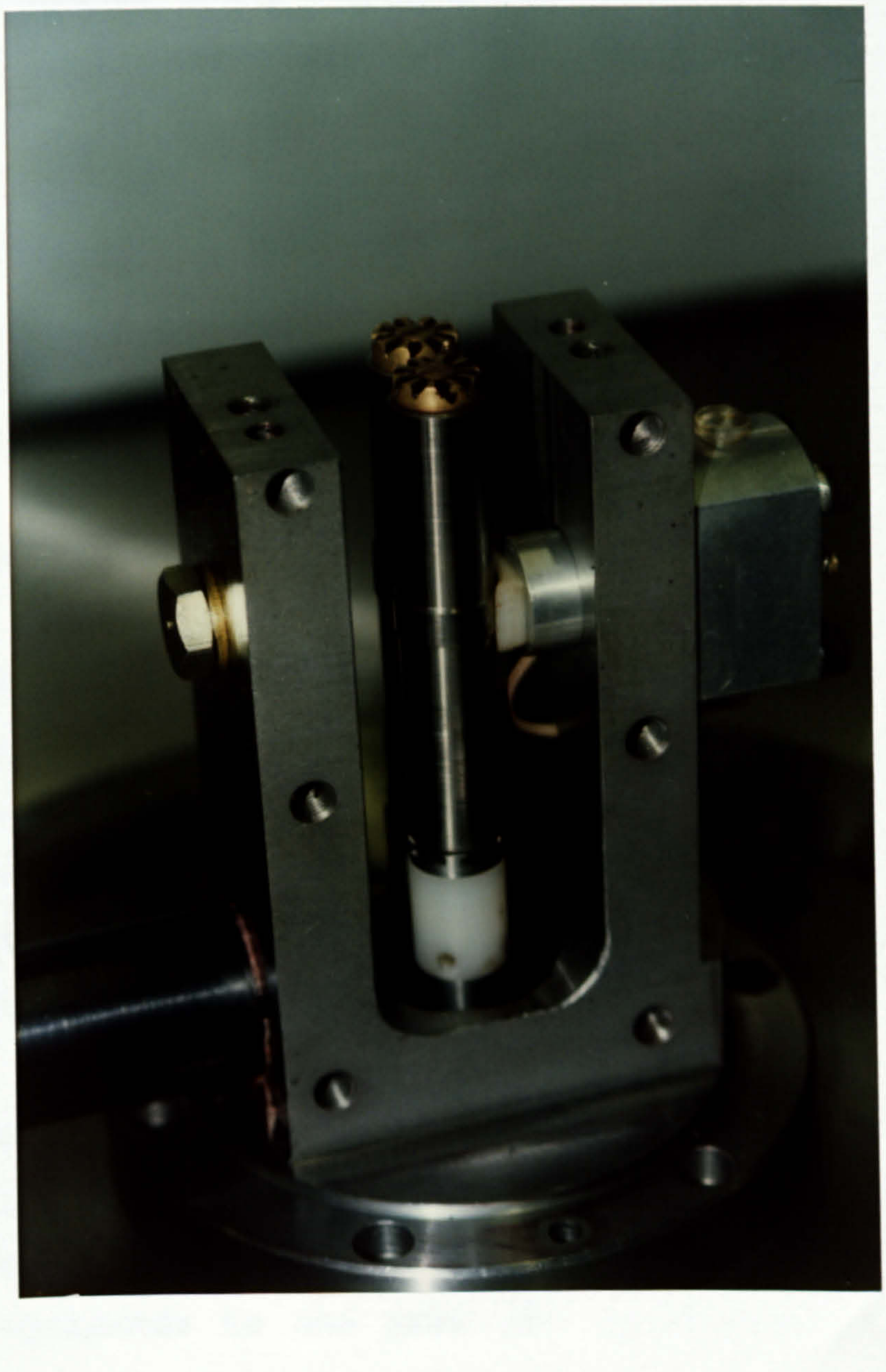


Figure 3.9 Modified oil shield attached to the vibration monitor

3.4.4 Tests on the Modified Laser System

To test the equipment before carrying out the main experiments, the test specimen described previously in Section 3.4.2 was used again. However, before it was used it was polished using 1 μm diamond paste so that there were no features on its surface other than the indentation.

The specimen was then run on the mini-disc machine as specimen D2, with the laser light falling on a reference disc as described in the previous section. Before running the vibration signal was calibrated by a static movement of the disc.

The operating conditions were as follows:-

Speed of indented disc = 10 Hz

Speed of reference disc = 10 Hz

Contact pressure = 1.02 GPa

Hertzian contact width = 215 μm

Maximum indentation depth = 20 μm

Indentation width = 300 μm

Figure 3.10 shows the vibration output signal for two revolutions of the test disc, with the response to the test disc clearly visible as two sharp peaks in the signal. This corresponds to one peak per revolution of the test specimen as the indentation crosses the contact zone.

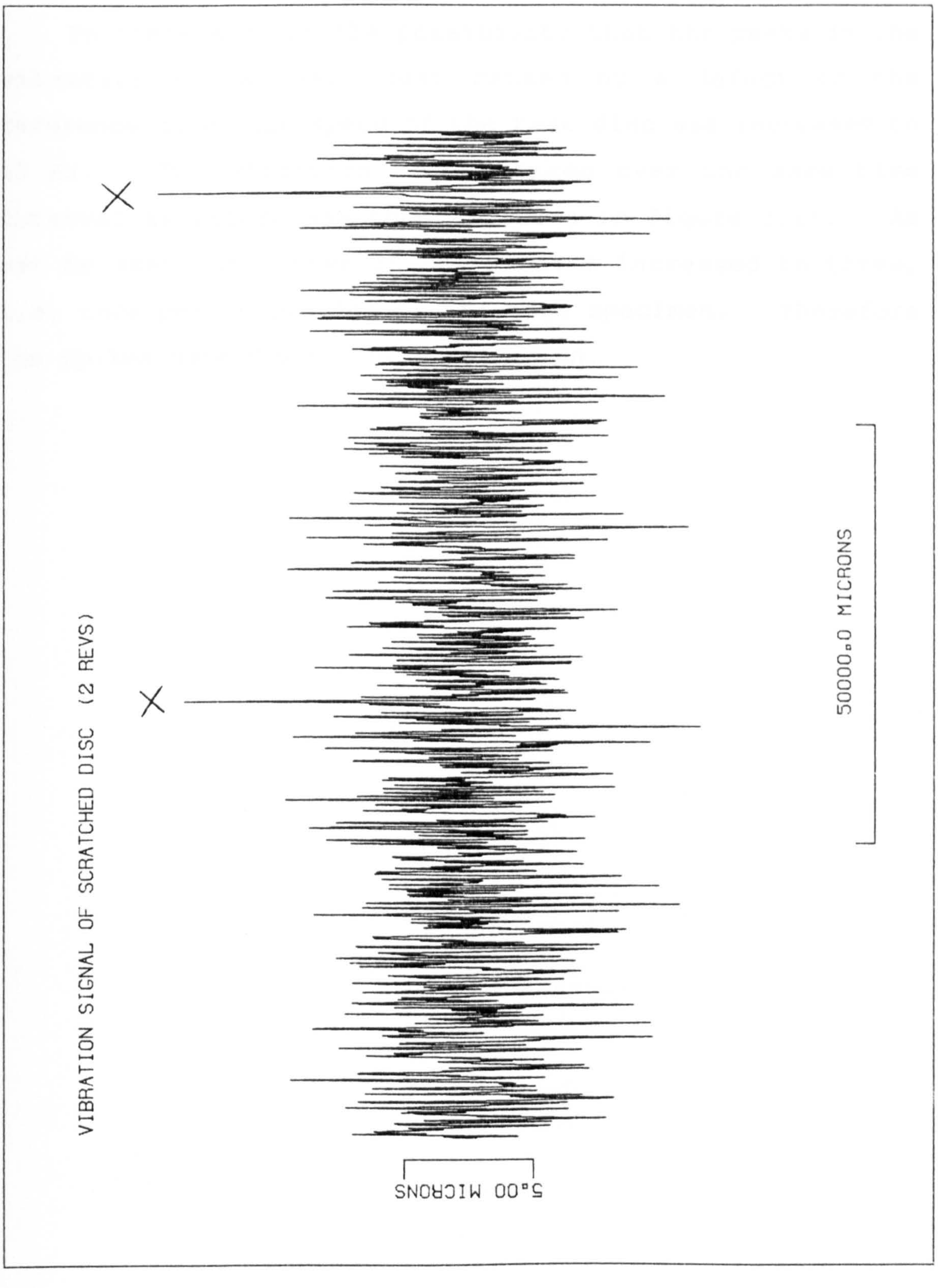


Figure 3.10 Vibration output for 2 revolutions of the test disc

To check against the possibility that the peaks in the vibration signal were just caused by a defect on the reference disc, the speed of the test disc was increased to 15 Hz. The vibration signal taken over the same time interval as before was then as shown in Figure 3.11. As can be seen the number of sharp peaks increased to three, i.e. once per revolution of the test specimen. Therefore the spikes were due to the indentation.

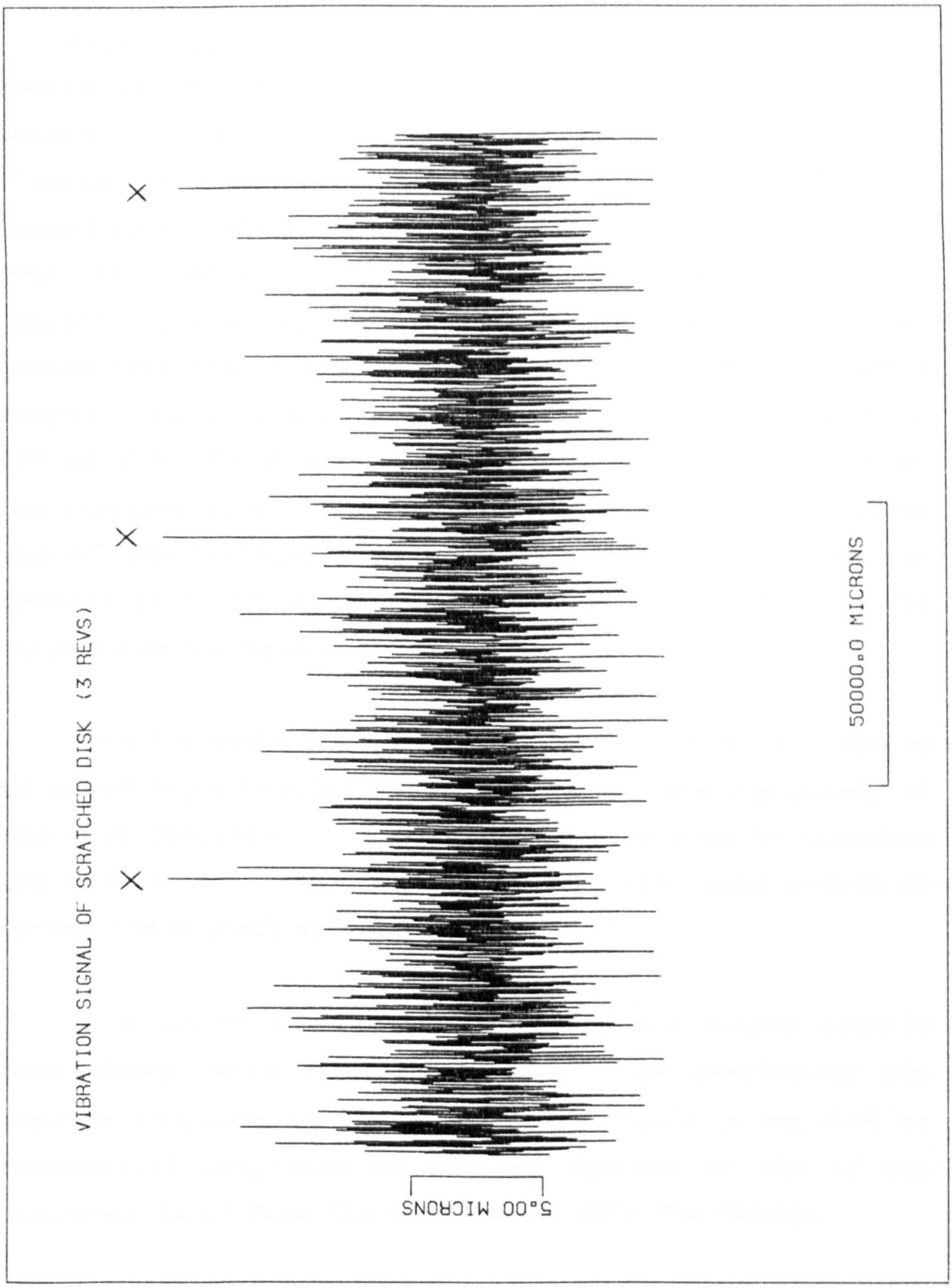


Figure 3.11 Vibration output for 3 revolutions of the test disc

From Figure 3.10 the movement of the reference disc centre is of the order of 12 μm due to the indentation. However, it should be noted that this movement is not towards the test specimen but away from it. Thus it is caused by the shock impulse as the reference system passes over the indentation. Figure 3.12 shows the detail of the vibration signal at the point where the reference specimen passed over the indentation. This looks very much like a damped impulse signal with a wavelength of approximately 500 μm (1200 Hz at a disc speed of 10 Hz). The fact that the response shown in Figure 3.11 with 20% relative sliding was of approximately the same magnitude even though the general level of vibration was higher, also indicates the cause to be an impulse.

Thus the tests proved that the laser system was capable of detecting vibration caused by the surface topography of the test specimen. It was therefore decided to commence the main experimental programme using the laser system to detect the contact vibrations.

It would be advantageous to study the output signals from these tests in more detail. In particular the impulse response to the indentation could be matched to theoretical responses to similar systems to see if the frequency found from the experiment, fits the theory.

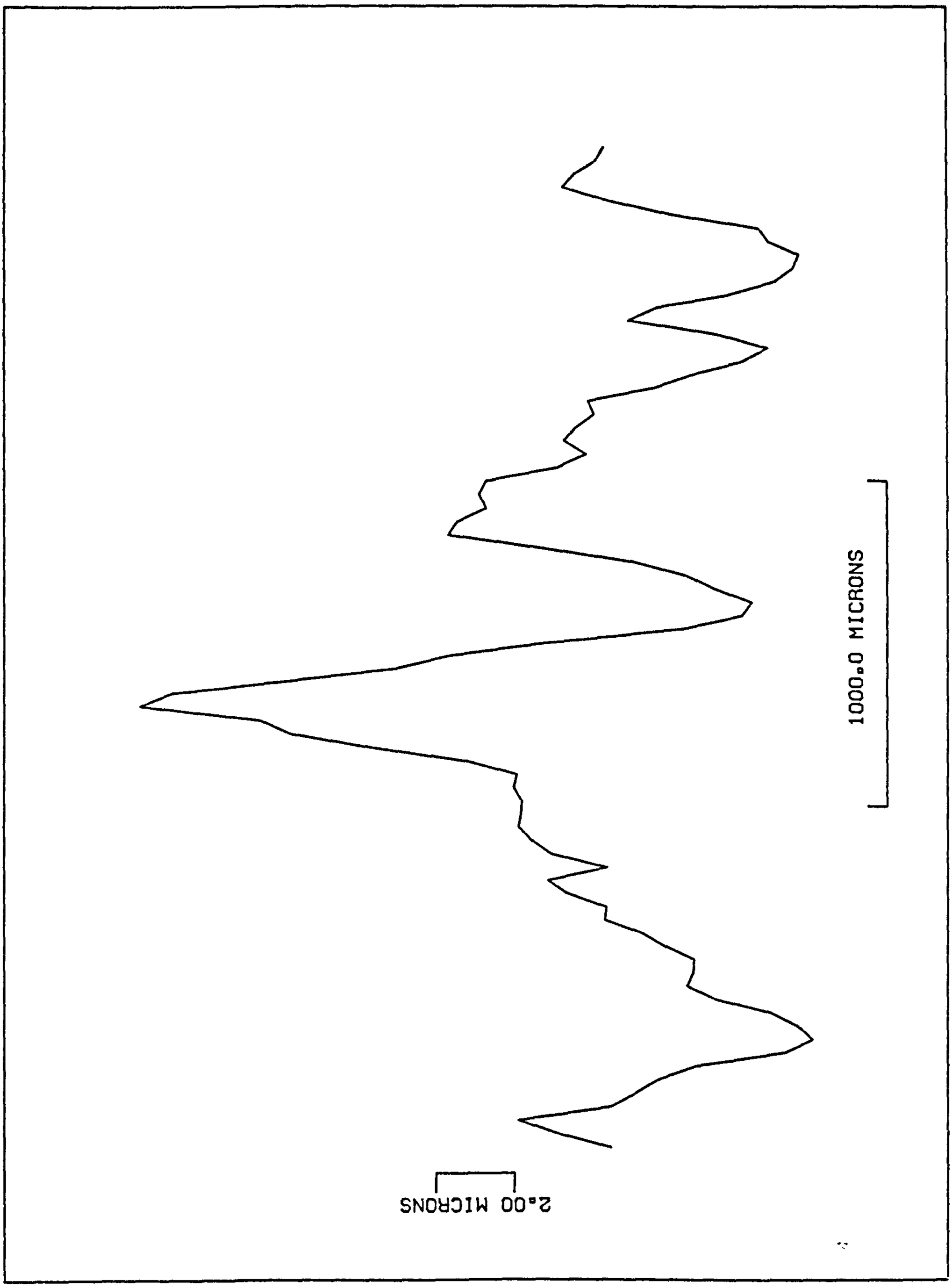


Figure 3.12 Detail of the vibration output as the groove comes into contact with the reference disc

CHAPTER 4

EXPERIMENTAL RESULTS AND ANALYSIS

4.1 INTRODUCTION

After the measurements of the fifteen pairs of test discs had been completed the main experimental programme was undertaken. Each pair of discs were mounted in the mini-disc machine as described in Chapter 3. The machine was then operated with various contact conditions and the pertinent parameters recorded. The test conditions, experimental results and subsequent analysis are described in the following sections.

4.2 TEST CONDITIONS

Conditions applying to all the tests on the 15 pairs of specimens were:-

Speed of test specimen = 10 Hz (0.6 m/s surface speed)

Lubricant = E.T.0.25 ENGRD2497

(5.5 cst synthetic turbine oil)

Bulk oil temperature = 45-47° C

Conditions varied for each specimen:-

Speed of reference specimen in range 10 to 15 Hz
Giving slide-roll ratios of 0 to 0.2
Mean contact pressures of 1.02 to 1.56 GN/m^2
Giving Hertzian contact widths of 215 to 330 μm.

For each pair of specimens, tests were carried out with a slide-roll ratio of zero and mean contact pressures of 1.02 GN/m^2, 1.32 GN/m^2 and 1.56 GN/m^2. Then at a mean contact pressure of 1.02 GN/m^2 tests were carried out at slide-roll ratios of 0, 0.048, and 0.2. Note, specimens 1, 3 and 15 additionally had tests carried out at slide-roll ratios of 0.091, 0.130 and 0.167.

During each of these tests, two recordings of the vibration monitor output were made, each covering one revolution of the test specimen. The average position of the reference disc and the average speckle output were also recorded.

As stated in Chapter 3 the torque monitoring device failed at an early stage of the experiments. However, it seemed to be indicating that at low values of slide-roll ratio (0.048) the coefficient of friction, μ was approximately 0.14, rising to 0.22 at a slide-roll ratio of 0.2.

4.3 POSITION AND SPECKLE RESULTS

The results from the average position of the reference disc did not show that any significant wear took place during the experiments (the initial readings could be obtained by returning to the original contact conditions).

Talysurf measurements taken of the specimens after the completion of the experiments indicated that on average the root mean square (RMS) roughness, σ of the specimen topographies had been reduced by 20%. Figure 4.1 shows the final RMS roughness divided by the original RMS roughness plotted against the original RMS roughness for all the test specimens. Figure 4.2 shows the structure functions of a typical specimen before and after the tests, the smoothing being indicated by a reduction in the structure function. The fact that the wear indicator signal had not changed would suggest that no material removal was involved in this smoothing of the topographies. It should be noted that each specimen was subjected to approximately 18,000 cycles during the experiments.

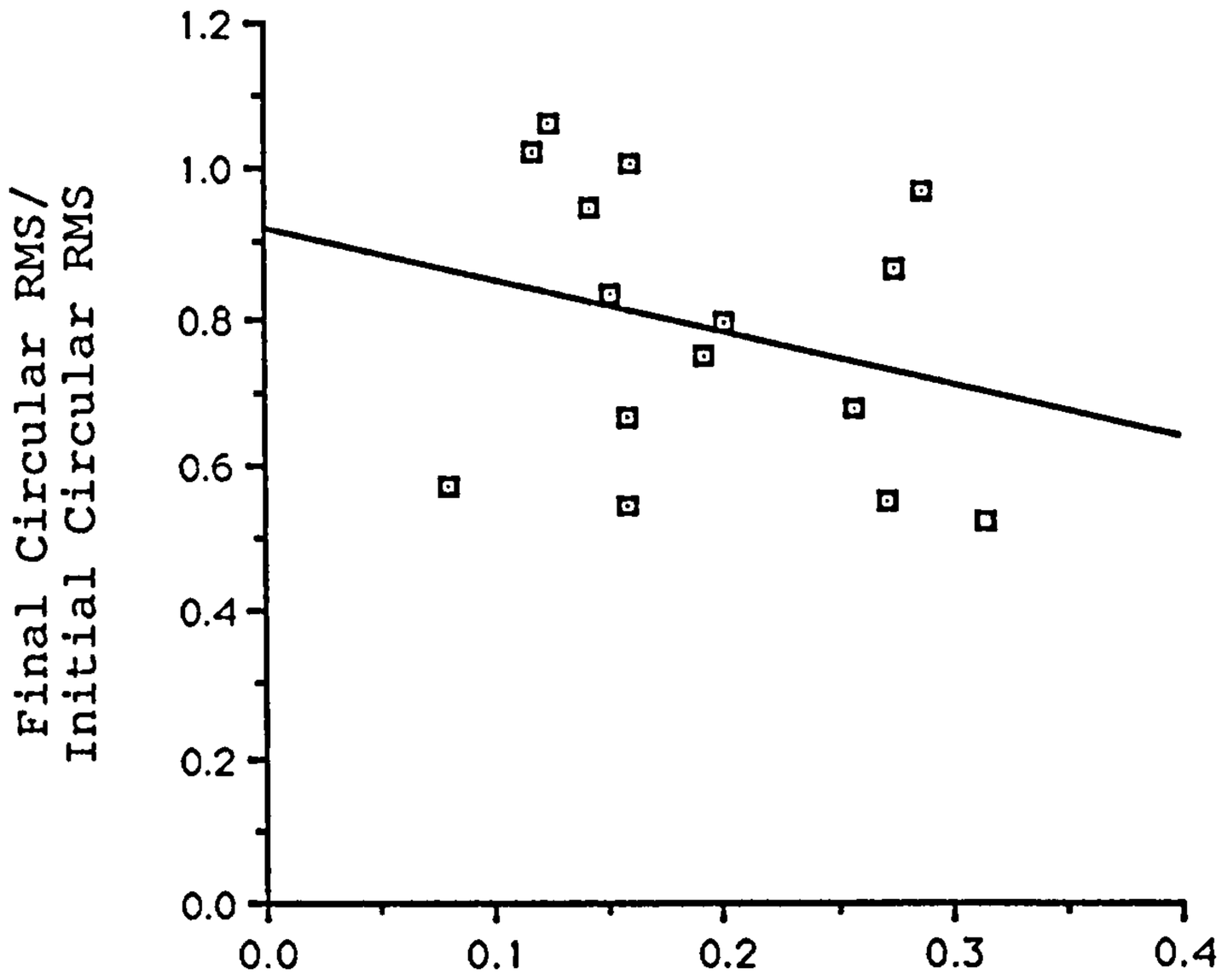


Figure 4.1 The change in the surface roughness of the test specimens during the experiments

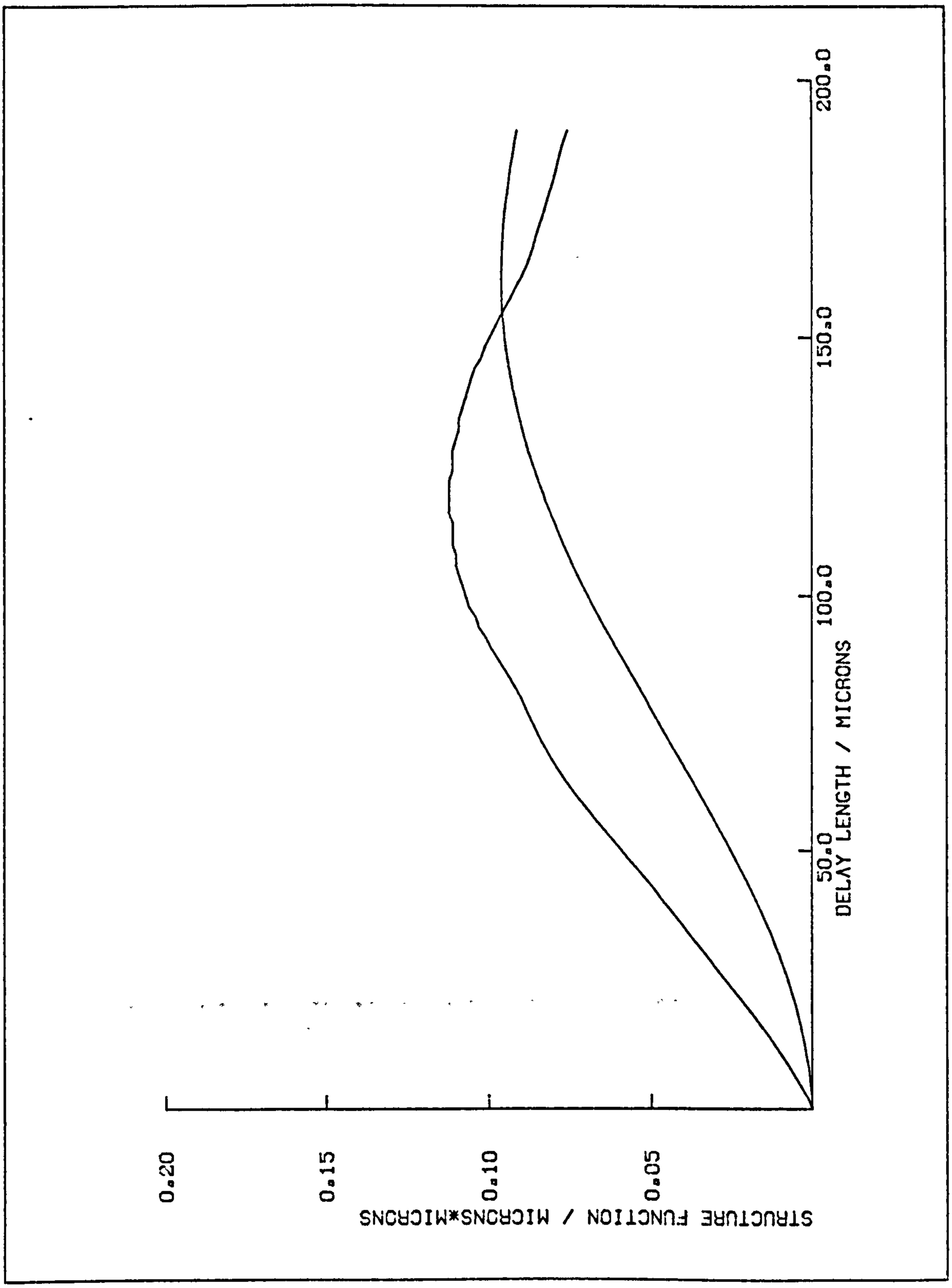


Figure 4.2 Structure Functions of a typical test specimen before and after experiments

In the pure rolling condition increasing the contact pressure from 1.02 GN/m^2 to 1.32 GN/m^2 resulted in a shift of the centre of the reference disc away from the test specimen of some 5 μm in most cases. Similar movements but in the opposite direction occurred if, at constant pressure, the slide-roll ratio was increased from 0.0 to 0.2.

Both these movements were probably due to realignments of the two discs within the hydrostatic bearing housing, rather than movements of the two centres of the discs towards or away from each other.

The output from the surface finish monitor (speckle output) was similar in all the experiments. As the surface finish of the reference specimens became on average 20% smoother, a change of this magnitude from the speckle output would have been expected. However, this could not be detected.

4.4 THE OPTICAL VIBRATION MONITOR RESULTS

4.4.1 Vibration Signals at Constant Pressure

The output signals of the vibration monitor were analysed in the same way as the Talysurf measurements of the specimens. The signals were high-pass filtered at 200μm (approximately the Hertzian contact width) and their root mean square (RMS) roughness, σ computed.

SPECIMEN NO.	RMS ROUGHNESS FOR PURE ROLLING σ(μm)	INCREASE IN σ FOR SLIDE ROLL RATIO 0.048 (%)	INCREASE IN σ FOR SLIDE ROLL RATIO 0.072 (%)
1	0.762	9	25
2	1.209	4	33
3	0.648	9	10
4	1.105	5	34
5	1.278	2	19
6	1.407	10	34
7	1.077	-4	26
8	1.252	3	32
9	1.104	1	19
10	1.243	8	34
11	1.480	1	27
12	1.367	6	42
13	1.067	8	23
14	1.186	10	29
15	1.196	11	29
AVERAGE CHANGE		6	28

Table 4.1 Analysis of vibration signals for a contact pressure of 1.02 GN/m^2, (signals high pass digitally filtered at 200 μm).

Table 4.1 gives the results at a mean contact pressure of 1.02 GN/m^2. The values of σ for a zero slide-roll ratio (pure rolling) and the percentage increases in this value for slide-roll ratios of 0.048 and 0.2 are shown. Note, that each value is the average of two output records and that 200 μm corresponds to a frequency of 3,000 Hz.

An attempt was made to correlate the vibration signals to the surface topographies of the specimens. The results from the experiments with a mean contact pressure of 1.02 GN/m^2 and pure rolling were studied in detail. Figure 4.3 shows the RMS value of the vibration signal plotted against the final RMS roughness of the specimens. The slope of the least squares fit mean line would seem to indicate that the amplitude of the vibration is reduced by increasing the roughness of the specimen surfaces. However, the scatter of the results give this result little meaning.

Figure 4.4 gives the same information as Figure 4.3 but this time the vibration signal is the raw laser signal, i.e. before the static calibration has been added (see section 4.4.1). Notice that the scatter is no greater than in Figure 4.3 and is perhaps less. This suggests that the static calibration is inadequate to compensate for the differing topographies and reflectivities of the reference specimens. A different technique is therefore required to convert the laser output into displacement in microns.

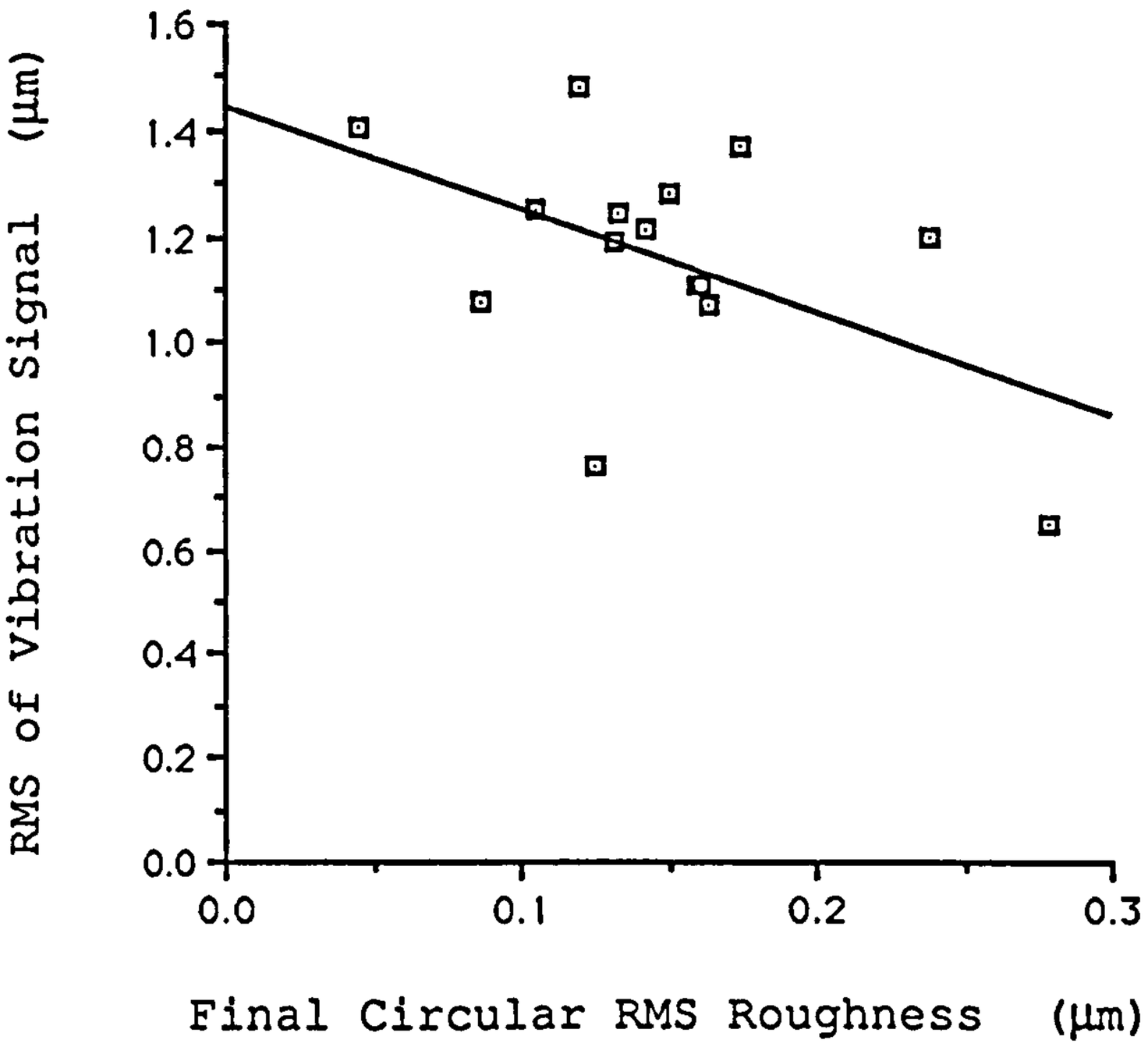


Figure 4.3 Vibration level against specimen roughness

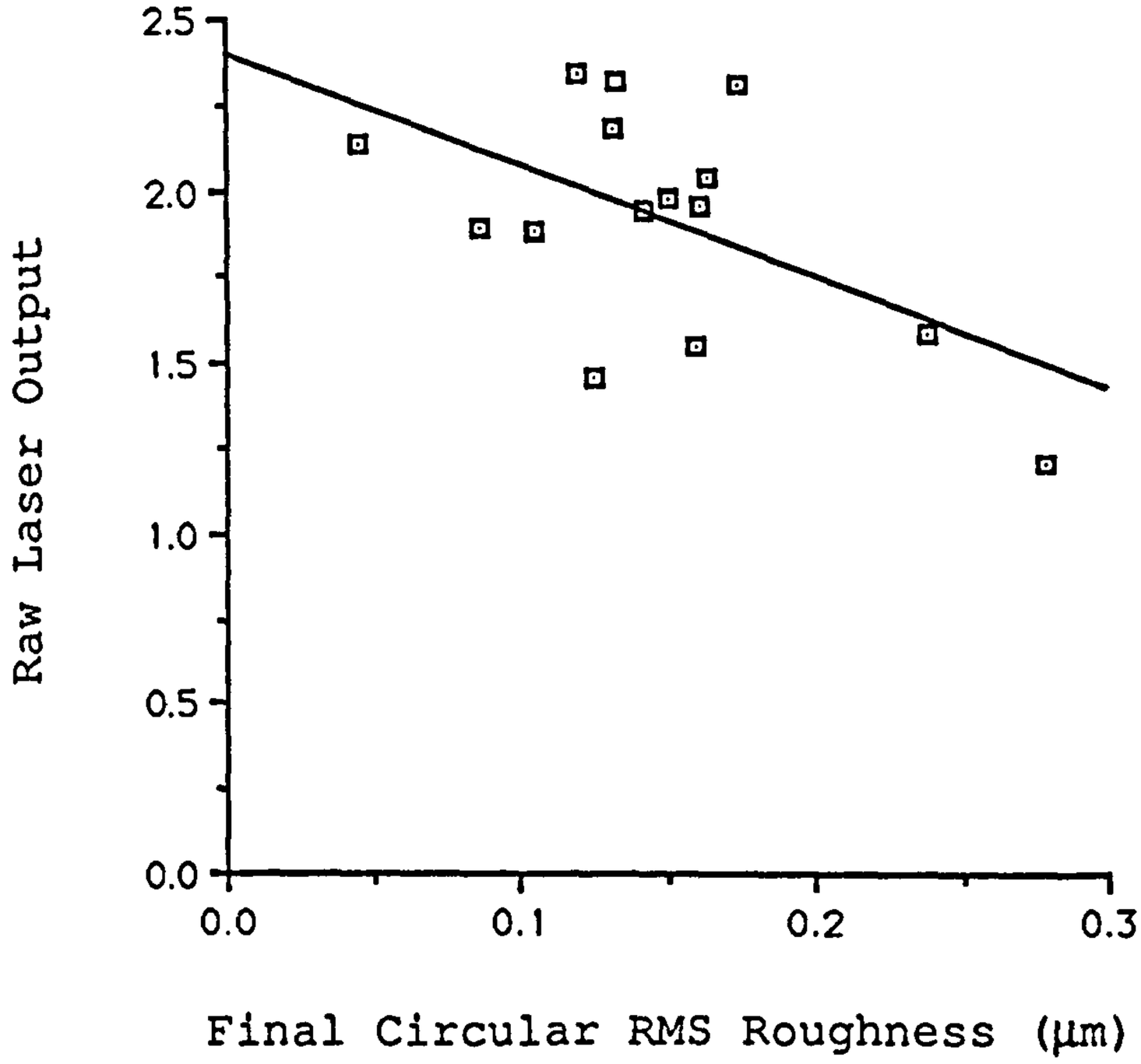


Figure 4.4 Raw laser output against specimen roughness

Figures 4.3 and 4.4 could be interpreted as indicating that the laser device could not detect changes in vibration levels due to different specimen topographies. However, the tests detailed in Chapter 3 using a grooved test specimen proved that the system was capable of detecting vibrations due to surface topography. This would suggest that the vibration level is relatively insensitive to the form of the surface topography, only major features having a significant effect.

A dimensional analysis was carried out to see if a trend could be found from a combination of the known parameters. However, no correlation was detected between any of the parameter combinations attempted. Figure 4.5 shows a typical result using one of the combinations which came from the dimensional analysis, i.e.:-

$$\frac{\text{RMS Vibration}}{\text{RMS Roughness}} = K \times \frac{\text{RMS Roughness}}{\text{Specimen Correlation Length } (\beta^*)} \qquad (4.1)$$

where K is a constant.

As can be seen from the figure no correlation seems to exist. This was the case with all combinations attempted.

The lack of a significant change in the vibration level due to differing test specimen topography may have been in part due to the fact that the reference specimens were not smooth but had comparable roughness to the actual test specimens. The composite roughness of the contacting discs was perhaps not of a sufficient range to give

detectable differences in vibration level with the resolution of the laser device being limited to a movement of 2 μm.

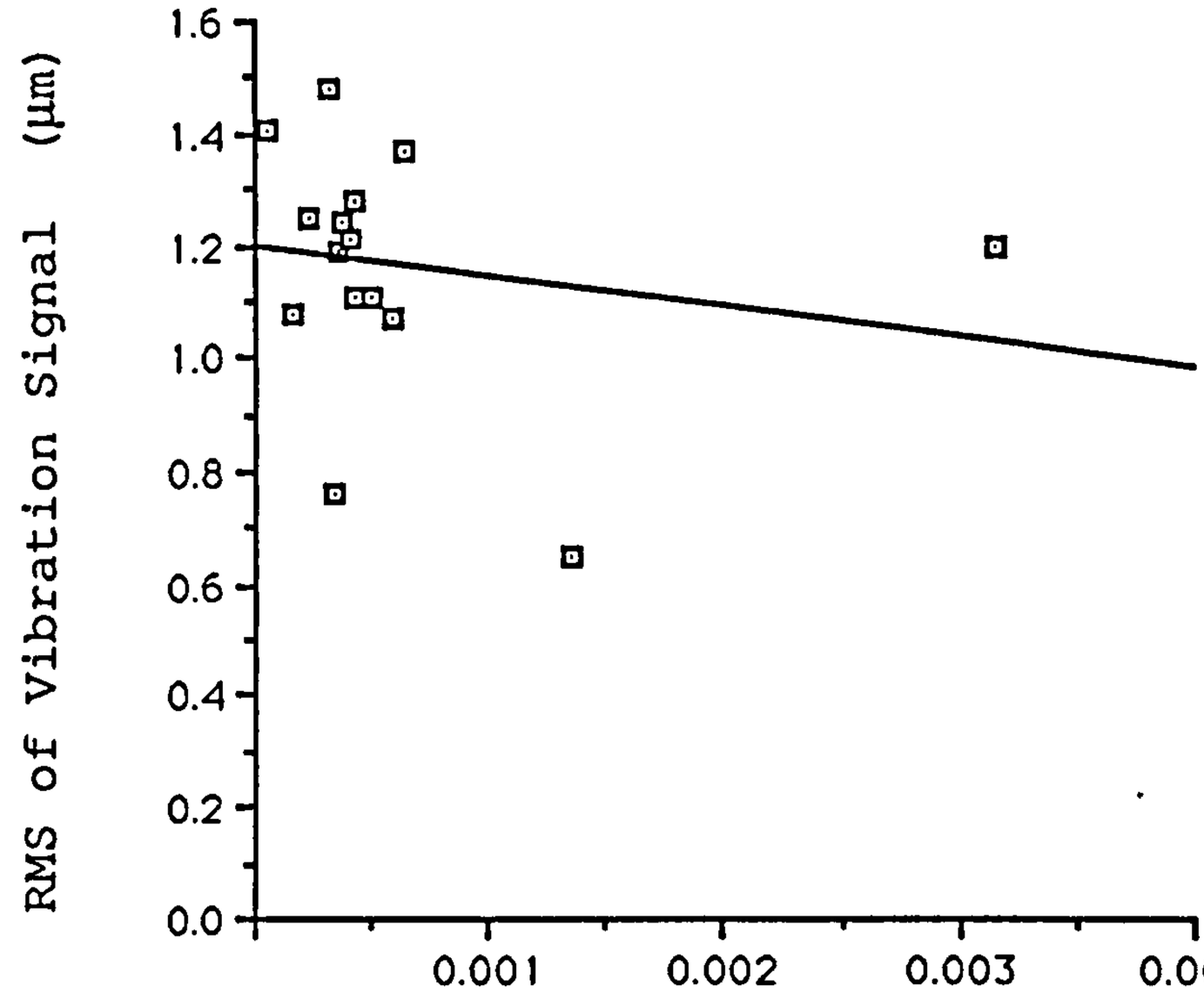


Figure 4.5 Attempt at finding correlation through dimensional analysis

From the results of Table 4.1 an increase in vibration level was detected with increasing slide-roll ratio at a constant contact pressure. Figure 4.6 shows the vibration signal structure functions for a typical test specimen for the pure rolling, 0.048 and 0.2 slide-roll ratio conditions. Note how the general level of vibration increases with increased slide-roll ratio.

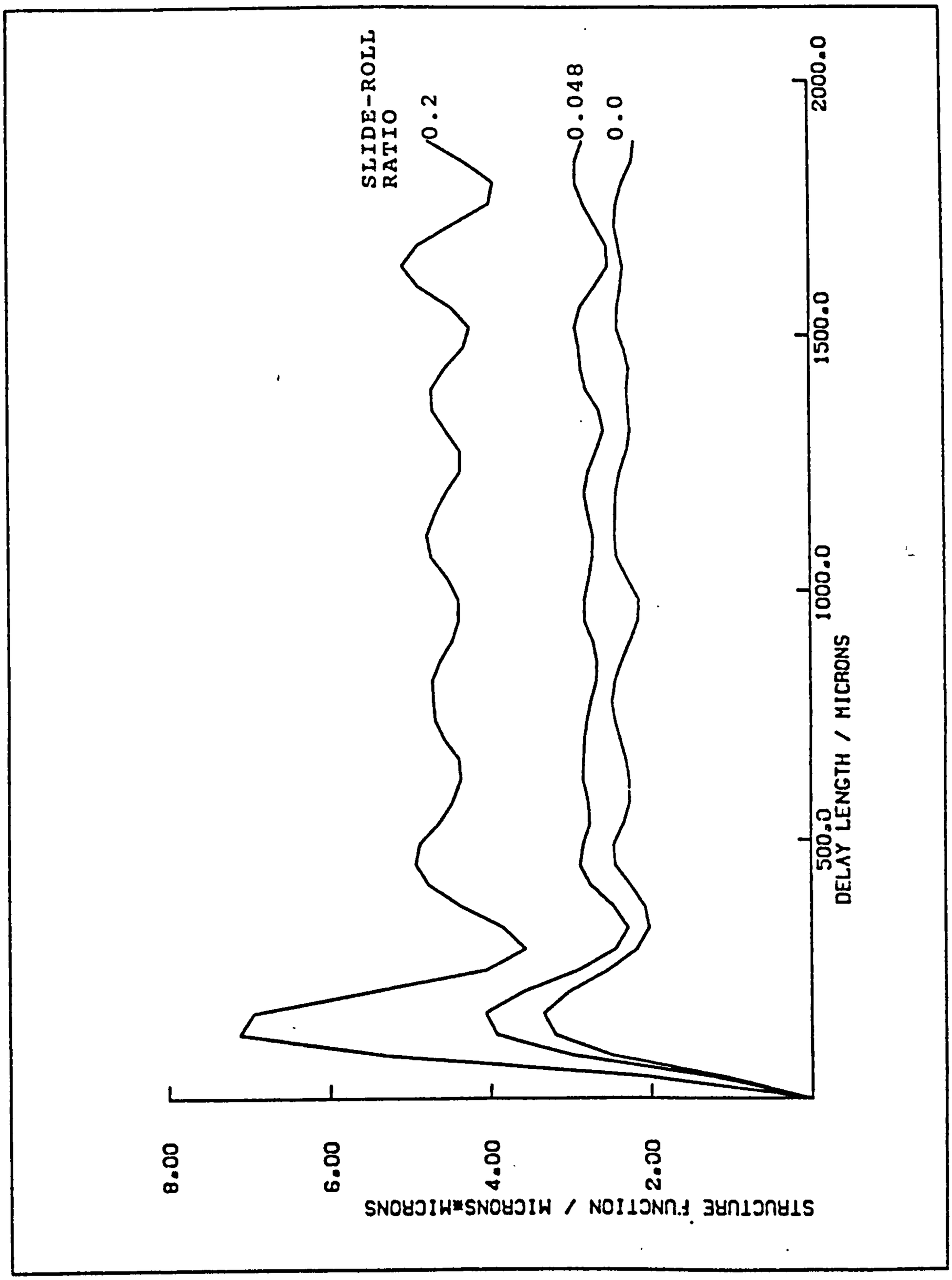


Figure 4.6 Noise structure functions for a typical specimen at differing slide-roll ratios

4.4.2 Vibration Signals at Constant Slide-Roll Ratio

When the slide-roll ratio was kept constant and the pressure increased from 1.02 GN/m^2 to 1.32 GN/m^2 and then to 1.56 GN/m^2 the vibration signal did not alter by more than 4% from its root mean square value for 1.02 GN/m^2. Thus the vibration of the specimens seemed relatively insensitive to contact pressure.

4.5 CONCLUSIONS DRAWN FROM EXPERIMENTAL PROGRAMME

These were as follows:-

1) The commissioning of the mini-disc machine used for the experiments is now complete. This provides a very versatile machine for many forms of test, with its speed controls giving infinitely variable slide-roll ratios to produce a whole range of contact conditions.

2) A prototype optical vibration monitor has been developed. Although it requires more investigation it has shown a high potential as an on line monitoring device of contact conditions. The vibrational level, wear rate and surface finish of the disc specimens can all be continuously monitored.

3) From the results presented, there appears to be little correlation between overall vibration levels and roughness parameters. However, this could simply reflect the complex nature of the asperity interaction

problem, overall roughness parameters being inadequate to quantify the range of asperity shapes and sizes which influence the problem. Attempts at isolating this range are presented, through parameters such as σ and β^*, but more work is required if the ideal gear tooth flank surface topography is to be defined.

4) As an aid to resolving the problems mentioned above, a theoretical model of a profile structure function has been developed and is presented in Chapter 5. This allows roughness characteristics to be studied on an asperity size basis. The model is shown to give promising results on the limited tests performed.

5) Chapters 6 and 7 detail a completely different approach to investigating the contact of rough surfaces through the development of a numerical three-dimensional surface contact model. Chapter 8 shows the application of this technique to the surfaces used in these experiments.

CHAPTER 5

SURFACE TOPOGRAPHY CHANGES DURING LUBRICATED WEAR

5.1 INTRODUCTION

From the experimental programme it appears that overall roughness parameters are unable to quantify the range of asperity shapes and sizes influencing complex asperity interactions occurring during the wear process. A method whereby roughness characteristics can be studied on an asperity size basis is required. The following sections describe the development of a theoretical model of the structure function of a surface profile. This can be used to identify features of particular sizes that may be causing damage during the wear process.

5.2 BACKGROUND

Results from investigations into contact fatigue and wear indicate that surface roughness can have an important effect (Li *et al.*, 1976; Ishibashi and Matsumoto, 1972), with the initial topography tending to dominate the fatigue life. However, during the wear process individual surface features undergo dramatic changes. Thus it would seem that these features, in undergoing changes, damage the

component by setting up residual stress fields (Olver, 1983; Olver *et al.*, 1985; Olver *et al.*, 1986). If these features could be avoided, wear rates would be slowed down and fatigue lives lengthened.

The importance of plastic deformation on an asperity scale has recently been highlighted by Olver in terms of its influence on micropitting (Olver, 1983; Olver *et al.*, 1986) and also on wear (Olver *et al.*, 1985) in rolling, and rolling and sliding contacts. The formation of inclined microcracks have been shown to arise from residual tensile stresses created as a result of local asperity plastic deformation. Olver has shown how such cracks can lead to micropitting, a rapid rate of wear phenomenon, and also to be responsible for an extrusive form of wear, where material is extruded from the cracks by subsequent surface contacts.

The magnitude of the residual stress fields created by these mechanisms depend upon the size of the plastically deformed zone, i.e. it depends upon the maximum size of asperity that will deform plastically. Thus, as with many other surface related problems, there seems to be a critical size of asperity which has a dominating influence on failure related problems.

The physical explanation for the existence of a critical range of asperity sizes, within a plastically related failure mechanism, is relatively straightforward. A surface can be said to be a broad band of wavelengths, or spatial sizes. Considering the extremes of this range, the shortest wavelength asperities are sharp and therefore

easily deformed, whereas the spatially larger asperities, though of greater height, do not have slopes or curvatures likely to induce plastic deformation.

Thus considering the range of spatial sizes available it is evident that the larger asperities deform elastically, but as their size reduces spatially, a size is reached where the slopes and curvatures are great enough to allow plastic deformation. It is this size range, i.e. the maximum size range that can create plastic deformation, that appears to be the most damaging as it creates the maximum plastically deformed zone and consequently maximises the residual tensile stress effects.

Identifying this critical size of asperity is not straightforward, but if it can be achieved in a reliable way, it may be possible to modify machining processes to reduce, or even remove, the magnitude of such effects.

To isolate these damaging features the topographical changes during contact have to be investigated. Recent work by Sugimura and Kimura (1984), using two types of cross-correlation coefficient, did not identify any particular size of damaging feature. This may have been due to the fact that the correlation coefficients are not independent of the mean plane of the surface from which they are measured and obviously during the wear process the mean plane changes. A similar function, that does not suffer from this drawback, is the structure function.

5.3 THE STRUCTURE FUNCTION

The use of the structure, or variance function, (SF) as a means of quantifying variations in surface texture has been proposed by Sayles and Thomas (1977). This is defined for a profile as:-

$$S(\tau) = E\left\{[z(x) - z(x+\tau)]^2\right\} \qquad (5.1)$$

$z(x)$ is height at position x, $z(x+\tau)$ is height at a position τ later, and E{} denotes expectation. Hence, the function represents the mean square of difference in height expected over any spatial distance τ.

In an integral form, equation (5.1) becomes:-

$$S(\tau) = \lim_{L\to\infty} \frac{1}{L} \int_0^L [z(x) - z(x+\tau)]^2 dx \qquad (5.2)$$

This is an idealised solution over an infinitely long profile length L. In practice the profile is of a finite length, and the full integration cannot be performed as $z(L+\tau)$ would be required for x=L. This is generally overcome by modifying equation (5.2) to become:-

$$S(\tau) = \frac{1}{(L-\tau)} \int_0^{(L-\tau)} [z(x) - z(x+\tau)]^2 dx \qquad (5.3)$$

Equation (5.3) has the disadvantage that the length over which the integration is performed is not constant. This can be overcome by replacing equation (5.3) with:-

$$S(\tau) = \frac{1}{(L - \tau_{max})} \int_{0}^{(L-\tau_{max})} [z(x) - z(x+\tau)]^2 dx \qquad (5.4)$$

In equation (5.4) τ_{max} represents the maximum value of τ calculated. To limit the errors from the ideal solution τ_{max} is limited to L/10. The following sections will use equation (5.4) as the definition of the structure function.

5.4 STRUCTURE FUNCTION CHANGES DURING WEAR

Figure 5.1 shows the structure function representation of worn and unworn profiles. The dramatic change in the function due to the wear process should obviously reveal information about the effects of wear on the surface. However, apart from the RMS roughness σ (the structure function asymptotes to $2\sigma^2$ for stationary data (Sayles and Thomas, 1977), the precise relationship between surface features and the resultant structure function is complicated. This makes prediction of changes in surface features due to wear processes, from the change in profile structure functions, difficult to accomplish. In an attempt to do this a theoretical model of the structure function has been constructed.

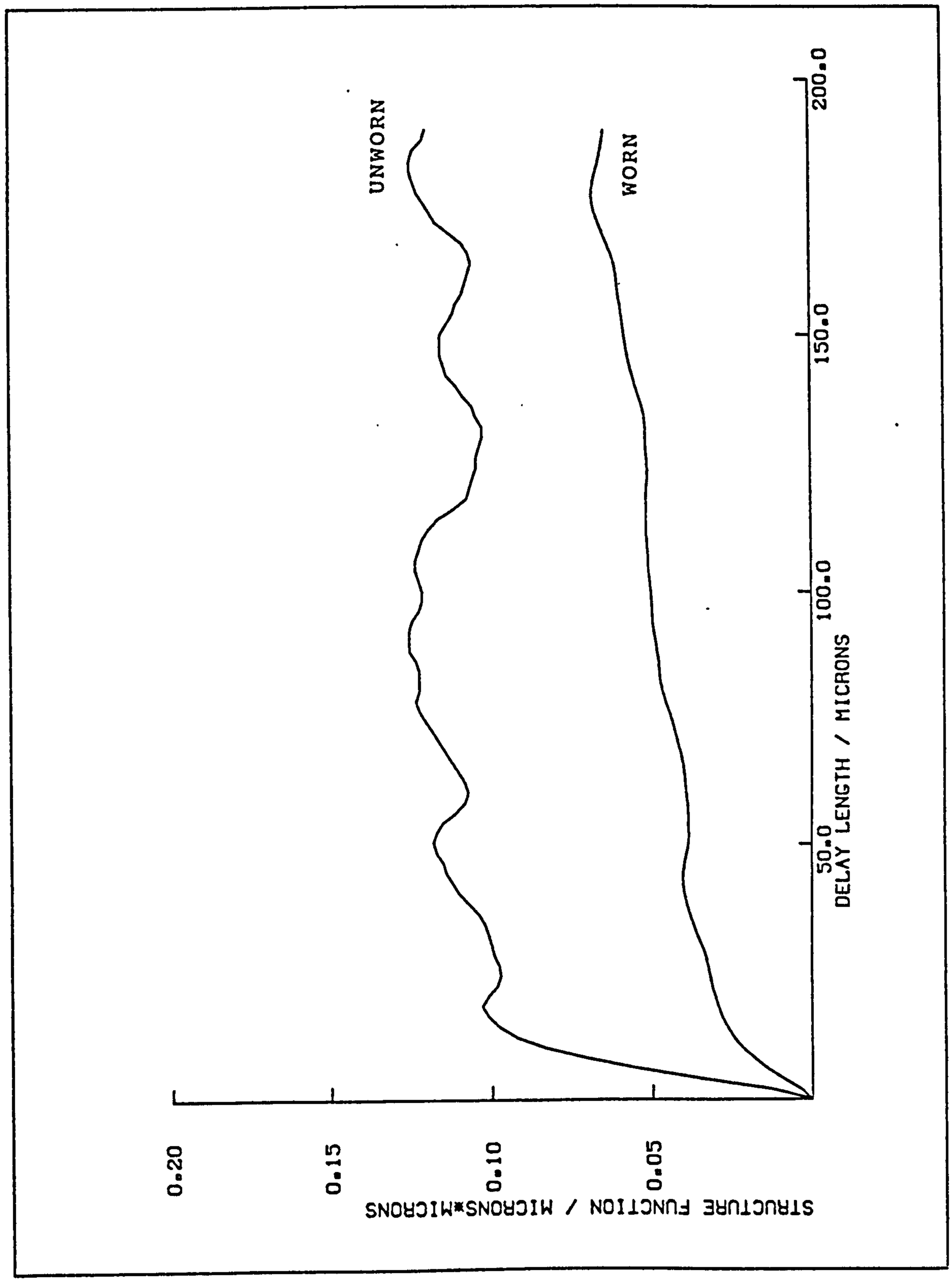


Figure 5.1 Structure functions of worn and unworn bearing surfaces.

5.5 THE STRUCTURE FUNCTION MODEL

5.5.1 Basis of the Structure Function Model

The model structure function described is based on the principle that a surface can be modelled by a series of triangular features or asperities of varying base angle, spatial size and position. The general asperity is shown in figure 5.2. The variables, θ and a, describe the shape of an individual asperity which can occur anywhere along a profile and be at any angle to the horizontal. Note, that θ can be positive or negative.

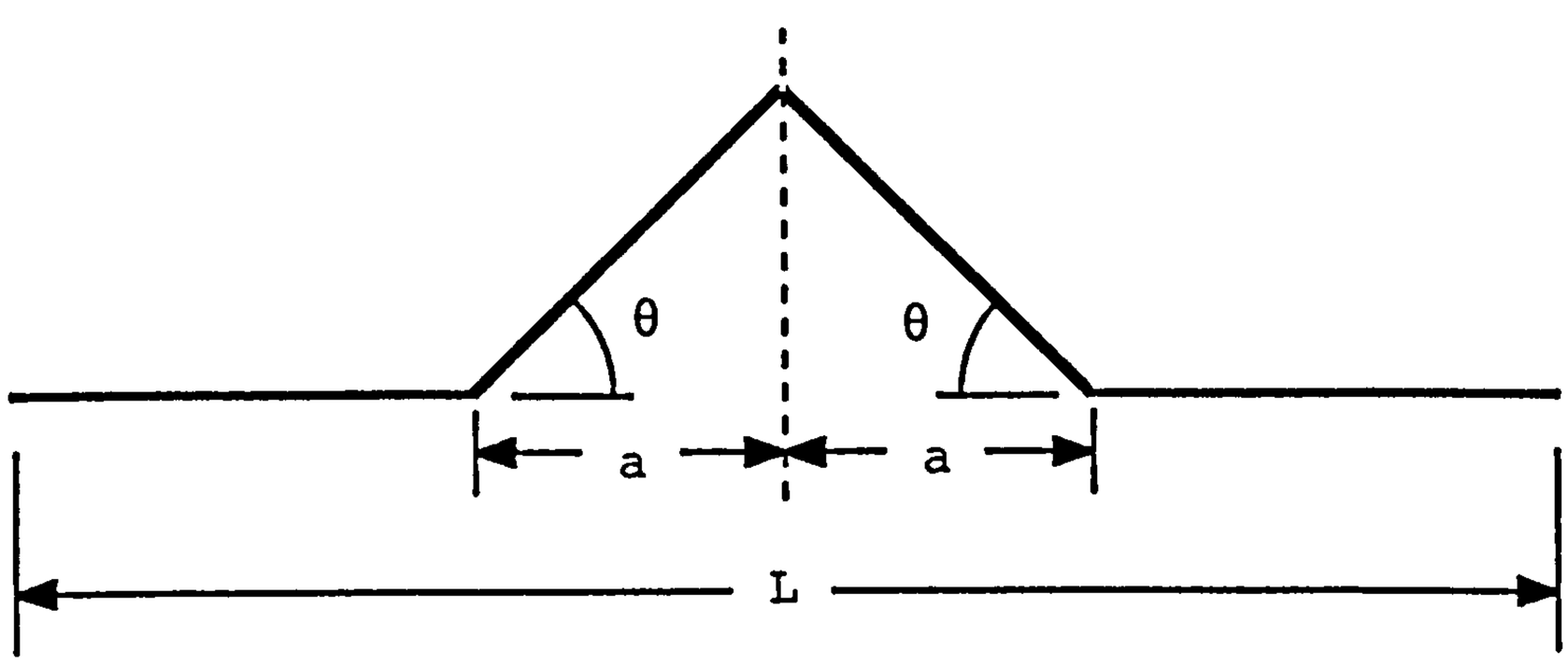


Figure 5.2 The general asperity in a profile.

5.5.2 Structure Function of General Asperity

To calculate the structure function for the general asperity of Figure 5.2, the following conditional

expectation relationship was employed (Bendat, 1958):-

$$E\left\{\begin{matrix}\text{some}\\\text{event}\end{matrix}\right\} = \sum_{\substack{\text{ALL STATES}\\\text{OF EVENT}}} E\left\{\begin{matrix}\text{PARTICULAR}\\\text{STATES OF}\\\text{EVENT}\end{matrix}\right\} \times \begin{matrix}\text{PROBABILITY}\\\text{OF BEING IN}\\\text{THE PARTICULAR}\\\text{STATE}\end{matrix} \qquad (5.5)$$

Combining equation (5.5) with equations (5.1) and (5.4) enables the structure function to be calculated for any value of τ, (see Appendix A):-

$$S(\tau) = \frac{\theta^2 4a^3}{3(L-\tau_{max})} \qquad \text{for } \tau > 2a \qquad (5.6a)$$

$$S(\tau) = \frac{\theta^2\left[4a^3-(2a-\tau)^3\right]}{3(L-\tau_{max})} \qquad \text{for } 2a > \tau > a \qquad (5.6b)$$

$$S(\tau) = \frac{\theta^2\tau^2(2a-\tau)}{(L-\tau_{max})} \qquad \text{for } \tau < a \qquad (5.6c)$$

Equations (5.6a), (5.6b) and (5.6c) completely define the structure function for a single asperity. Note that from equation (5.6a) the structure function becomes a constant if τ is greater than 2a. The data is then stationary and this value is equal to $2\sigma^2$ for the profile, as predicted by Sayles and Thomas (1977).

5.5.3 The Complete Structure Function Model

Extending the model to include two or more asperities in the profile complicates the mathematics. However, for the case of two asperities a general pattern seems to

emerge whatever the relative positions and sizes of asperities as follows:-

$$S(\tau) = S(\tau) \text{ for first asperity alone} + S(\tau) \text{ for second asperity alone} + \text{correction factor} \qquad (5.7)$$

$$\text{where correction factor} = \frac{2}{(L-\tau_{max})} \sum_{i}^{\text{No. of Asperities}} a_i^2 \theta_i (h_{Ai} + h_{Bi}) \qquad (5.8)$$

Here, h_{Ai} and h_{Bi} are the heights of the centre of the base of asperity i measured relative to the positions of the underlying τ steps, as shown in Figure 5.3. Equation (5.7) suggests that to model the structure function of a profile of several asperities, the exact location of each asperity has to be known so that the correlation factor can be calculated. However, if some assumptions are made about the distribution of the asperities then this problem can be overcome. Notice that the correction factor depends on θ and not θ^2. Therefore, if θ forms a symmetrical distribution with a zero mean, then the overall correction factor is zero. This implies that for a random distribution of asperities the correction factor is zero and the structure function is simply the sum of the structure functions for each individual asperity alone.

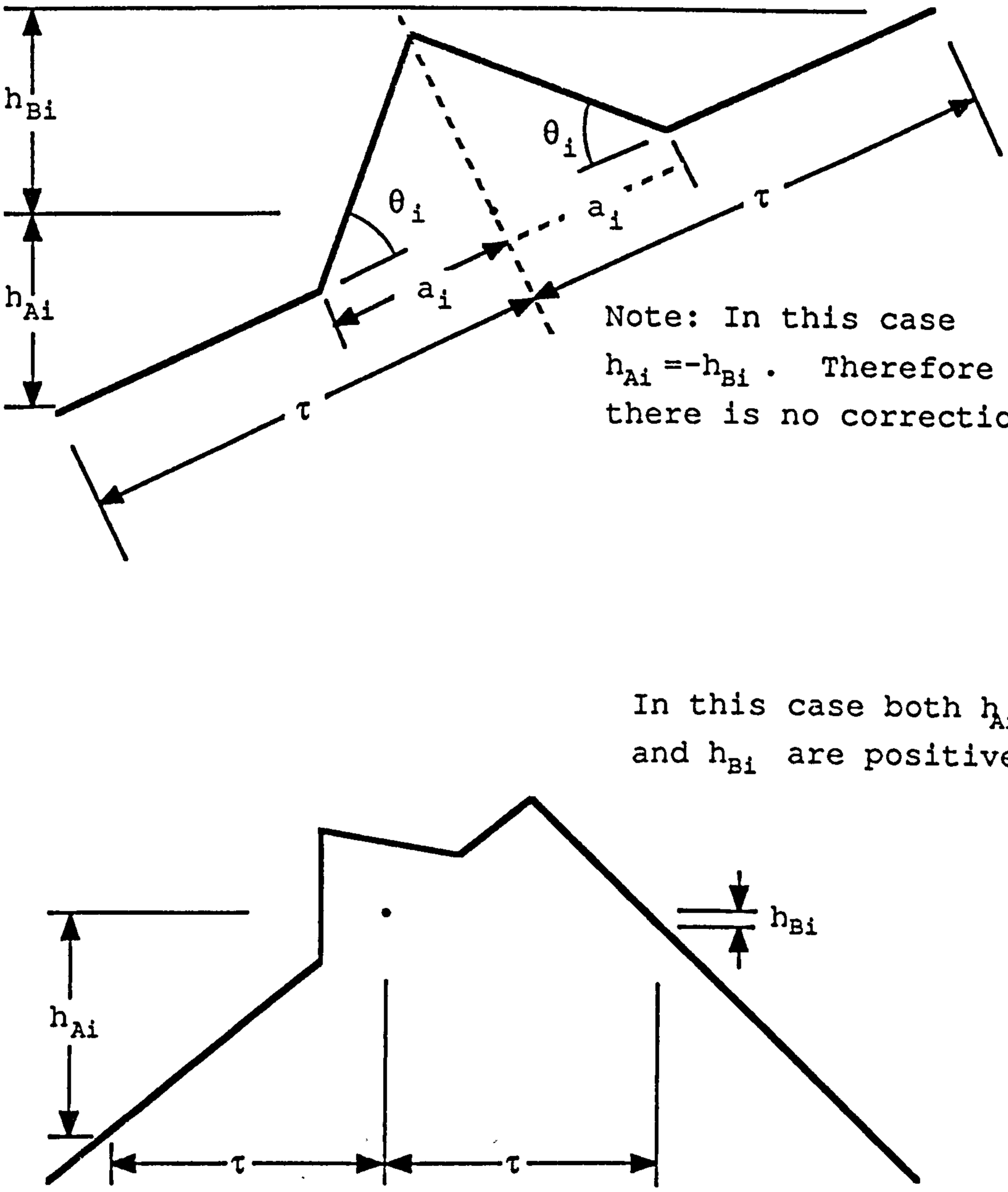


Figure 5.3 Two examples of the formation of the correction factor

If the profile is formed from asperities of differing half base lengths a, then depending on the relationship of a to τ, the contribution to the structure function of any particular asperity will be found from one of equations

(5.6a), (5.6b) or (5.6c). The contribution of all asperities with a particular half base length a can be calculated by replacing θ^2 in equations (5.6a), (5.6b) and (5.6c) with $N_a\theta^2{}_a$ where N_a is the number of asperities of half base length a and $\theta^2{}_a$ is the mean square base angle of these asperities. The complete structure function for any particular value of τ can then be obtained by summing the modified equations (5.6a), (5.6b) and (5.6c) over the limits for a, giving:-

$$S(\tau) = \frac{1}{3(L-\tau_{max})}\sum_{a=0}^{\frac{\tau}{2}} N_a\theta^2{}_a a^3 + \frac{1}{3(L-\tau_{max})}\sum_{a=\frac{\tau}{2}}^{\tau} N_a\theta^2{}_a\left[4a^3-(2a-\tau)^3\right]$$

$$+\frac{\tau^2}{(L-\tau_{max})}\sum_{a=\tau}^{a_{max}} N_a\theta^2{}_a(2a-\tau)$$

which can be rewritten as:-

$$S(\tau) = \frac{1}{(L-\tau_{max})}\left[\frac{4}{3}\sum_{a=0}^{\frac{\tau}{2}} N_a\theta^2{}_a a^3 - \frac{1}{3}\sum_{a=\frac{\tau}{2}}^{\tau} N_a\theta^2{}_a(2a-\tau)^3 + \tau^2\sum_{a=\tau}^{a_{max}} N_a\theta^2{}_a(2a-\tau)\right] \qquad (5.9)$$

Equation (5.9) is the complete modelled structure function in its most general form. To solve it we need to know the distribution of $N_a\theta^2{}_a$.

5.6 TESTS ON THE MODELLED STRUCTURE FUNCTION

One method for estimating the distribution of $N_a\theta^2{}_a$ is

to use fourier analysis. This assumes that a surface profile can be constructed from a series of sine and cosine waves. If these waves are converted into saw-tooth waves of the same power, then the waves can represent a series of asperities joined together. From this the number and base angle of each particular half-base length of asperity can be evaluated, thus giving an estimation of $N_a\theta^2_a$.

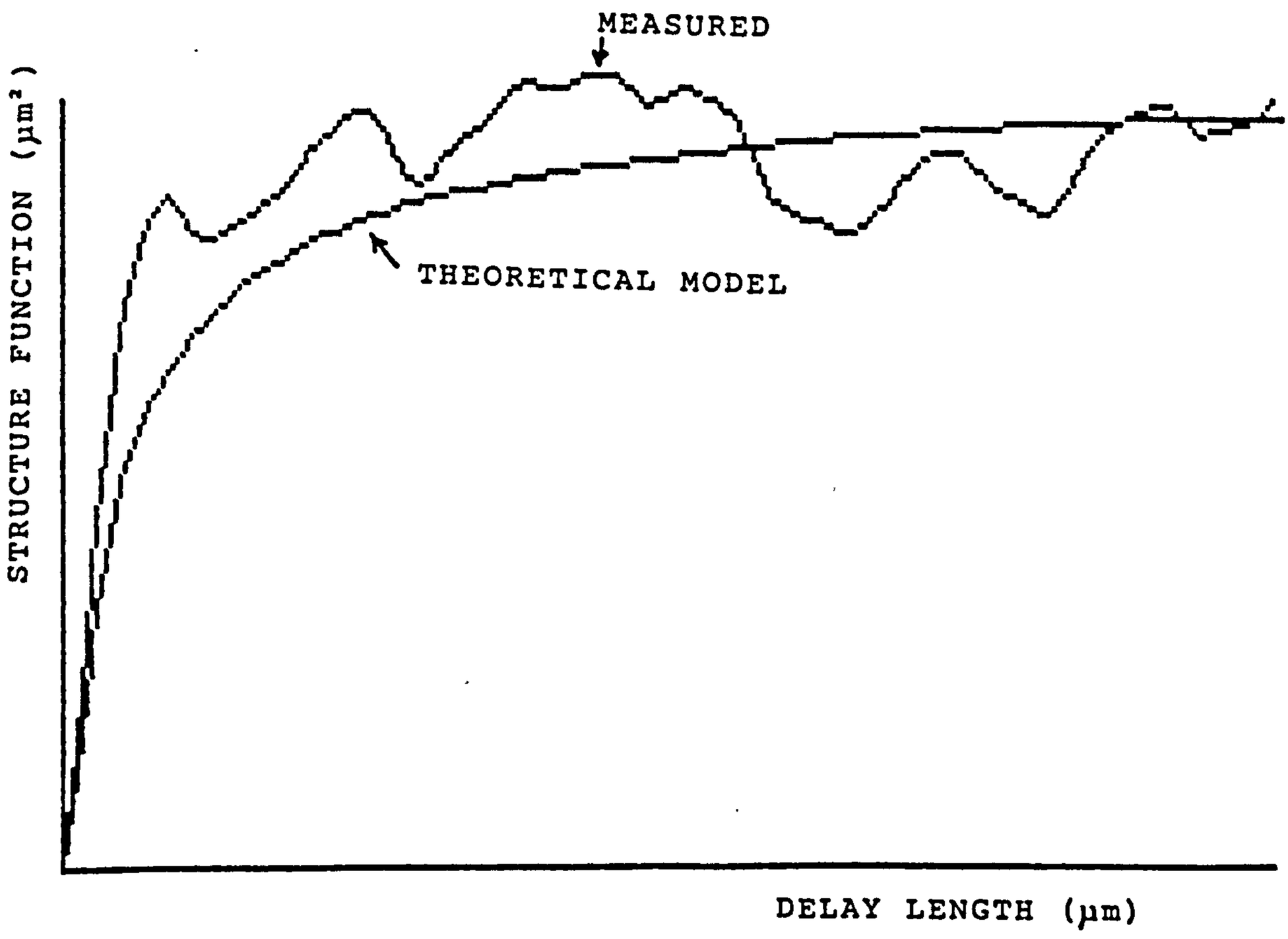


FIGURE 5.4 Comparison of measured and modelled structure functions of a bearing profile

Figure 5.4 shows the structure function of a shot-blasted surface profile calculated in the normal way. Superimposed on this is the structure function for the same profile, found by using fourier analysis and the modelled structure function. As can be seen the two structure functions are very similar. The modelled structure function can therefore be used to investigate the changing of surface features during wear processes.

Similar tests using fractal analysis (Mandlebrot, 1983) to estimate the distribution of $N_a\theta^2_a$ also showed that the modelled structure function was capable of producing the correct structure function. However, if the modelled structure function is to be used to give a greater insight into the surface topography changes during wear processes other methods of estimating the distribution of $N_a\theta^2_a$ are required. Otherwise a direct use of fourier analysis could be used to give the same information. Perhaps the difference of a direct calculation and the modelled calculation of the structure function using fourier analysis could be used to predict the correlation factor and hence the actual position of asperities in the profile.

Note that the values of the modelled structure function for differing distributions of $N_a\theta^2_a$ and Figure 5.4 were produced using a computer program written for a BBC model B micro-computer using the BASIC language.

CHAPTER 6

THREE-DIMENSIONAL ROUGH SURFACE CONTACT

6.1 INTRODUCTION

The traditional statistical models of rough surface contact using assumed shapes and distributions for the asperities, such as those by Greenwood and Williamson (1966) or Nayak (1971), have the disadvantage that they do not give the actual deformed geometries of the bodies, especially away from the contact zones. They also tend to become inaccurate for significant levels of penetration between the surfaces.

Another approach to the problem of surface contact analysis is the use of numerical methods. Until the advent of modern computing facilities, these methods were impractical as the complex algorithms required for their solution could not be solved. However, now that these facilities are available much work is being carried out in this field as they allow the contact stresses and deformed geometries to be defined exactly.

The availability of powerful micro computers has led to the development of surface measurement and analysis systems where real measured data can be used as input for numerical

contact analysis. One such three-dimensional mapping system (Webster *et al.*, 1986 - reproduced as Appendix B) uses a Talysurf stylus instrument to record a matrix of surface heights from flat or curved surfaces. Two-dimensional plane strain contact analysis can then be carried out on selected profiles (Webster, 1986; Webster and Sayles, 1986). Chapter 7 describes work carried out to develop a similar full three-dimensional contact model, including the effects of plasticity, whilst this chapter is a review of the different approaches which have been used for rough surface contact analysis.

6.2 THEORETICAL ROUGH SURFACE CONTACT ANALYSIS

6.2.1 Elastic Contact

The first satisfactory analysis of the stresses at the contact of two elastic solids is due to Hertz (1896), which is reviewed by Johnson (1982). However, this and later analyses suffer from the need to confine the original undefined shapes of the contacting bodies to known polynomial functions.

Attempts to model the contact of real rough surfaces have tended to use extensions of these techniques. The measured surface topography is reduced to its probability density functions and their moments, which are then used to model the real surface by a population of idealised asperity shapes, for which the contact behaviour is well known.

One of the earliest such theoretical models was proposed by Archard (1957) in a study of friction. Until this time it was generally assumed that the linearity between applied load and contact area in Amonton's laws of friction could only be satisfied by plastic contact (see Section 6.2.2). The idea that the linearity could be explained by elastic contact was first proposed by Lincoln (1953) and Archard (1953).

In his contact model Archard originally assumed that the surface roughness could be represented by spherical protuberances of constant radius of curvature. The relationship between the contact area and the applied load for a smooth sphere loaded against a flat surface being:-

$$A \alpha F^{2/3} \qquad (6.1)$$

where A is the contact area and F the applied load.

This result does not give the proportionality between contact area and normal force required to satisfy Amonton's laws of friction. Archard therefore extended his model to consider the effect of spherically capped asperities of smaller radius placed on top of the original asperities. Continuing this process with further levels of smaller radius spherical caps brings the proportionality between contact area and applied load nearer to unity. For a surface consisting of three scales of spherically capped asperities the relationship becomes:-

$$A \alpha F^{26/27} \qquad (6.2)$$

Similarly the proportionality between the number of individual contact spots and the applied load becomes more linear as the number of scales of asperity are increased. For a smooth sphere the number of contact spots is constant at one, but for three scales of spherically capped asperities the relationship between the number of contact spots, N and the applied load, F is:-

$$N \alpha F^{8/9} \tag{6.3}$$

Archard concluded that Amonton's laws of friction (that the frictional force is proportional to load and hence that the true area of contact is proportional to the applied load) could be satisfied by a rough surface, if the surface is modelled by superimposed layers of asperities of successively smaller scales. The final layer of micro-asperities support the load and as the scale of the asperities decreases their individual contact area becomes independent of load and the number of individual contact spots becomes proportional to load.

Although this approach of using spherically capped asperities is unrealistically simplifying the problem, it does demonstrate the important concept of roughness existing over a large spatial size range. This must be taken into account in any study of the true nature of surface contact (Mandlebrot, 1977).

The first statistical analysis of rough surface contact was presented by Greenwood and Williamson (1966). They again assumed the surface to consist of spherically capped

asperities of constant radius, but the height of the tips of the asperities varied randomly according to an assumed distribution. From real measurements it had been found that many rough surfaces closely followed a Gaussian height distribution. The exponential height distribution is a good approximation to this over the upper 25% of the surface and so expressions relating applied load, area of contact and number of contact spots were found using this distribution. Again direct proportionality was obtained between applied load and contact area and between applied load and number of contact spots. Furthermore the average contact size was found to be nearly independent of load.

From this model Greenwood and Williamson went on to look at the important consideration of plastic contact. They developed the so called plasticity index, ψ defined as:-

$$\psi = \frac{E^*}{H} \left(\frac{\sigma_s}{R_s}\right)^{\frac{1}{2}} \qquad (6.4)$$

The plasticity index indicates the proportion of the contact which will undergo plastic deformation. A value of ψ greater than 1 giving a greater than 2% proportion of plastic contact and $\psi<0.6$ indicating less than 2% is plastic. Significantly, the applied load was shown to have little influence on whether the contact is elastic or plastic. However, the exact significance of the plasticity index remains a subject for debate.

The Greenwood and Williamson model was extended by Greenwood and Tripp (1967) to consider the effect of concentrated contacts by looking at the contact of rough spheres. They found that the load was proportional to the contact area for low loads, but as the load was increased the contact conditions become increasingly like a smooth Hertzian contact. This work was further extended by Lo (1969) to consider the line contact of rough cylinders.

The original Greenwood and Williamson model was also extended by Greenwood and Tripp (1970) to account for both surfaces being rough, taking into account misalignment of asperities. The results were qualitatively similar to those of the simpler model, and it was shown that the contact of rough surfaces can be simulated by the contact of a single equivalent rough surface with a plane.

In an attempt to make the asperity models more realistic, models based on profile measurements were developed. These use the bearing area curve (developed by Abbott and Firestone (1933)), giving the proportion of a surface below any particular height, to define the contact area as a function of surface separation. Profilometric theories provide information on load-compliance laws and both surfaces can be rough. Their disadvantage is that simple contact stress laws have to be used. A discussion of the relative merits of contact theories based on asperity models and profile measurements has been presented by Greenwood (1967).

To link the approaches of the asperity and

profilometric methods, Whitehouse and Archard (1970) proposed a method using parameters gathered from measured surfaces. They assumed that a surface could be described by a height and a spatial parameter, the height parameter being the root mean square roughness, σ and the spatial parameter being the correlation length, β^* defined as the distance for the autocorrelation function to drop to a value of 1/e (see Chapter 2 Section 2.3.1.1).

Comparing their results with those of Greenwood and Williamson, an expression was found for the plasticity index in profile measurement terms (after a correction by Onions and Archard (1973)):-

$$\psi = 0.69 \frac{E^*}{H} \frac{\sigma}{\beta^*} \qquad (6.5)$$

Nayak (1971) proposed a model of surface roughness based on the theory of statistical geometry developed by Longuet-Higgins (1957*a* and 1957*b*) in a study of ocean waves. The validity of this theory is dependent on one assumption, namely that the surface heights, slopes and curvatures possess a multi-Gaussian probability density. It is a standard result in random-process theory that a linear combination of Gaussian processes is itself a Gaussian process. Thus, since surface slopes are essentially the difference between two heights, the surface slopes will possess a Gaussian distribution. Using a similar argument, the second derivatives (curvatures) of the surface will also be of Gaussian distribution.

Nayak (1973) was able to show that the statistics of a single profile are markedly different from those of a surface but that it is possible to calculate the parameters necessary for the analysis of the surface from single profiles.

Another random-process approach to the elastic contact of rough surfaces has been presented by Bush *et al.* (1975). They assumed that the rough surface consists of a mean plane with hills and valleys randomly distributed on it. The summits of the hills were approximated by paraboloids, the distributions of heights and principal curvatures of which were obtained from the random-process model. At large separations, their model achieved an exact proportionality between applied load and contact area and indicated that the contact area for any given separation of the surfaces is half the bearing area at that separation.

The equation relating the applied load and contact area for an isotropic surface was found to be:-

$$\frac{F}{A} = \sqrt{\frac{m_2}{\pi}}\, E^* \qquad (6.6)$$

where m_2 is the second moment of the power spectral density function (effectively the mean square slope of the surface).

Bush *et al.* (1979) extended the analysis to cover the case of strongly anisotropic materials. The circular paraboloid hill summits being replaced by elliptic

paraboloids (ie. the asperities have horizontal sections which are ellipses and vertical sections which are paraboloids). The results from this analysis were similar to those gained from the isotropic case but the applied load against contact area equation required extra terms involving elliptic integrals.

Most recent work on the elastic contact analysis of rough surfaces has tended to extend the random-process approach (Whitehouse and Phillips, 1978 and 1982). However, the added complexities introduced in order to relax some of the restrictive assumptions have tended to make the theory difficult to apply in real situations.

To combat this Greenwood (1984) has presented work attempting to simplify the theory so that real data can be directly compared with theoretical results. He concluded that, in non-dimensional form, the properties of the peaks and summits of a rough surface depend only to a minor degree on the parameters describing the surface or on the sampling interval used for the data. However, the absolute values are strongly dependent on the sampling interval.

6.2.2 Plastic Contact

Some of the first studies of surface contact, taking into account the fact that surfaces are rarely perfectly smooth on a microscopic scale, were carried out by Bowden and Tabor (1954 and 1964) in connection with friction, and by Holm (1967) in connection with electrical contact

resistance.

They assumed that the solids are in contact only at isolated points and that asperities are crushed into the plastic region. Plasticity theory was considered justified because the local strains are very high and the theory predicts that the real contact area is proportional to the applied load.

6.2.2.1 Onset of plastic yield

The load at which plastic yield begins in the complex stress field of two solids in contact is related to the yield point of the softer material in a simple tension or shear test through an appropriate yield criterion. The yield of most ductile materials is usually taken to be governed either by von Mises' shear strain-energy criterion:-

$$\frac{1}{6}\left\{ (\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2 \right\} = k^2 = \frac{Y^2}{3} \qquad (6.7)$$

or by Tresca's maximum shear stress criterion:-

$$\max\left\{ |\sigma_1-\sigma_2|,\ |\sigma_2-\sigma_3|,\ |\sigma_3-\sigma_1| \right\} = 2k = Y \qquad (6.8)$$

in which σ_1, σ_2 and σ_3 are the principal stresses in the state of complex stress, and k and Y denote the values of the yield stress of the material in simple shear and simple tension (or compression) respectively (Johnson, 1985).

Refined experiments on metal specimens, carefully controlled so as to be isotropic, support the von Mises criterion of yield. However, the difference in the predictions of the two criteria is not large and is hardly significant when the variance in the values of k or Y and the lack of isotropy of most materials are taken into account. It is justifiable, therefore, to employ the Tresca criterion where its algebraic simplicity makes it easier to use.

The simplest assumption about the onset of plastic yield is that it occurs at a pressure equal to the indentation hardness, H of the softer material. The value of H is normally taken to be approximately equal to 3Y. A number of researchers have investigated in detail the plastic deformation of a single asperity (Ling, 1958*a*; Huffington, 1962; Thomas *et al.*, 1971; Uppal and Probert, 1972*a* and 1975) and it has been noted that the pressure at which yield occurs may differ considerably from the indentation hardness. This has been explained in terms of slip-line fields (Edwards and Halling, 1968; Uppal and Probert, 1972*b*).

Obviously as the proportion of the true area of contact increases relative to the apparent area, the likelihood of interaction between neighbouring asperities increases, thus affecting the onset of plastic yield.

6.2.2.2 Plastic contact surface models

It is not easy to apply slip-line fields to a three-dimensional solid, and a two-dimensional surface model is generally used consisting of an assembly of parallel ridges triangular in cross-section (Ling, 1958*b*; Mitchell and Rowe, 1969; Bay *et al.*, 1975). Three-dimensional plastic asperity models are limited to cones (Kragelsky and Demkin, 1960; Thomas *et al.*, 1971; Uppal and Probert, 1973) and pyramids (Hisakado, 1969; Childs, 1973).

A number of researchers have considered the effect of applied load on a surface covered with plastically yielding cones with their heights normally distributed (Ling, 1958*a*; Tsukizoe and Hisakado, 1968; Hisakado, 1975). They have shown that the real area of contact is proportional to the applied load, that the number of contacts increases at a rate of approximately the 0.9 power of the load and that the average contact spot size increases at a rate of about the 0.1 power of the load.

That is, the average contact spot size is almost independent of the load and the increase in the real contact area is almost entirely due to an increase in the number of individual contacts. This result has been confirmed by experiments on real surfaces (Thomas and Probert, 1970; Uppal and Probert, 1972*b*).

Many asperity models assume implicitly that each asperity deforms independently. If asperities are sufficiently close together, however, the stress patterns

set up when they deform will interact in such a way as to affect their behaviour as an ensemble.

The interaction of closely spaced asperities was studied by Bay *et al.* (1975) and some experimental evidence of asperity interaction has been found by Uppal and Probert (1973). Pullen and Williamson (1972) have suggested that material displaced from the peaks under load is redistributed so as to reappear as a uniform rise in the valley floors. However, Thomas (1975) argues that this is based on a misinterpretation of experimental results involving profile measurements where a constant reference height has not been maintained. However, most theories of plastic deformation simply assume that displaced material disappears, i.e. volume is not conserved. The resulting errors are small at large separations but become more serious as the surfaces approach more closely.

It has been noted that asperities transmit stresses greater than the flow pressure without themselves completely flattening. This phenomenon of asperity persistence has been studied by several researchers including Pullen and Williamson (1972). Williamson and Hunt (1972) have shown that it is not connected with work-hardening, whilst Chivers *et al.* (1974) do not dismiss the effects of work-hardening at higher loads and deformations. Asperity persistence has been explained as an effect of the distribution of sizes of asperities on a real surface by Childs (1977), with work-hardening only having a minor effect.

6.2.3 Concluding Remarks on Theoretical Methods

Although the complexities of the various elastic and plastic theoretical surface contact models all differ, in general, they all share the following shortcomings:-

1) They require a number of assumptions regarding, among other things, the probability distributions of asperity heights, slopes and curvatures, or they require the use of simple geometric asperity shapes.

2) Each discrete contact spot is assumed to deform independently.

3) They cannot predict the deformed shape of the surfaces, particularly for non-contacting areas.

4) When applied to almost all real rough surfaces, prior knowledge is required about the size range of topography that is important to the problem.

It is to overcome these difficulties that numerical methods of contact analysis have been developed over the years.

6.3 NUMERICAL ROUGH SURFACE CONTACT ANALYSIS

To find analytical solutions to elastic contact problems, assumptions have to be made about the undeformed shapes of the contacting bodies. Hence, in the previous

section, each contact model assumed a particular geometrical shape for the asperity tips.

The relaxation of this restriction has led to a number of numerical techniques being developed to consider the contact of arbitrarily shaped surfaces.

The essence of the problem of contact analysis is to determine, both inside and outside the contact area, the distributions of normal and tangential tractions which satisfy the normal and tangential boundary conditions at the interface, where the shape and size of the contact is not known at the outset. In general the tractions are coupled, but a considerable simplification can be achieved, with only a small loss of precision, for most engineering surfaces where the surface slopes are relatively small, by neglecting the effect upon the normal pressure of the tangential traction which arises if the moduli of the two bodies are different. Thus the contact is assumed to be frictionless.

In general the approach which has been used to solve this numerically has been to divide the contacting surfaces into small elements. Over each so called "traction element" (Johnson (1985)) an assumption is made about the local distribution of normal pressure.

Taking line contact as an example, the simplest representation for the pressure distribution is an array of locally concentrated forces. However, at each point of application of the force the deformation of the surfaces will be infinite. This can be overcome by representing

the pressure as adjacent columns of uniform pressure acting on discrete segments of the surface: the surface displacements are finite everywhere, but the displacement gradients are infinite between adjacent elements, where there is a step change in pressure. A piecewise-linear distribution of pressure, on the other hand, produces surface displacements which are everywhere smooth and continuous: such a distribution of pressure elements may be built up by the superposition of overlapping triangular pressure elements.

In three-dimensional contact the piecewise-linear pressure element is a regular pyramid on a hexagonal base. An array of such pyramids, erected on an equilateral triangular mesh and overlapping in such a way that every apex coincides with a mesh point, adds up to a resultant distribution of traction comprising plane triangular facets. The pressure distribution is then specified completely by discrete values of the heights of the pyramid shaped pressure elements.

From the shape of the traction elements, and the assumed elemental pressure distribution, it is possible to use half space theory to derive expressions for the displacements anywhere on the surface, due to an applied pressure on a particular element. The total displacement at a given point is merely the sum of the displacements due to all the elemental pressures. An application of this procedure is explained in Chapter 7.

In order to find the values of the traction elements which best satisfy the boundary conditions two different

methods have been developed (Johnson, 1985).

1) The Variational method in which the values of the pressure elements are chosen to minimise an appropriate energy function.

2) The direct, or Matrix Inversion, method in which the boundary conditions are satisfied exactly at specified "matching points", usually the mid-points of the boundary elements.

6.3.1 Variational Methods

Variational methods have been applied to contact problems for two reasons: (i) to establish conditions which will determine the shape and size of the contact area and contact stresses uniquely and (ii) to enable well-developed techniques of optimisation such as quadratic programming to be used in numerical solutions.

Kalker (1974) reviewed the variational theory and points out that it was developed almost exclusively in Italy and France. In their book, Duvaut and Lions (1972) show that for two bodies having continuous profiles, pressed into contact by an overall displacement δ, the true contact area and surface displacements are those which minimise the total strain energy U_E, provided there is no interpenetration of the surfaces.

For the numerical solution of contact problems it is

more convenient to work in terms of unknown tractions rather than displacements. Kalker (1977) has therefore proposed an alternative principle in which the true contact area and distribution of surface pressure are those which minimise the total complementary energy V^*, subject to the constraint that everywhere the contact pressure is positive.

This method has been applied to frictionless non-Hertzian contact problems by Conry and Seireg (1971) and Kalker and Van Randen (1972). The effects of the finite geometries of the contacting bodies have been considered by De Mul *et al.* (1986), the basic half space equations being modified for both the length and depth of the bodies.

De Mul *et al.* (1987) have used real measured surface data to look at the effect of debris dents on the fatigue life of roller bearings. Before carrying out the analysis they removed the small scale features from the surface roughness by filtering their data. This was necessary because the resolution of the traction elements they were using was not fine enough to model the small scale features. However, the results of their analyses were in close agreement with those found by experiment.

6.3.2 Matrix Inversion Method

The principle behind this is that if two bodies are pressed into contact by an overall displacement δ, and a guess is made about which traction elements are in contact, then it is possible to solve directly a set of simultaneous

equations for the unknown pressures. If the resulting pressure over any element is found to be negative then the system of equations is solved again after removing this element from the assumed contact geometry. Eventually convergence will be obtained with a particular set of contacting elements defining the area over which the bodies are in contact. This is more fully explained in Chapter 7.

Hartnett (1979 and 1980) used this technique with rectangular traction elements to find the surface pressure distributions for rolling element bearings. No account was taken of the roughness of the surfaces and ill-conditioning problems had to be overcome by splitting the contacting region into sub-regions. A solution was then found for the pressures due to the displacements within each sub-region whilst also taking account of the overall displacements of all other sub-regions.

This technique has been extended to consider the contact of layered solids (Chiu and Hartnett, 1983) and also the cylinder effect of the contact of cylindrical bodies (Chiu and Hartnett, 1987).

With the advent of modern computing facilities it has become feasible to use real measured topographic data as the input to numerical contact models. Such a model of two-dimensional plane strain contact has been presented by Francis (1982). He used a piece-wise linear approximation for the pressure distribution and a partitioned Gauss-Seidel method of solution for the simultaneous equations.

The contact of a ground surface profile with a flat surface profile was simulated and the resultant pressure distributions and contact geometries found for a number of applied loads.

Another plane strain model using recorded surface profiles has been presented by Webster (1986). This model uses constant pressure traction elements and Choleski factorisation to solve the simultaneous equations. The model has been used to study the contact of damaged bearing surfaces and the applied load against contact area relationships of sets of digitised rough surface profiles (Sayles and Webster, 1985; Webster *et al.*, 1986 - reproduced as Appendix B; Webster and Sayles, 1986; West *et al.*, 1987 - reproduced as Appendix C).

A similar model, using rectangular constant pressure traction elements, simulating the contact of three-dimensional digitised surface data has been presented by West and Sayles (1987) - reproduced as Appendix D. The derivation of this model is the subject of Chapter 7.

CHAPTER 7

THE THREE-DIMENSIONAL CONTACT MODEL

7.1 INTRODUCTION

Chapter 6 reviewed the different approaches used for rough surface contact analysis. It was shown that numerical techniques have only recently been applied to the problem of finding the surface stresses and exact contact geometries of real measured surfaces in contact. In general these techniques have been limited to two-dimensional contact of real rough surfaces or three-dimensional contact of smooth surfaces.

Webster and Sayles (1986) have developed a two-dimensional plane-strain elastic contact model using data gathered from a three-dimensional mapping system (Webster *et al.*, 1986 - reproduced as Appendix B). This chapter describes work carried out developing a similar three-dimensional contact model, including the effects of plasticity, which uses data from the same mapping system.

7.2 BASIS OF THE MODEL

7.2.1 The Geometry of Contact

The aim of the contact model is to predict the surface pressure distribution and deformed contact geometry of a real measured rough surface topography in contact with a smooth geometrically shaped surface. This can be used as a simulation of the contact of two rough surfaces because the contact of two uncorrelated rough surfaces can be modelled by the contact of a smooth surface and a rough surface having the sum of the variances of the two original surfaces. This can be proved as follows:-

Referring to Figure 7.1, the gap between the two bodies must be kept the same for the equivalent surface to represent the two rough surfaces in contact, therefore

$$gap = d-z = d-x_1-x_2 \tag{7.1}$$

giving

$$z = x_1+x_2 \tag{7.2}$$

but the variance of the equivalent surface, σ_z^2 is

$$\sigma_z^2 = E\{z^2\} = E\{(x_1+x_2)^2\}$$

$$= E\{x_1^2\} + E\{x_2^2\} + 2E\{x_1x_2\} \tag{7.3}$$

however, if surfaces 1 and 2 are uncorrelated the cross-

correlation function is zero, ie.

$$2E\{x_1x_2\} = 0$$

leading to

$$\sigma_z^2 = \sigma_1^2 + \sigma_2^2 \qquad (7.4)$$

The variance of the equivalent surface is therefore the sum of the variances of the two original surfaces (see also Thomas (1982)).

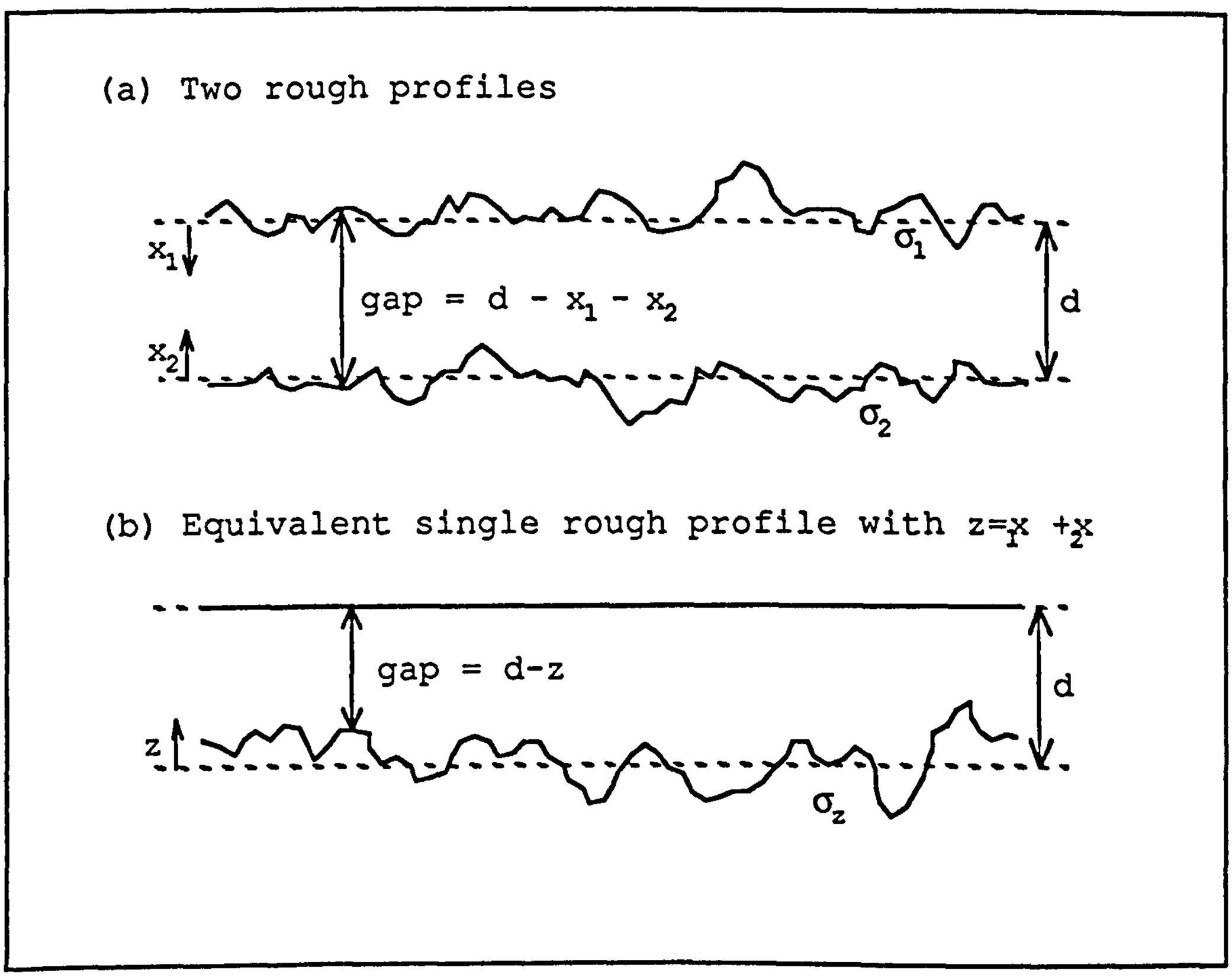


Figure 7.1 Two rough surface profiles and their equivalent single profile.

The geometry of the smooth body used in the contact simulation is that of a surface having principal radii of curvature in orthogonal directions, giving elliptical contours in the horizontal direction. This allows the simulation of most gear and bearing contacts to be attempted.

7.2.2 Constraints of the Model

Even though the model has been kept as general as possible some assumptions have been made, namely:-

1) Strains are small compared to the overall surface geometries, leading to the usual linear elastic theory equations.

2) The elastic contact area is contained within the geometric overlap region and there is no contact outside the region defined by the model.

3) Inside the contact region the penetration is just equal to zero and the surface pressures are greater than or equal to zero.

4) Outside the contact region there is no penetration and the normal surface pressures are zero.

5) The surface slopes are reasonably small so that forces act normally to the surfaces.

6) Tangential stresses and displacements are ignored, ie.

it is a frictionless contact.

The model is principly based on elastic contact but, so that the contact conditions can be made more realistic, a simple model of plasticity has been incorporated.

7.3 THE CONTACT MODEL

7.3.1 Contact Equations

The classical approach to finding the stresses and displacements in an elastic half-space due to surface forces is due to Boussinesq (1885) and Cerruti (1882), who made use of the theory of potential. The approach is well presented by Love (1952). The Boussinesq theory gives the normal surface deflection at a general surface point due to a concentrated contact at some other point. If the deflection of a surface point is required due to a number of concentrated contact points over an area S, then the total normal deflection u_z in polar coordinates (s,ϕ) is:-

$$u_z = \frac{(1-\nu^2)}{\pi E} \iint_S p(s,\phi)\, ds d\phi \qquad (7.5)$$

As the surface data to be used with the contact model is in the form of a rectangular matrix of height data, it would seem appropriate to use each known data point to be the centre of a rectangular area of sides 2a and 2b over which the normal pressure is constant. Note that the

sampling intervals in the x and y directions are 2a and 2b respectively (Figure 7.2). Pressure elements similar to these have been used in other contact models (see Chapter 6).

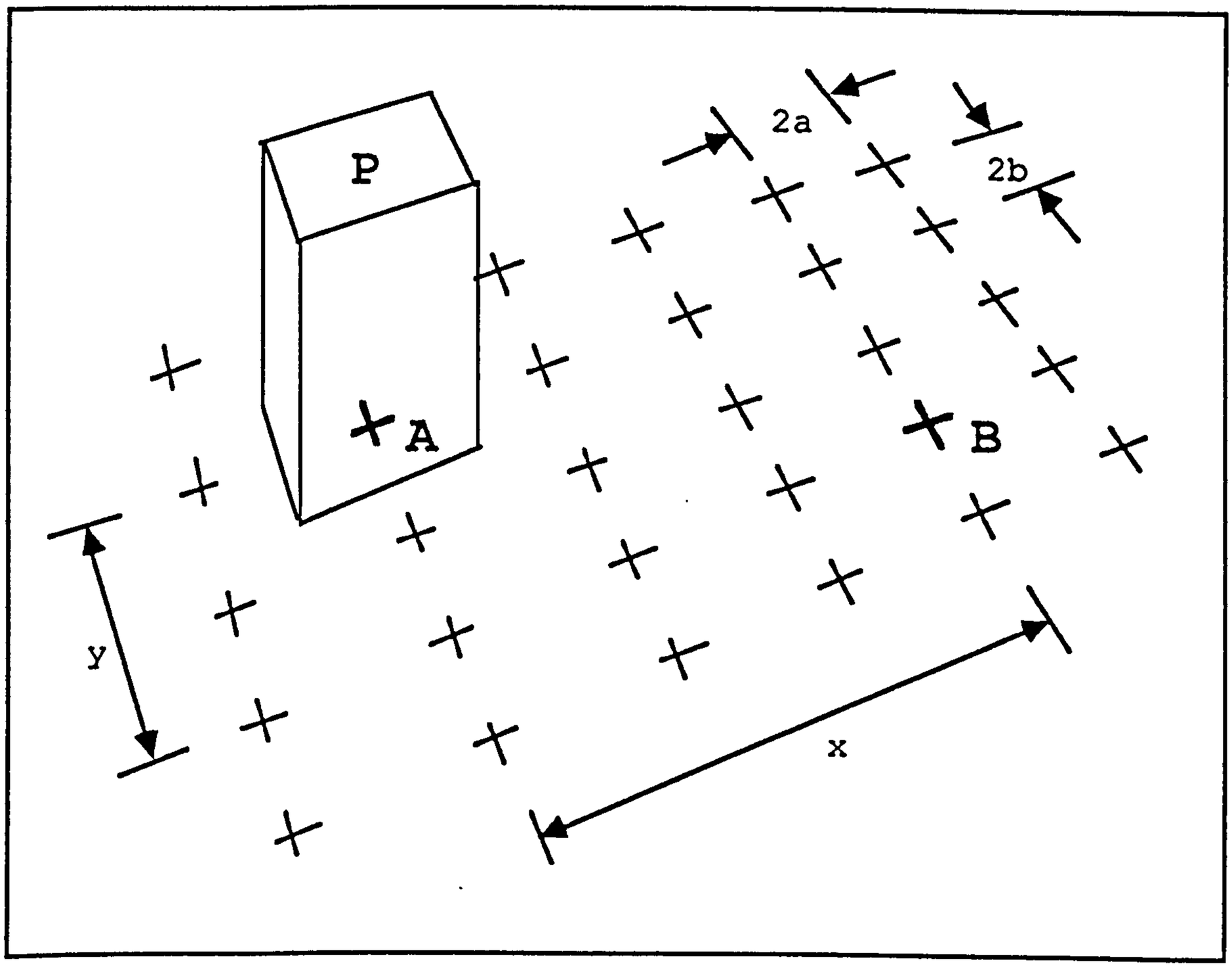


Figure 7.2 Detail of pressure element used:-
Each cross represents a data point and
P is the normal pressure on element.

The disadvantage of this method is that there is a discontinuity in the value of the pressure at the boundary of each element, thus a piecewise linear variation of pressure is not achieved. However, to extrapolate the

measured data onto the triangular grid required for a piecewise linear variation of pressure would add complications and inaccuracies to the calculation that would outweigh any increase in accuracy achieved by doing this.

The effect of a uniform pressure acting on a rectangular area of 2ax2b has been analysed in connection with rectangular foundations by Love (1929). Referring to Figure 7.2, the normal deflection of a general point B at (x,y) on the surface due to a uniform pressure p over a rectangle centred at the origin A is:-

$$u_z = \frac{(1-\nu^2)}{\pi E}\, p \left\{ (x+a)\ln\left[\frac{(y+b)+\left\{(y+b)^2+(x+a)^2\right\}^{1/2}}{(y-b)+\left\{(y-b)^2+(x+a)^2\right\}^{1/2}}\right] \right.$$

$$+ (y+b)\ln\left[\frac{(x+a)+\left\{(y+b)^2+(x+a)^2\right\}^{1/2}}{(x-a)+\left\{(y+b)^2+(x-a)^2\right\}^{1/2}}\right]$$

$$+ (x-a)\ln\left[\frac{(y-b)+\left\{(y-b)^2+(x-a)^2\right\}^{1/2}}{(y+b)+\left\{(y+b)^2+(x-a)^2\right\}^{1/2}}\right]$$

$$\left. + (y-b)\ln\left[\frac{(x-a)+\left\{(y-b)^2+(x-a)^2\right\}^{1/2}}{(x+a)+\left\{(y-b)^2+(x+a)^2\right\}^{1/2}}\right] \right\} \qquad (7.6)$$

This was used as the basis for the equations in the contact model. This equation, adjusted by putting the pressure as unity, is used to form the influence coefficient of any contacting point on any other point on the surface. The total deflection at any point can therefore be calculated by summing each element of pressure times it's influence coefficient.

$$\text{Total deflection of a point} = \sum \text{elemental pressures} * \text{influence coefficients} \qquad (7.7)$$

Thus, if the pressures on each contacting element are known, the deflection of any point on the surface and hence the deformed geometry can easily be found. However, the elemental pressures are not known. This can be overcome by giving the bodies a known bulk overlap δ and noting that the sum of the deflections in the two bodies at any contacting point is just equal to the geometric overlap.

The deflection in each body can then be found using the relationship:-

$$\frac{u_{z_1} E_1}{(1-\nu_1^2)} = \frac{u_{z_2} E_2}{(1-\nu_2^2)} \qquad (7.8)$$

By assuming some of the points to be in contact, then from the deflections at these points and their influence coefficients on all other points of the surface, a set of linear equations can be set up to find the unknown pressures at the contacting points. In matrix terms:-

$$(u_z) = [C]\ (p) \tag{7.9}$$

Where (u_z) is the vector of normal surface displacements at the contacting points, [C] is the matrix of influence coefficients and (p) is the vector of pressures at the contacting points.

Once the solution for the elemental pressures have been found then any elements where the pressures are negative are removed from the assumed contact region (tension is not permitted as adhesive forces are assumed to be relatively small). The deformations are then found everywhere to check that no contact occurs at any point outside the assumed contact region. Any points where this would occur are then added to the assumed contact region and the system of linear equations is then solved again for the new assumed contact points.

This process is repeated until convergence on a particular set of contact points is achieved. The deformed shapes of the bodies can then be found and the total normal load calculated by integrating the pressure distribution. The general principles of this method of solution by "Matrix Inversion" are outlined in Chapter 6 and Johnson (1985).

7.3.2 The Plasticity Model

Using a linear elastic contact model, such as the one described, will produce large pressure spikes on most

topographies even at very low loads due to the finite sampling and infinite pressure gradients between adjacent points of the surface data. These spikes, often well above the elastic limit of the materials being simulated, are obviously not realistic.

One partial solution to this problem is by the use of "Functional Filtering" (Thomas and Sayles, 1973). Both the highest and lowest frequencies present in the data can be filtered out just leaving the bandwidth of interest to the problem being solved. The method of three-dimensional filtering employed on the data is that of a moving average box filter (Webster *et al.*,1986 - reproduced as Appendix B). However, if within the scale of roughness being studied plasticity may play a part, it must be included.

7.3.2.1 The onset of plastic yield

On the basis that the contact of two rough surfaces begins with single asperity against asperity interaction it would seem appropriate to model each contact spot as the contact of two solids of revolution as the pressure required to cause plastic yield can be calculated for this situation using both the von Mise and Tresca yield criteria (see Chapter 6).

The maximum shear stress in the contact stress field of two solids of revolution occurs beneath the surface on the axis of symmetry. The maximum principal stress differences at the point with the greatest absolute

principal stress difference, assuming ν=0.3, are $0.62p_0$, $0.62p_0$ and 0.0 at a depth of 0.48a, where a is the Hertzian semi contact width.

Thus by the von Mises criterion the value of the maximum Hertzian pressure, p_0 for yield to occur is given by

$$p_0 = 2.8k = 1.6Y \qquad (7.10)$$

whilst by the Tresca criterion

$$p_0 = 3.2k = 1.6Y \qquad (7.11)$$

where k and Y are the yield stresses in simple shear and tension respectively.

From this it can be seen that whichever criterion is employed sub-surface yield begins when the maximum contact pressure reaches a value of 1.6Y.

7.3.2.2 The plasticity criterion

The plasticity criterion incorporated into the contact model is based on the assumption that as the applied load increases the material deforms elastically until a critical value of contact pressure, p_y is reached at which plastic yield begins. If the applied load is further increased the contact pressure does not rise above p_y but the surface of the material continues to deform and a zone of plastically deforming material is formed within the material below the point of application of the load.

Assuming the contact to be that between solids of revolution (see Section 7.3.3.1), the point of plastic yield is reached at a value of pressure, p_y such that:-

$$p_y = 1.6Y \qquad (7.12)$$

The plastic zone is contained within the interior of the body and the plastic yield manifests itself as a sub-surface elastic movement of the material axi-symmetrically away from the point of loading.

This is in agreement with the Johnson 'cavity' model of mixed elastic-plastic deformation (Johnson, 1985). In this model plastically deforming material appears to be lost within an imaginary cavity within the body of the material until the pressure becomes so great that the plastic zone breaks out onto the surface.

This criterion is added to the calculation by modifying the shape of the original undeformed surfaces at any point at which the pressure is above p_y so that the strain, and hence the pressure, is reduced. The elastic equations are then solved again for this new shape.

This process is repeated until convergence on a solution is reached with no pressures above p_y, and all points where plastic deformation has occurred are subjected to a constant pressure of p_y. The shapes of the modified original bodies then define the elastically recovered surfaces after contact.

It should be noted that this model of plastic

deformation is not valid for gross plastic flow as it does not generate the observed build-up of material around the periphery of the contacting region.

7.4 IMPLEMENTATION OF THE MODEL

The contact simulation program is a set of FORTRAN 77 routines implemented on an Acorn A443 Workstation micro computer. This fast 32 bit machine with 4 megabytes of RAM, although not as powerful or fast as a main-frame computer, was chosen so that it could form part of a stand alone surface data acquisition and analysis system.

The size of the memory of this computer effectively limits the surfaces to be analysed to approximately 20,000 data point heights with a maximum of 1250 data points in contact. However, this has been found to be enough for representative results to be obtained in most cases.

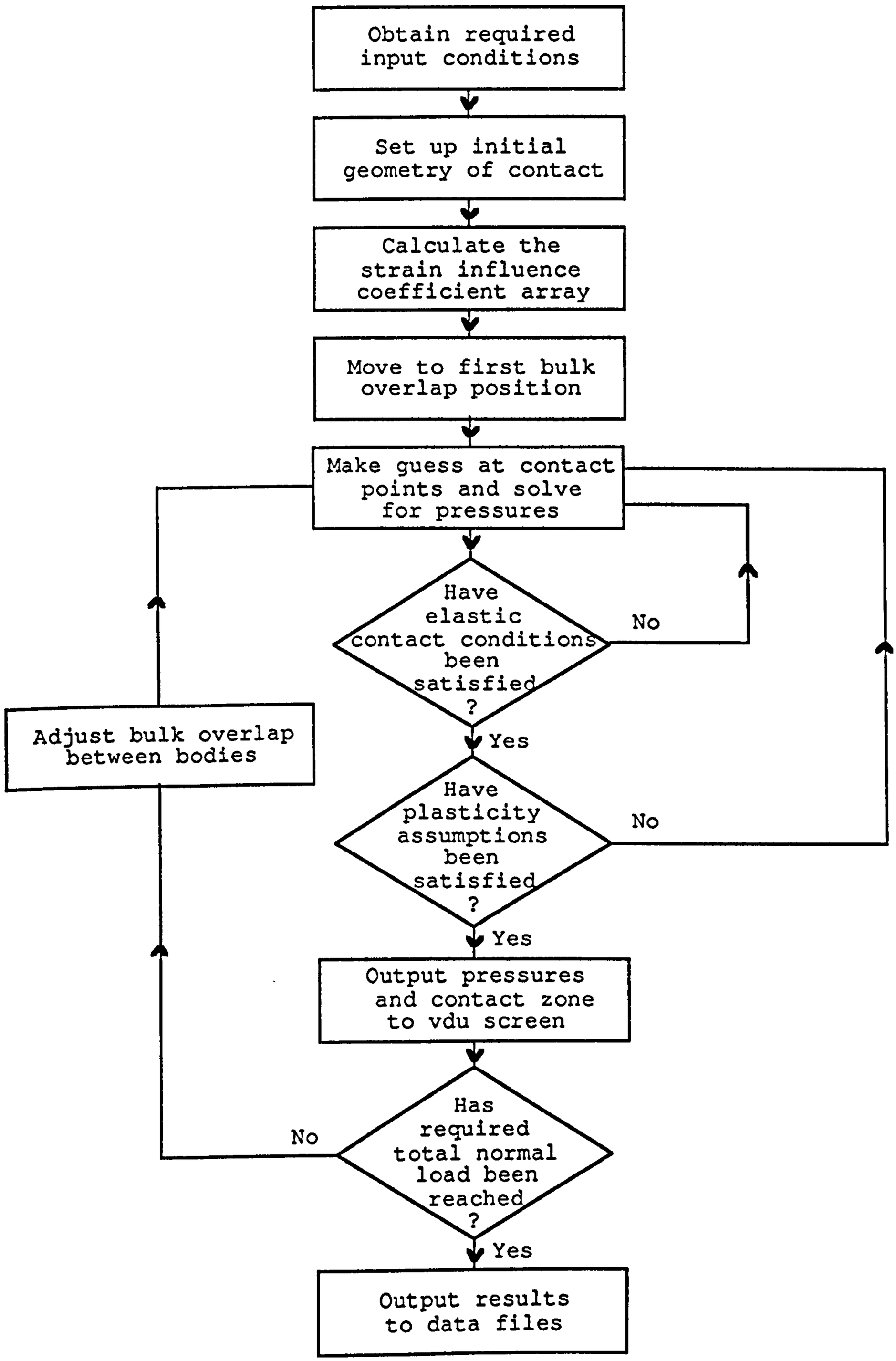


Figure 7.3 Flow diagram of contact program 3DCONT

7.4.1 The Contact Program

Figure 7.3 is a flow diagram of the contact simulation program which is called 3DCONT. The major features of this program are as follows:-

1) Obtain required input conditions

The operator is given a number of choices about the contact simulation to be performed. The file name of the measured rough surface has to be entered. The shape of the smooth body is defined by its principal radii of curvature parallel to and perpendicular to the direction in which the rough surface was measured (R_x and R_y) and the position on the surface above which the centre of these radii lie (C_x and C_y). The height of the smooth body z is therefore:-

$$z = \left(\frac{1}{2R_x}\right)(x-C_x)^2 + \left(\frac{1}{2R_y}\right)(y-C_y)^2 \qquad (7.13)$$

The material properties of the two bodies have to be input, namely the values of their Young's moduli E and their Poisson's Ratios ν.

If the effects of plasticity are to be included, then the yield pressure Y of each body has to be entered so that the maximum supportable pressure p_y can be calculated.

The program offers the choice of simulating the contact at a given applied normal load F or a calculation of the applied load and real contact areas at a series of bulk

overlaps δ (where δ increases by constant intervals) until the maximum applied normal load F is reached.

2) Set up initial geometry of contact

The required rough surface data file is read and a least-squares mean plane fitted to remove any gross form errors from the simulation. The coordinates of the smooth body are calculated using equation 7.13 and its height adjusted so that the two bodies are just touching but without any overlap.

3) Calculate the deflection influence coefficient array

The deflection influence coefficients (equation 7.6 with p taken as unity) that are used in the matrix of influence coefficients [C] (equation 7.9) can only have particular values. This is because there are only a limited number of different horizontal distances between data points within the contacting surfaces. In fact there are the same number of horizontal distances as data points.

An array of the possible values for the deflection influence coefficient is calculated so that the matrix of influence coefficients [C] for any particular set of contacting points can be constructed by using the appropriate elements without any further calculation.

The array is also used to find the deformed shapes of the contacting bodies by summing the deflections caused by each contacting point at all other points in the bodies.

4) Move to first overlap position

If the load versus area option has been selected then the bodies are moved to a bulk overlap δ equal to the incremental value in δ.

If the direct solution option has been selected then a guess at the required bulk overlap δ is made. Using the Hertz theory for the elastic contact of a smooth body with a flat surface the bulk overlap δ can be expressed as:-

$$\delta = \left(\frac{9F^2}{16R_e E^{*2}} \right)^{1/3} F_2(R_x/R_y) \qquad (7.14)$$

where the equivalent radius of curvature $R_e = (R_x R_y)^{1/2}$, and the shape function $F_2(R_x/R_y) = 1$ to a first approximation.

Thus using equation 7.14 an approximate value for the required bulk overlap can be found and the bodies are moved to this overlap.

5) Find contact points and pressure distribution

As a first guess at which points are in contact it is assumed that any point where the overlap is more than half the bulk overlap δ is in contact. The set of linear equations shown in equation 7.9 are then constructed, using the relevant strain influence coefficients and strains, and then solved to find the normal pressures.

The matrix of influence coefficients [C] is symmetric and the method of solution used is that of Choleski

Factorisation. This makes use of the symmetry so that only just over half of the full matrix has to be stored at any one time, thus making efficient use of the computer's memory. The leading diagonal of the matrix contains the greatest numerical values (the influence of a contacting point on itself) which suits the Choleski method as this condition must be met for the greatest accuracy of solution.

To check if the correct contact points have been assumed the pressures are checked for negative (tensile) values indicating a point which should not be in contact. The geometric overlap area is then searched for any assumed out of contact points where the solution indicates contact is occurring. These points are added to the assumed contact area.

If the assumed points seem to be correct a final check is made by calculating the deformed body shapes at the contacting points using the solution found and making sure that the average error from the surfaces making exact contact is less than one nanometer. If the error is larger than this the point with least pressure on it is removed from the assumed contact region.

If the effects of plasticity are to be included in the calculation then the pressure distribution is searched for individual pressures greater than p_y. If any are found the shape of the undeformed surfaces at that point are modified to remove a volume of material equal to the effective strain caused by the pressure being greater than p_y. If

the two bodies have the same yield stress Y then the removal of material is shared equally between them, otherwise all the material is removed from the weaker body. The pressure at the point is then reduced to p_y.

If any point has previously had material removed from it because the pressure was greater than p_y but now the pressure is below p_y, then material is added back to the bodies and the pressure raised to p_y.

If for any of the reasons detailed above the solution was not complete for this bulk overlap, then the modified assumed contact points are used for another solution of the simultaneous equations of equation 7.9 and the checks carried out again.

With the solution at this overlap complete graphics showing a contour map of the contact zone and a three-dimensional view of the normal pressure distribution are drawn on the vdu screen.

6) Check for final solution found

The pressure distribution is integrated to give the total normal load between the bodies for the given bulk overlap.

If the load versus area option has been selected then the normal load and real contact area are saved to a data file. If the normal load is less than the required final load the bulk overlap is increased by its incremental value and the solution attempted for this new contact condition.

If the direct solution option has been selected then the difference between the required and calculated normal load is found. If the error is greater than 1% then an estimate is made for the bulk overlap to give the correct normal load. The bodies are moved to this overlap and the solution attempted again.

7) Solution found print results

When the final solution has been found the final contact area and bulk overlap are recorded to the data file along with the results expected due to Hertz smooth body theory. The pressure distribution and deformed shapes of the bodies can also be recorded for later viewing through contour maps and three-dimensional views.

If the effects of plasticity were included then the modified shapes of the original bodies can be recorded and the proportion of the contact that has undergone plastic deformation found.

7.5 VALIDATION OF THE MODEL

7.5.1 Smooth Plane Tests

The model was tested against theoretical Hertzian elastic deformation results by simulations using a smooth plane instead of a rough surface.

Equation 7.14 gives the Hertzian relationship for the bulk overlap δ. Other Hertzian relationships are as

follows:-

$$\frac{a}{b} \approx \left(\frac{R_x}{R_y}\right)^{2/3} \tag{7.15}$$

where a and b are the contact ellipse x and y direction semi-axes.

$$(ab)^{1/2} = \left(\frac{3FR_e}{4E^*}\right)^{1/3} F_1(R_x/R_y) \tag{7.16}$$

where the shape function $F_1(R_x/R_y) = 1$ to a first approximation.

$$p_0 = \frac{3F}{2\pi ab} = \left(\frac{6FE^{*2}}{\pi^3 R_e^{\ 2}}\right)^{1/3} \left[F_1(R_x/R_y)\right]^{-2/3} \tag{7.17}$$

where p_0 is the maximum contact pressure.

For all shapes of smooth curved bodies used for simulated contact with a plane, the contact patches and pressure distributions were as predicted by theory within the accuracy of the finite data. Figure 7.4 shows the normal pressure distribution obtained when a smooth body with principal radii of curvature in the ratio 1:4 (R_x:R_y) is in contact with a smooth plane. For the bulk overlap used in this simulation the contact model overestimates the pressures and hence the normal load by 0.5% with the

contact ellipse semi-axes being correct to the nearest sampling interval.

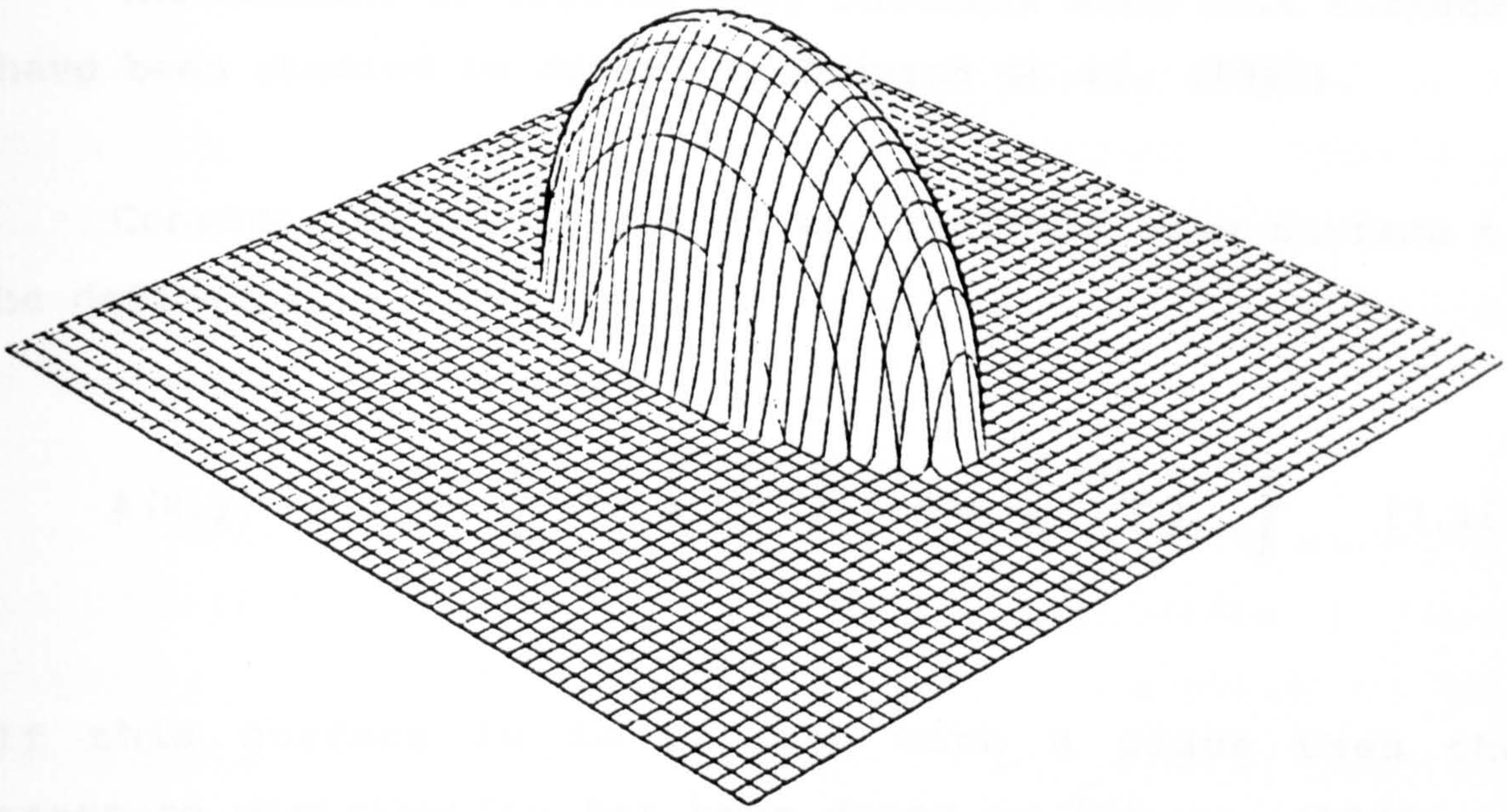


Figure 7.4 SMOOTH BODY IN CONTACT WITH A PLANE:-
SURFACE PRESSURE DISTRIBUTION
Area shown = 1.0 mm square
Load = 80 N, Hertz P_0 = 1.320 GPa
Smooth Body Radii of curvature = 5 mm/ 20 mm
Hertz semi-contact widths = 107 μm / 270 μm
Maximum pressure spike = 1.327 GPa

7.5.2 Geometric Surface Tests

As a first test of the contact model on rough surfaces, a simulation of contact with a geometrically shaped surface was attempted. The simplest model of a rough surface is a regular wavy surface which has a sinusoidal profile. The deformation will remain elastic provided that the amplitude Δ is small compared with the wavelength λ.

The contact of regular wavy surfaces with flat surfaces have been studied in detail by Johnson *et al.* (1985).

Considering the height of an isotropic wavy surface to be defined as:-

$$z(x,y) = \Delta \left\{ \cos\left(\frac{2\pi x}{\lambda}\right) + \cos\left(\frac{2\pi y}{\lambda}\right) \right\} \qquad (7.18)$$

If this surface is in contact with a plane then the pressure distribution has been found by Johnson (1985) to be:-

$$p(x,y) = \bar{p} + \frac{p^*}{2} \left\{ \cos\left(\frac{2\pi x}{\lambda}\right) + \cos\left(\frac{2\pi y}{\lambda}\right) \right\} \qquad (7.19)$$

where $p^* = 2\pi E^* \Delta/\lambda$ is the mean pressure required to cause contact everywhere between the surfaces.

Even for this simple contact situation we are unable to predict the shape of the areas of contact. The contact

areas are approximately circular at light load, become approximately square at higher load and eventually only small circular areas remain out of contact. With contact patch shapes varying in such a way an analytical solution to the problem seems improbable. However, asymptotic solutions can be found for early contact using Hertz theory and for nearly full contact by assuming the areas of no-contact to behave as penny-shaped cracks.

Johnson *et al.* (1985) used the numerical minimum complementary energy method due to Kalker (see Chapter 6) as a general solution to the contact problem. They also carried out experiments with a perspex and rubber model of the contact, which seemed to confirm that the numerical solution obtained was correct.

A comparison with this numerical solution was attempted by carrying out a simulated contact between a large diameter smooth ball and a surface consisting of one wavelength in each direction constructed using equation (7.18). The highest point on the surface was in the centre and coincided with the point of first contact with the ball. A ball rather than a plane was chosen for the simulation to compensate for the contact model taking no account of contact outside the area of the surface whereas the Johnson et al method used flat surface theory.

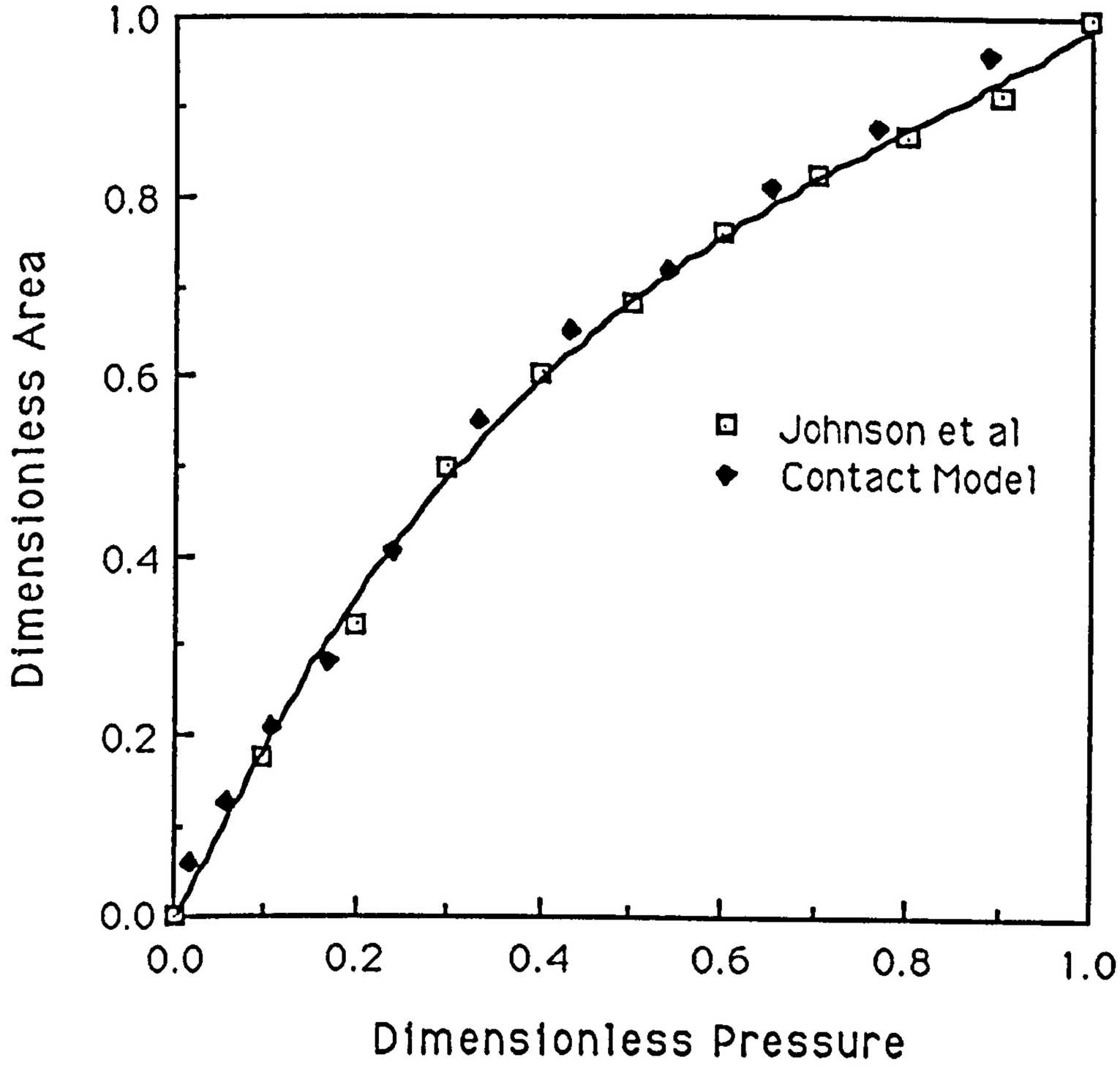


Figure 7.5 Real area of contact of a two-dimensional wavy surface with an elastic half-space.

Figure 7.5 is a plot of the variation in contact area and mean contact pressure plotted in non-dimensional form. The results using both the contact model and the minimum complementary energy method are shown. As can be seen the agreement between the two methods is very good which would seem to indicate that the contact model can cope with surface roughness.

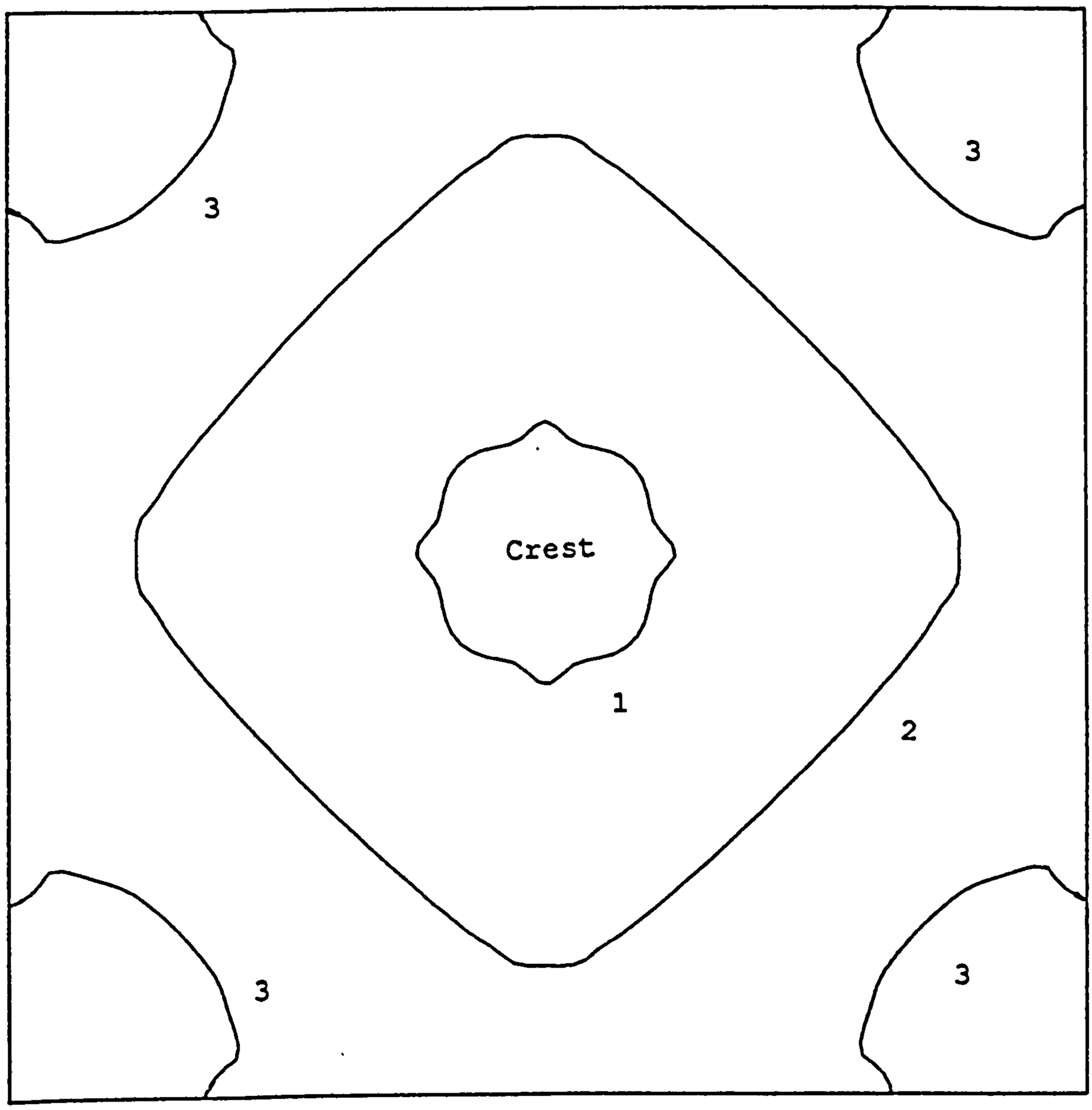


Figure 7.6 Changing shape of contact area with load using the contact model:-

1) 0.01 of pressure to cause complete contact
2) 0.25 of pressure to cause complete contact
3) 0.9 of pressure to cause complete contact

Figure 7.6 shows how the shape of the contact patches change as the load is applied. At low loads the contact patch is circular, but as the load increases it becomes square and then at high load the areas of no-contact become circular.

7.6 EXAMPLES USING REAL DATA

7.6.1 Worn Bearing

A 0.5 mm square area of worn bearing surface with some debris damage was used for a series of contact analyses (Figure 7.7). In all cases the normal load and theoretical contact area were the same but the ratio of the smooth body major to minor radius of curvature was varied.

At the load employed it was found that the actual contact area in all cases was approximately 57% of the theoretical smooth surface value and the maximum pressure spikes were approximately 4.5 times the theoretical Hertzian maximum. Assuming the material to be gear steel with a yield stress in simple tension of 2.6 GPa then the maximum pressure (4.5 GPa) would not be enough to cause any significant plastic deformation. Note that the surface used was that of a worn bearing where any plastic deformation should already have occurred.

Figures 7.8, 7.9 and 7.10 show the actual contact patch, deformed shape of the smooth body and pressure distribution for one of the test cases.

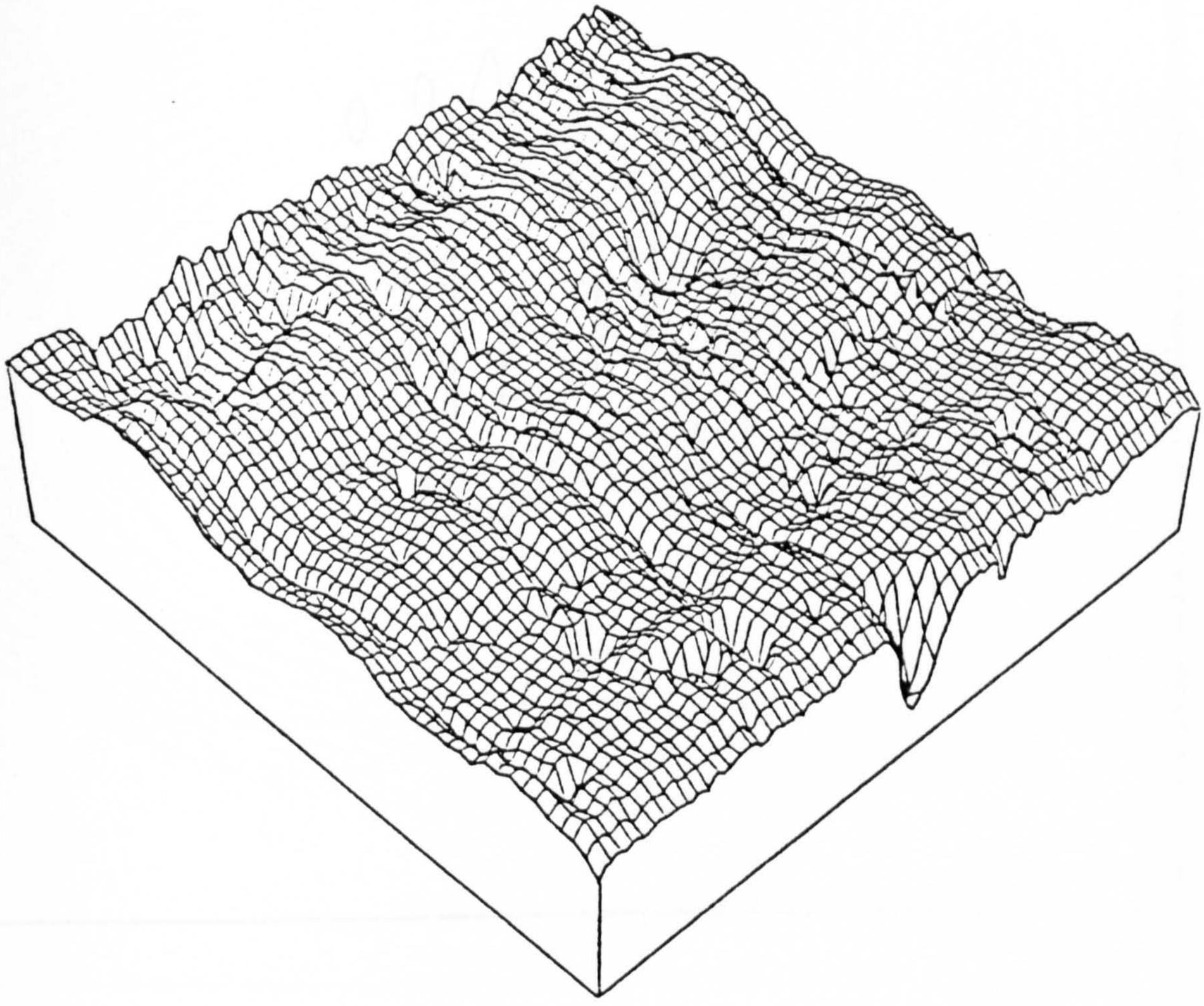


Figure 7.7 WORN BEARING SURFACE:-

Area = 0.5 mm square

Height range of data = 2.6 μm

Rolling direction = downwards to right

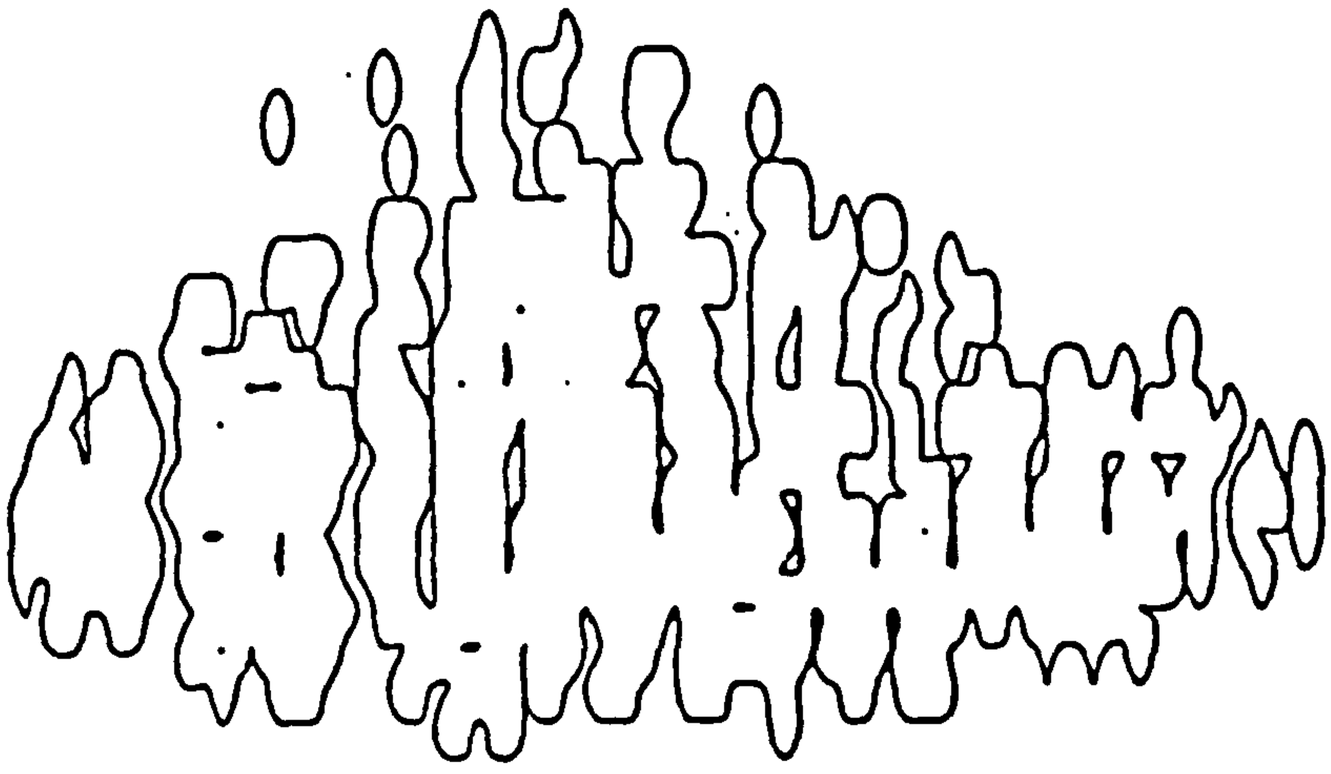


Figure 7.8 SMOOTH BODY IN CONTACT WITH WORN BEARING:-
CONTACT PATCH
Area shown = 0.5 mm square
Load = 25 N, Hertz P_0 = 1 GPa
Smooth Body Radii of curvature = 15 mm/ 5 mm
Hertz semi-contact widths = 162 μm / 78 μm
Ratio of true to apparent contact area = 0.57

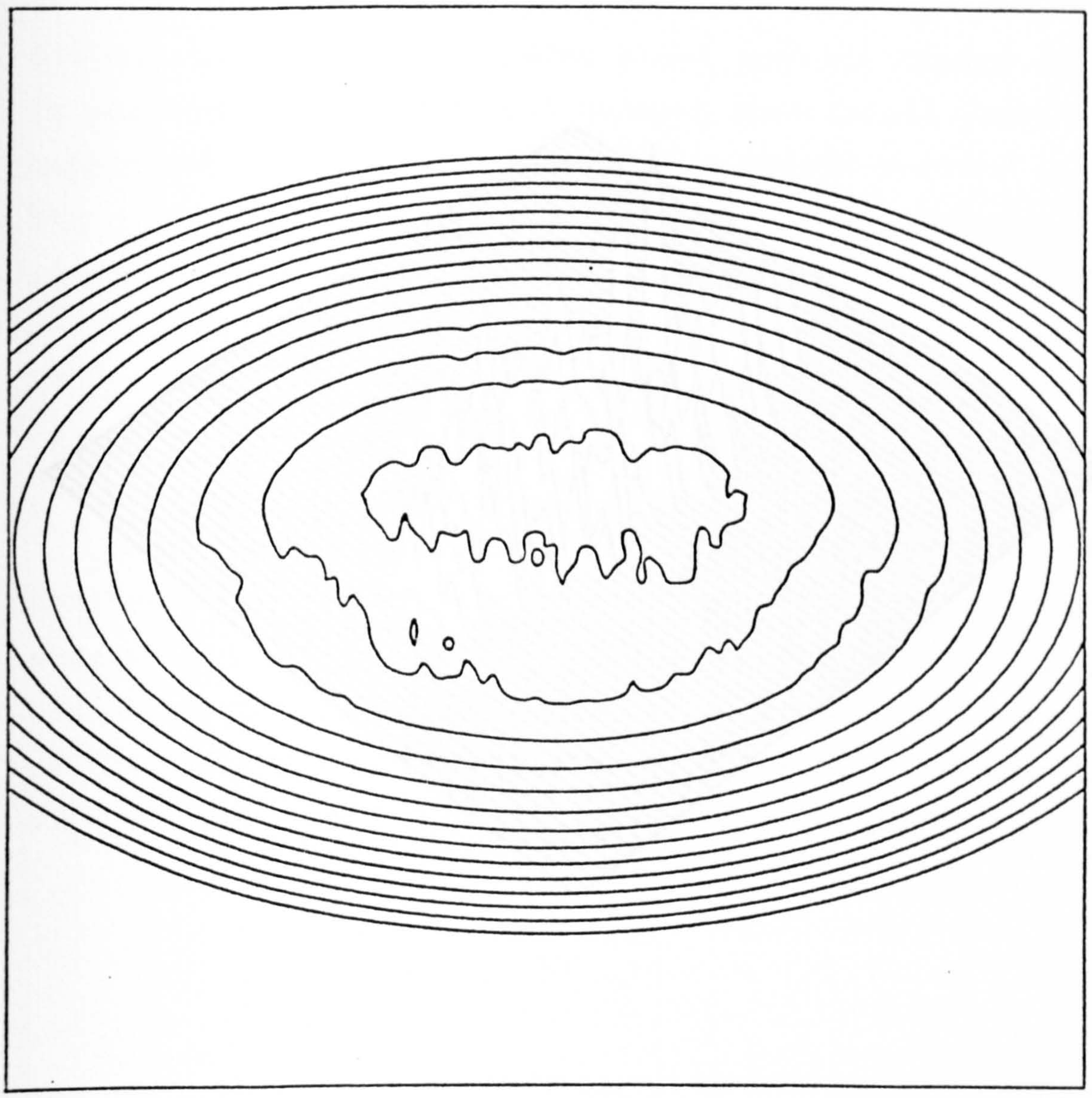


Figure 7.9 SMOOTH BODY IN CONTACT WITH WORN BEARING:-
DEFORMED SHAPE OF SMOOTH BODY
Load = 25 N, Hertz P_0 = 1 GPa
Smooth Body Radii of curvature = 15 mm/ 5 mm

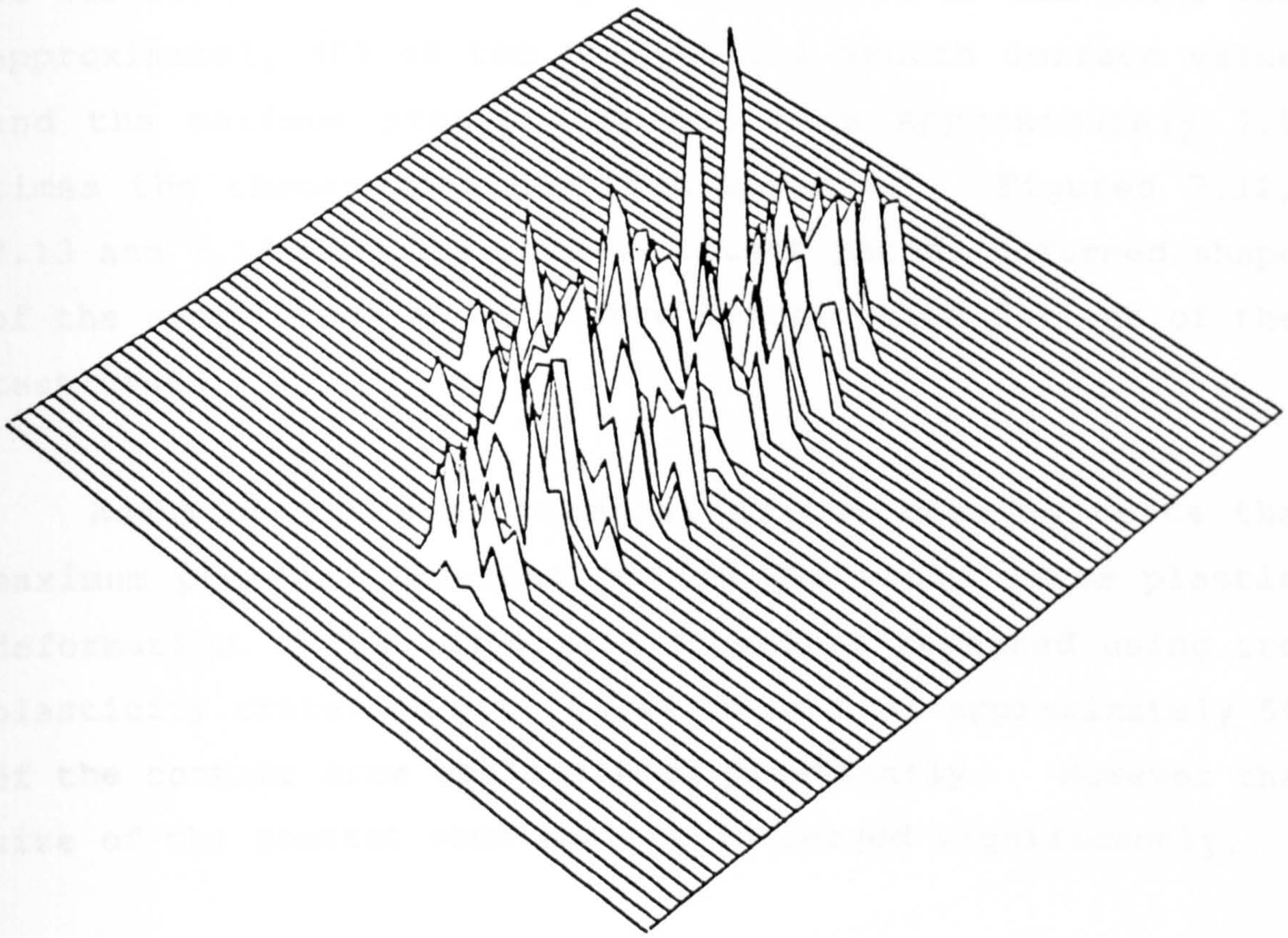


Figure 7.10 SMOOTH BODY IN CONTACT WITH WORN BEARING:-
SURFACE PRESSURE DISTRIBUTION
Load = 25 N, Hertz P_0 = 1 GPa
Smooth Body Radii of curvature = 15 mm/ 5 mm
Hertz semi-contact widths = 162 μm / 78 μm
Maximum pressure spike = 4.5 GPa

7.6.2 Abraded Steel Surface

Similar contact analyses were carried out using a 0.5 mm square area of abraded steel surface Figure 7.11. It was found that the actual contact area in all cases was approximately 40% of the theoretical smooth surface value and the maximum pressure spikes were approximately 7.0 times the theoretical Hertzian maximum. Figures 7.12, 7.13 and 7.14 show the actual contact patch, deformed shape of the smooth body and pressure distribution for one of the test cases.

Assuming the same material as the bearing tests the maximum pressure recorded of 7.0 GPa would cause plastic deformation. The tests were therefore repeated using the plasticity criterion and it was found that approximately 5% of the contact area would deform plastically. However the size of the contact area was not increased significantly.

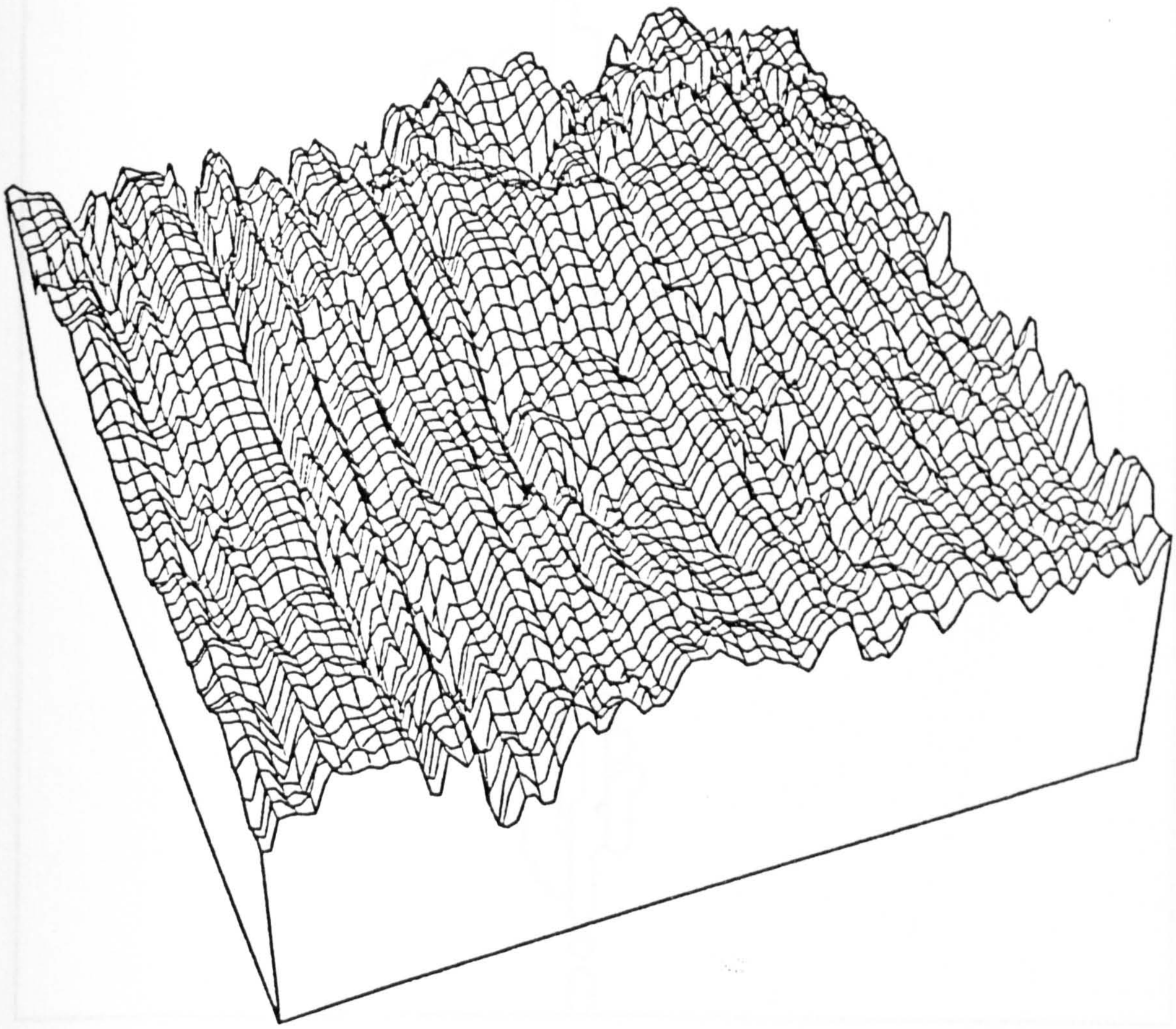


Figure 7.11 ABRADED STEEL SURFACE:-

Area = 0.5 mm square

Height range of data = 4.4 μm

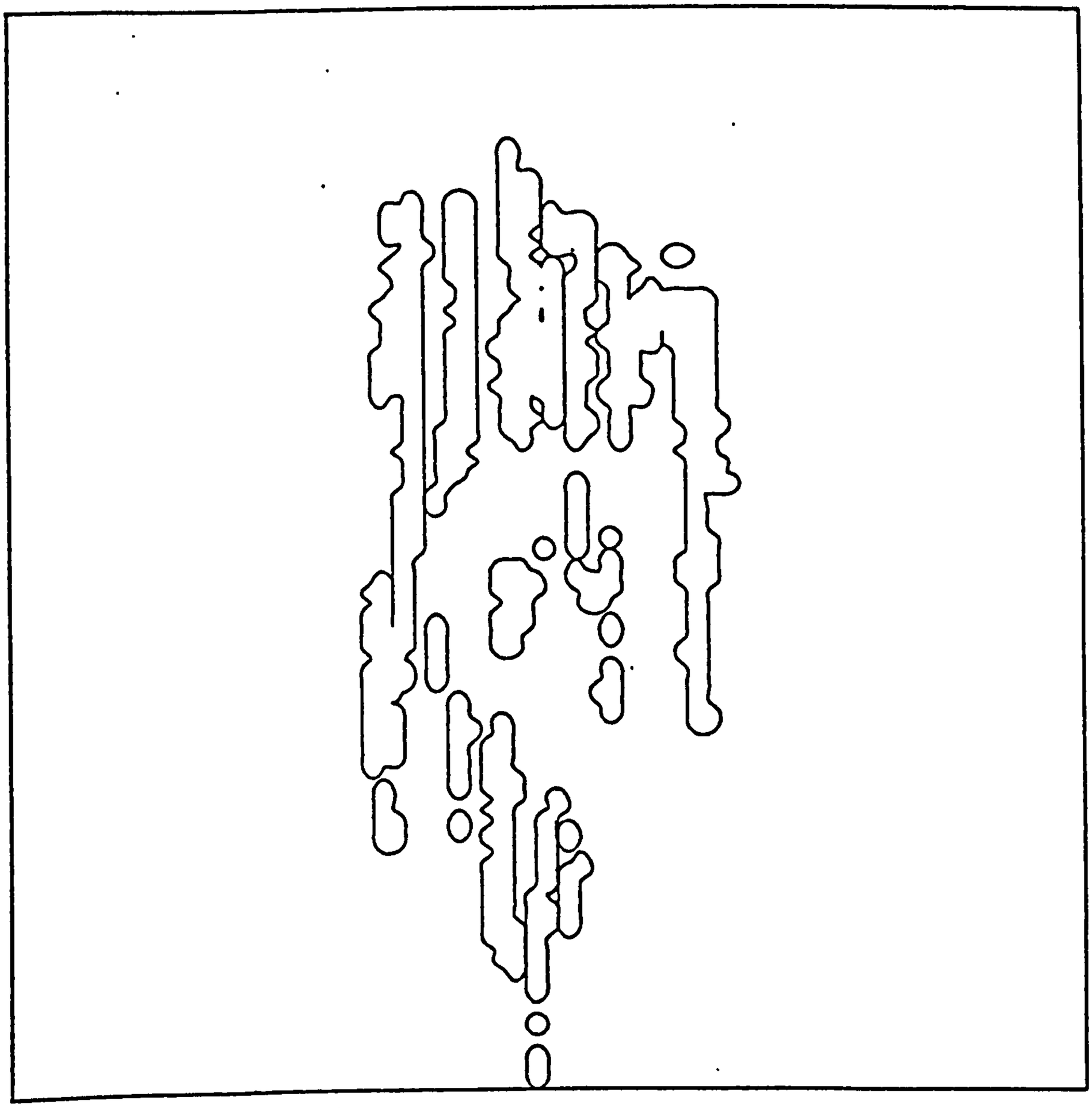


Figure 7.12 SMOOTH BODY IN CONTACT WITH ABRADED STEEL:-

CONTACT PATCH

Load = 25 N, Hertz P_0 = 1 GPa

Smooth Body Radii of curvature = .5 mm/ 15 mm

Hertz semi-contact widths = 78 μm / 162 μm

Ratio of true to apparent contact area = 0.39

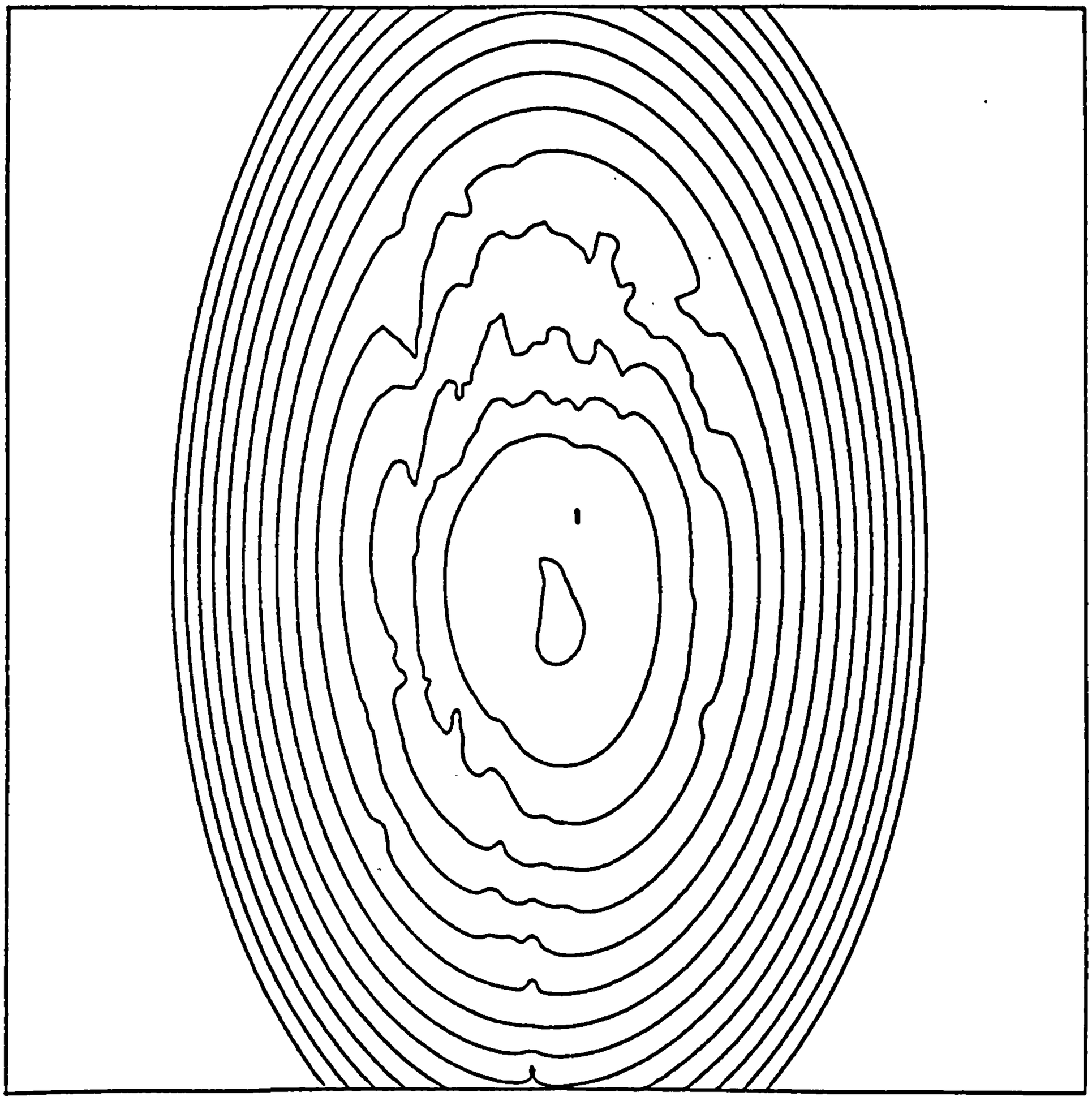


Figure 7.13 SMOOTH BODY IN CONTACT WITH ABRADED STEEL:-
DEFORMED SHAPE OF SMOOTH BODY
Load = 25 N, Hertz P_0 = 1 GPa
Smooth Body Radii of curvature = 5 mm/ 15 mm

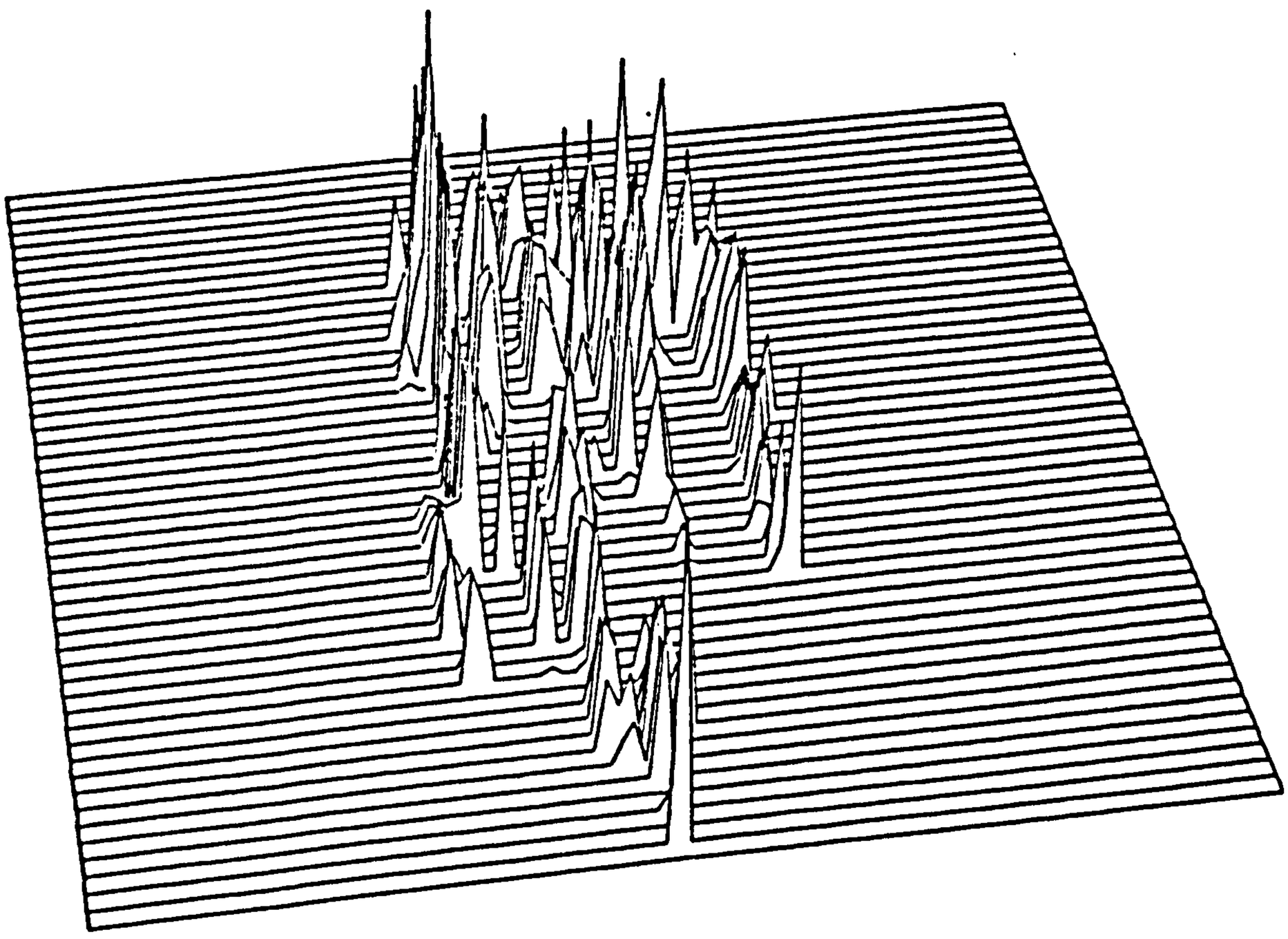


Figure 7.14 SMOOTH BODY IN CONTACT WITH ABRADED STEEL:-
SURFACE PRESSURE DISTRIBUTION
Load = 25 N, Hertz P_0 = 1 GPa
Smooth Body Radii of curvature = 5 mm/ 15 mm
Hertz semi-contact widths = 78 μm / 162 μm
Ratio of true to apparent contact area = 0.39
Maximum pressure spike = 7.0 GPa

7.7 THE EFFECT OF PLASTICITY

To study the effect of plasticity a simulation of the contact of a smooth 0.1 mm radius sphere on a smooth steel surface was carried out. The normal load was taken as 25 N as in the earlier tests so that this could represent a debris particle becoming trapped in the contact.

Taking the material to be the same as before virtually the whole contact undergoes plastic deformation. Figure 7.15 shows the 4 μm deep dent left in the surface after the load has been taken away. For comparison Figure 7.16 shows some detail of the debris damage on the worn bearing surface drawn to the same scale.

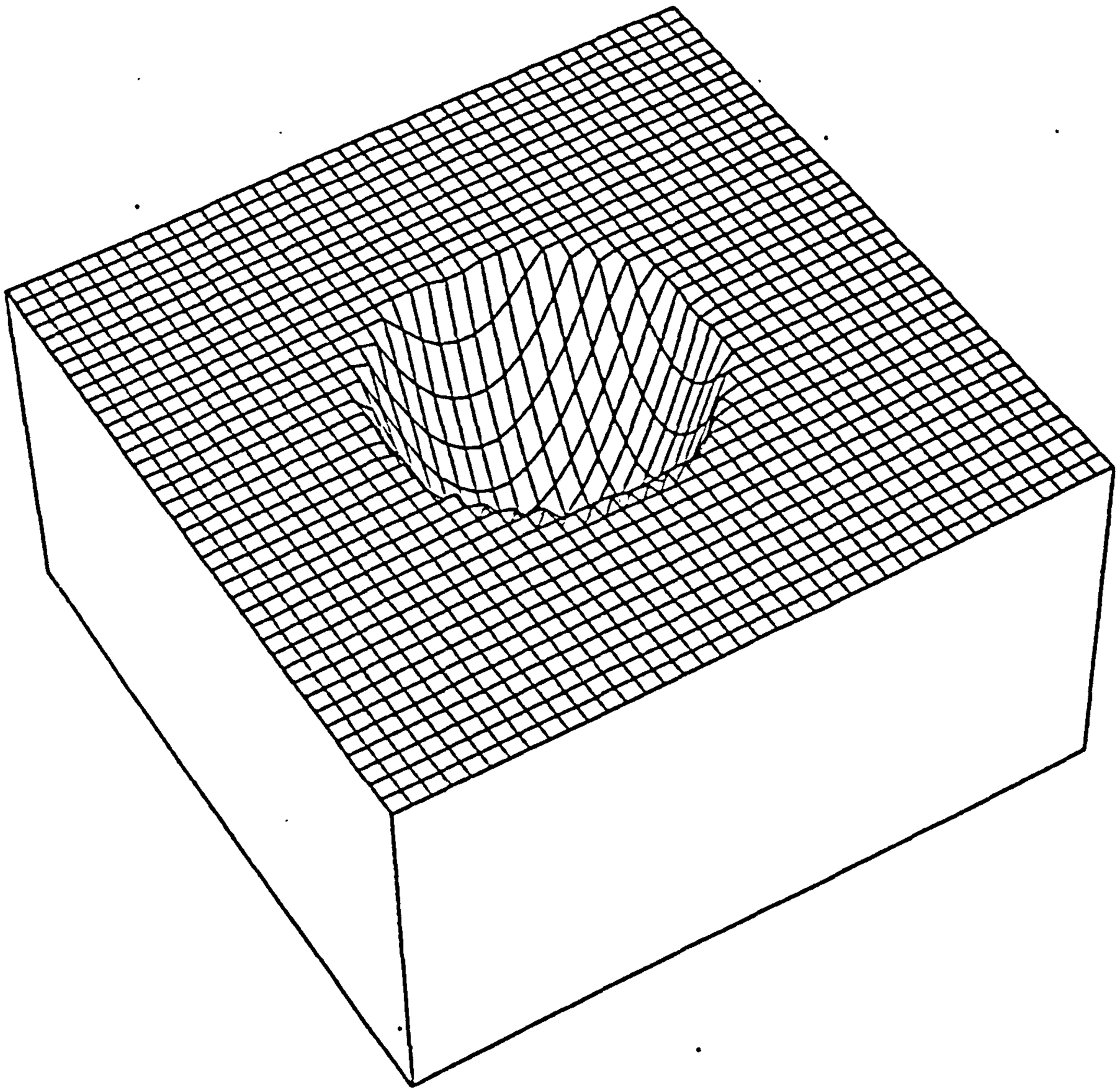


Figure 7.15 DENTED SURFACE:-

Dent caused by simulation of 0.1 mm radius sphere being pressed into a smooth surface

Area = 0.2 mm square

Height range of data = 4.0 μm

Load = 25 N, Hertz P_0 = 18.4 GPa

Yield Stress = 2.6 GPa

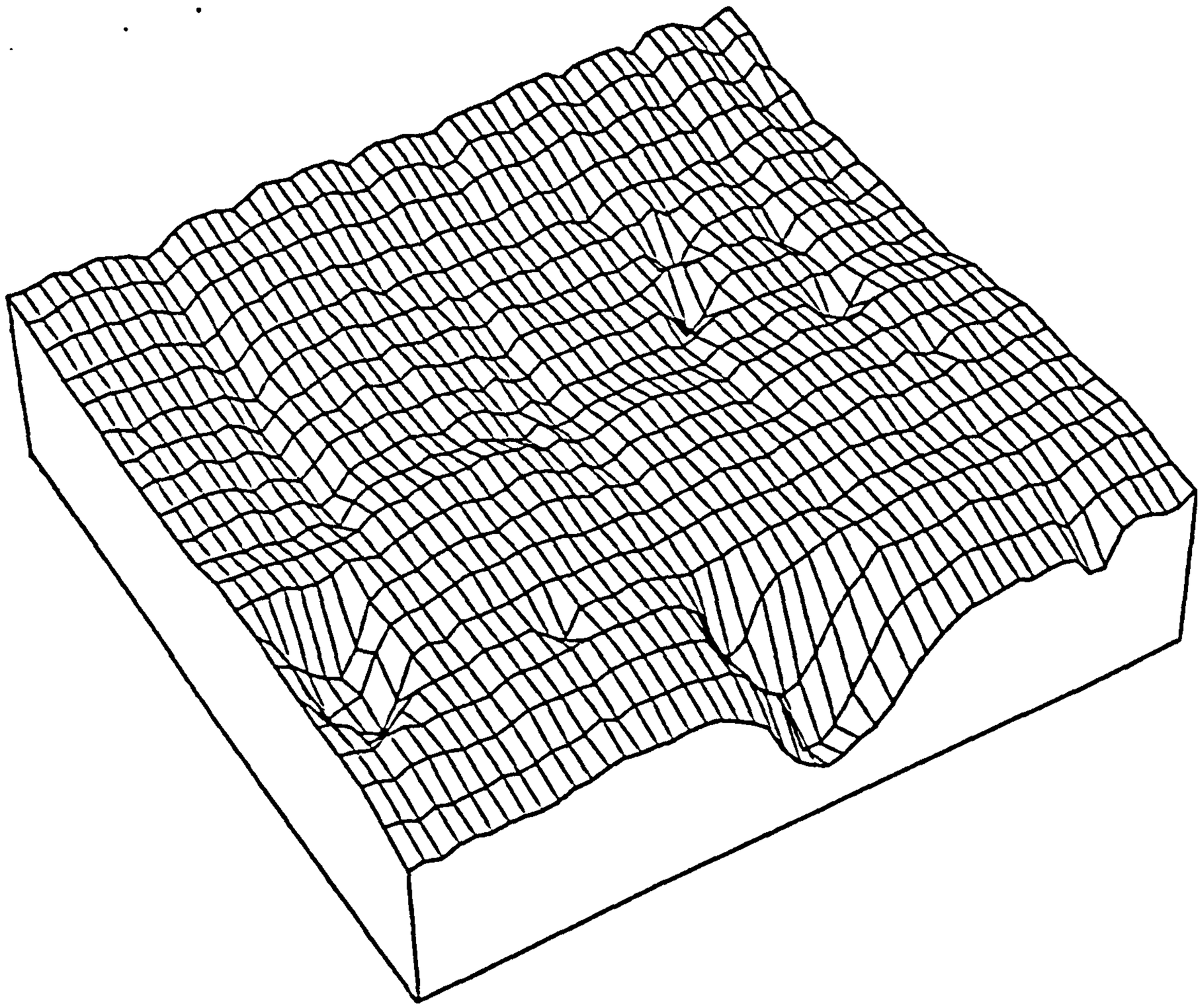


Figure 7.16 DETAIL OF WORN BEARING SURFACE:-

Area = 0.2 mm square

Height range of data = 2.6 µm

Rolling direction = downwards to right

7.8 CONCLUSIONS

A numerical technique is described which is capable of simulating the elastic contact of rough surfaces and estimating the effect of plasticity. The method is shown to be capable of creating the real pressure distribution together with a complete picture of local deformation and real contact area on an asperity scale and also to define the shapes of the recovered surfaces after plastic deformation. The technique is illustrated by its application to the contact in several typical engineering surfaces.

CHAPTER 8

3-D CONTACT ANALYSIS OF THE TEST SPECIMENS

8.1 INTRODUCTION

The following sections detail initial computer simulations of three-dimensional contact of the mini-disc specimens used for the experimental programme. This illustrates the use of the three-dimensional contact model described in Chapter 7 with realistic data.

8.2 THE CONTACT SIMULATIONS

Chapter 2 describes the topography measurements made of the test specimens used for the experimental programme. In particular three-dimensional surface maps of three of the specimens (specimens 1, 3 and 15). These three surface maps were used as input data to the three-dimensional rough surface contact model.

To simulate the contact conditions between the disc specimens, the smooth body brought into contact with the measured surfaces was given a radius of curvature of 5 mm in the circumferential direction and 20 mm in the axial direction. This approximated to the relative radius of

curvature in the circumferential direction and a line contact in the axial direction.

In each of the three simulations the loading was such that the maximum Hertzian smooth surface contact stress was 0.8 GPa with the Hertzian semi-contact widths being approximately 70 mm and 175 mm respectively.

Note that in all simulations it was assumed that no plastic flow occurred so that the plasticity criterion was not used.

8.3 THE CONTACT SIMULATION RESULTS

Figures 8.1 to 8.3 show the resultant surface pressure distributions obtained from the three analyses and Figures 8.4 to 8.6 the corresponding contact patches. Note how different the contacts are in each case.

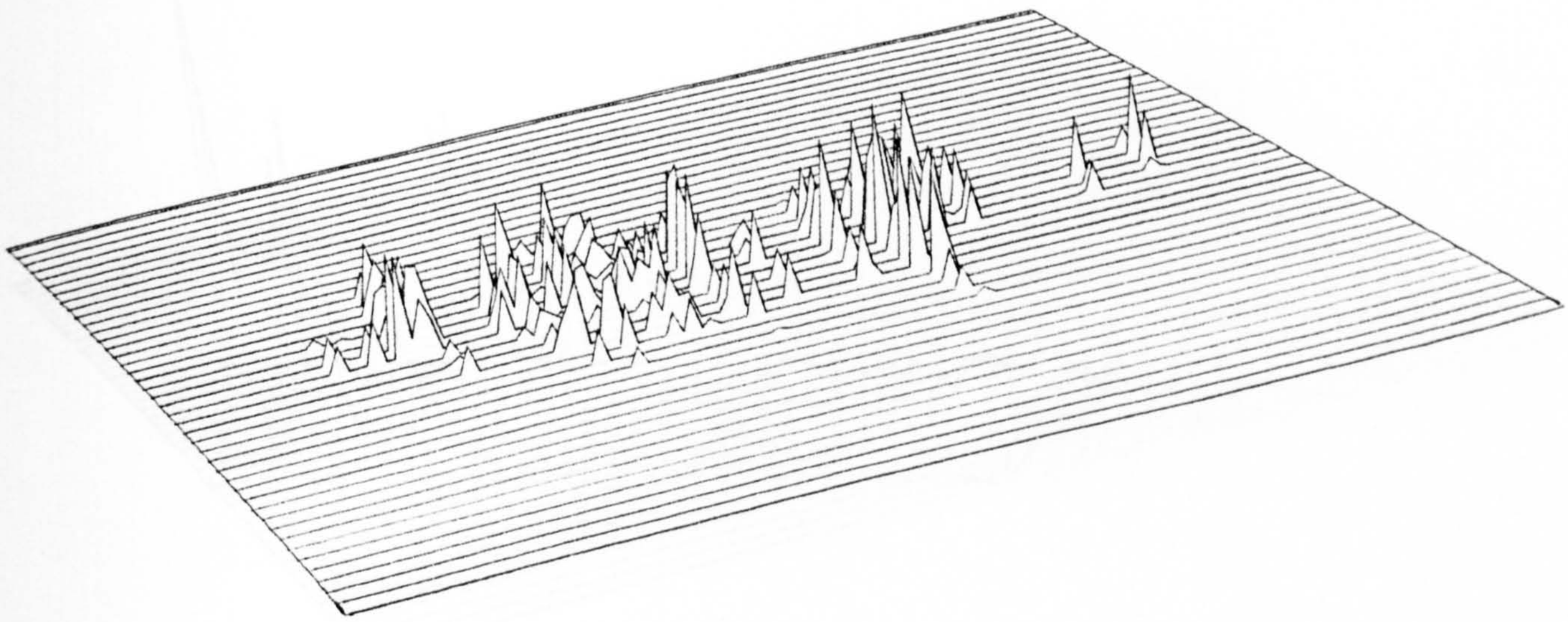


Figure 8.1 Surface pressure distribution between smooth surface and test specimen 1:-
Mapped area = 0.5 mm * 0.5 mm
Maximum pressure spike = 8.01 GPa

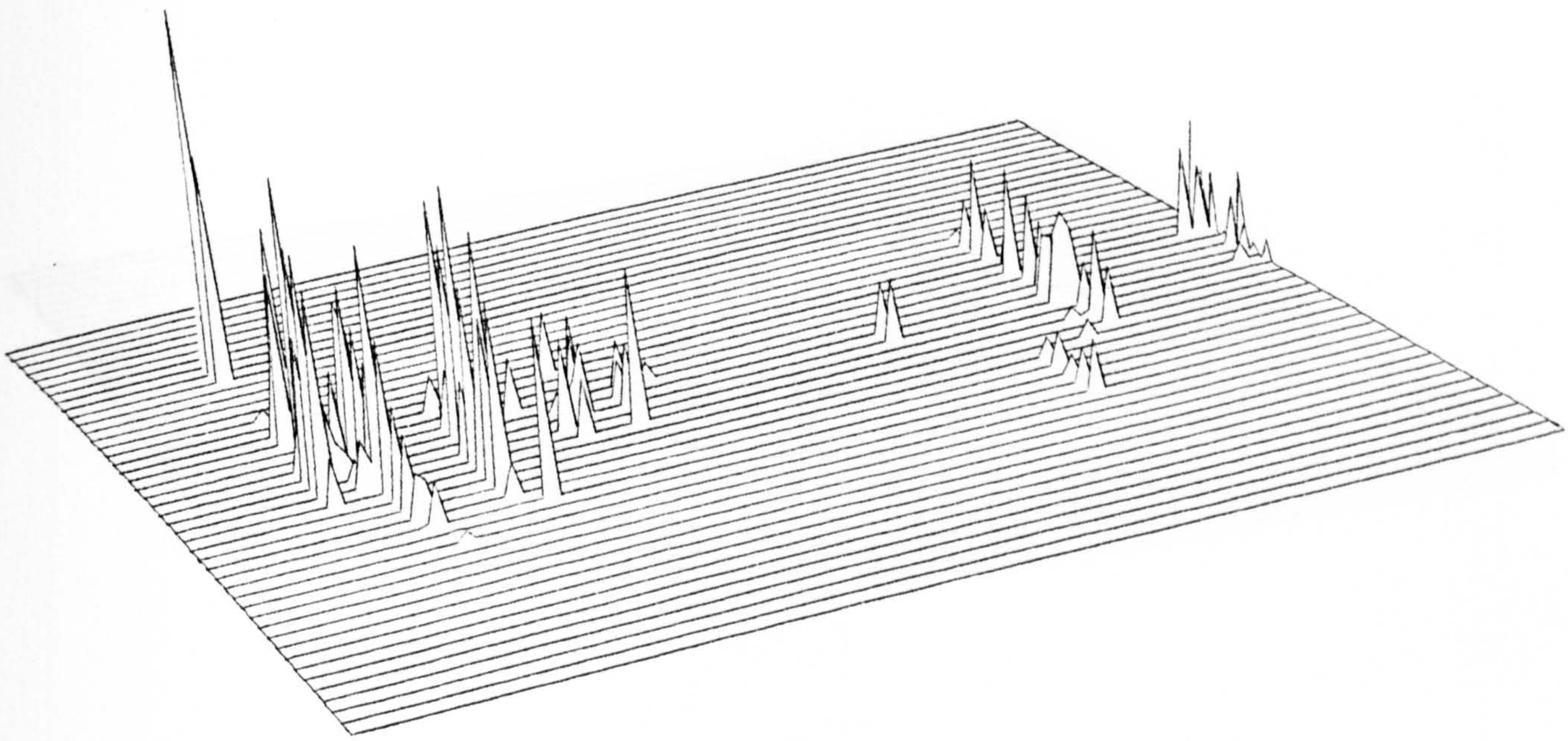


Figure 8.2 Surface pressure distribution between smooth surface and test specimen 3:-
Mapped area = 0.5 mm * 0.5 mm
Maximum pressure spike = 19.95 GPa

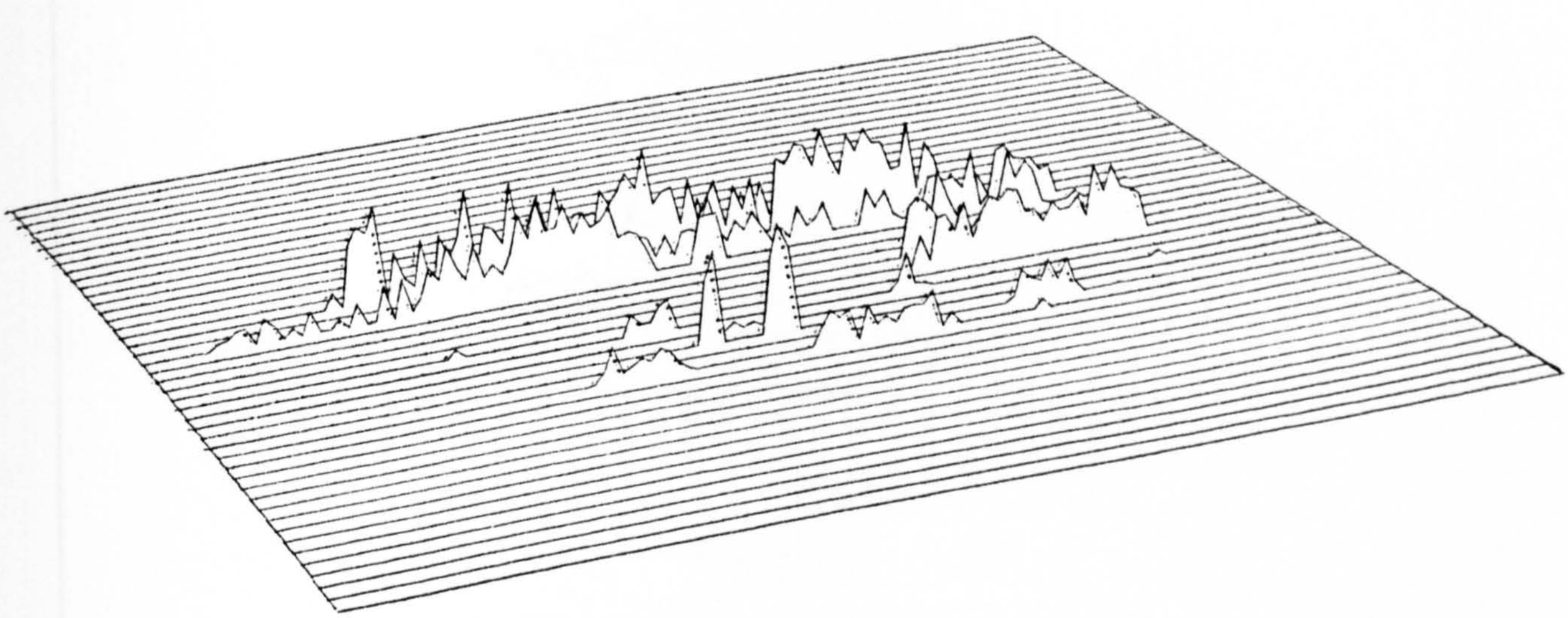


Figure 8.3 Surface pressure distribution between smooth surface and test specimen 15:-
Mapped area = 0.5 mm * 0.5 mm
Maximum pressure spike = 6.47 GPa

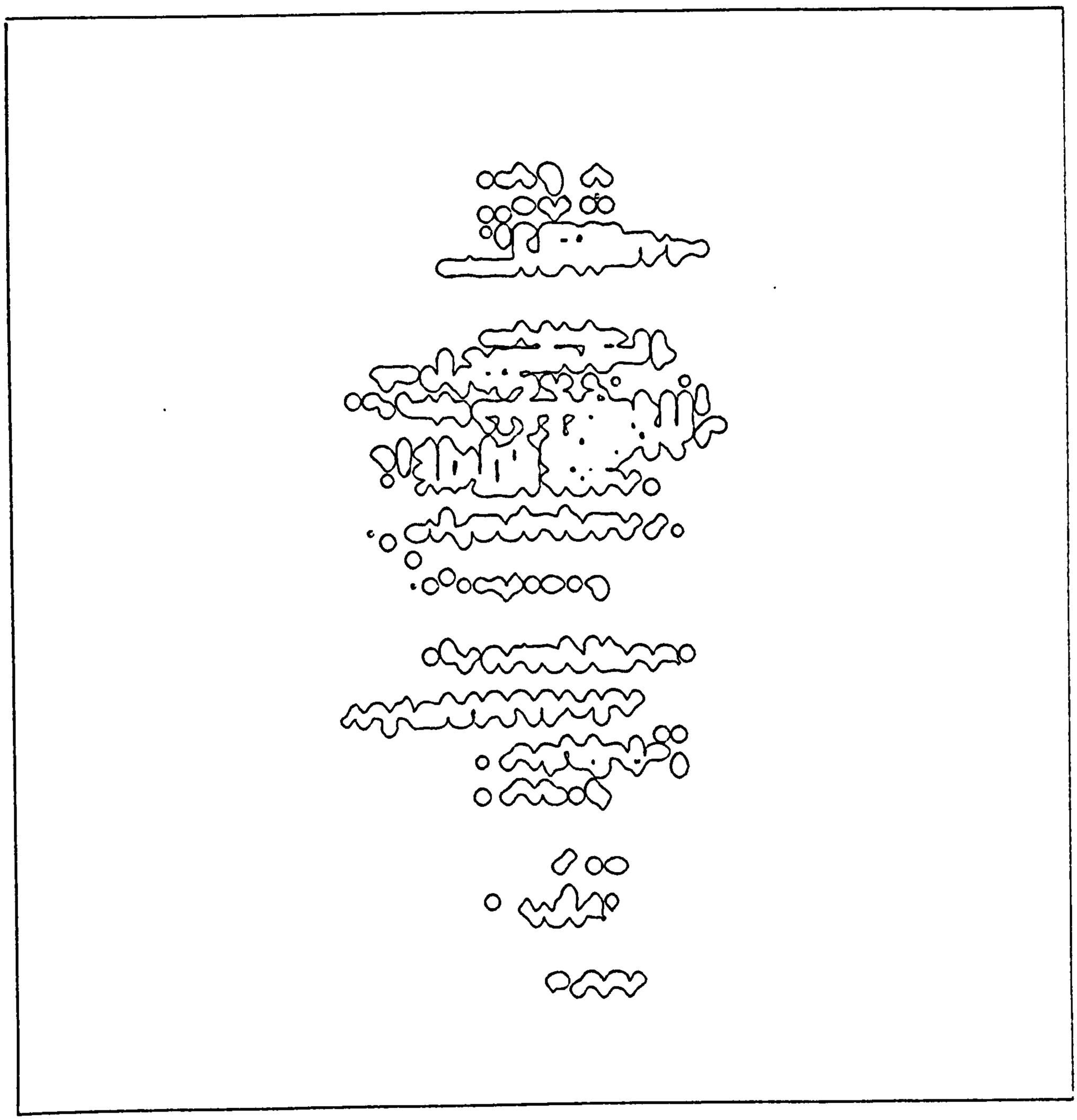


Figure 8.4 Contact patch between smooth surface and test specimen 1:-
Mapped area = 0.5 mm * 0.5 mm
Maximum pressure spike = 8.01 GPa

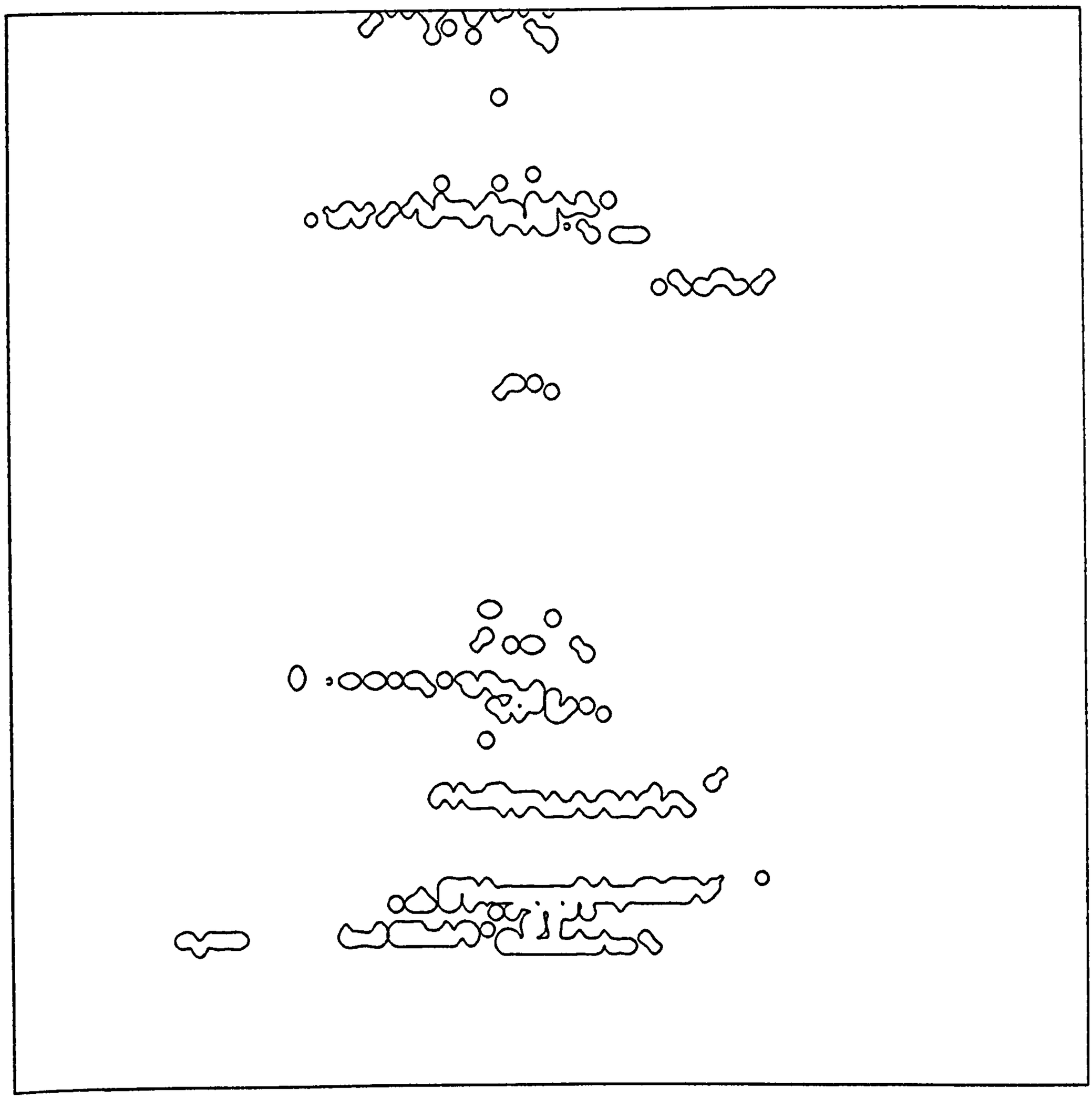


Figure 8.5 Contact patch between smooth surface and test specimen 3:-
Mapped area = 0.5 mm * 0.5 mm
Maximum pressure spike = 19.95 GPa

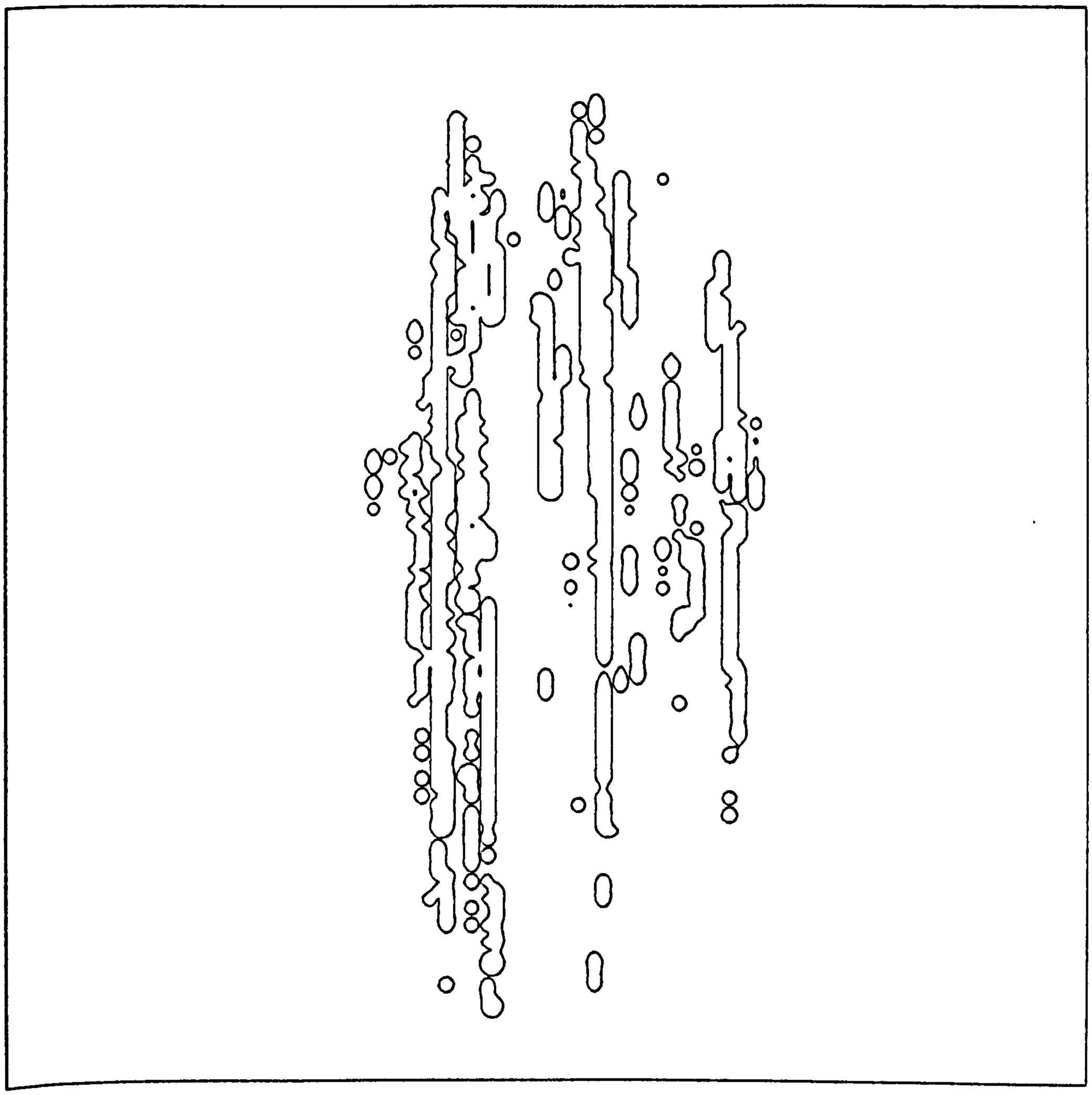


Figure 8.6 Contact patch between smooth surface and test specimen 15:-
Mapped area = 0.5 mm * 0.5 mm
Maximum pressure spike = 6.47 GPa

Parameter	Specimen 1	Specimen 3	Specimen 15
Max Pressure (GPa)	8.01	19.95	6.47
Max Hertzian Pressure (GPa)	0.80	0.80	0.80
True to Apparent Contact Ratio	0.277	0.132	0.320
Average RMS Roughness (μm)	0.212	0.529	0.163

Table 8.1 Summary of results of three-dimensional contact simulations.

Table 8.1 summarises the results of the contact analyses and includes a surface RMS roughness parameter. This roughness was obtained from the circumferential and axial roughnesses of the specimen surfaces taken from Table 3.1 as:-

$$\text{Average RMS Roughness} = \sqrt{\text{Circumferential RMS Roughness} \times \text{Axial RMS Roughness}} \qquad (8.1)$$

Figure 8.7 plots this parameter against the ratio of the maximum surface pressure recorded to the smooth surface

Hertzian maximum pressure for each specimen.

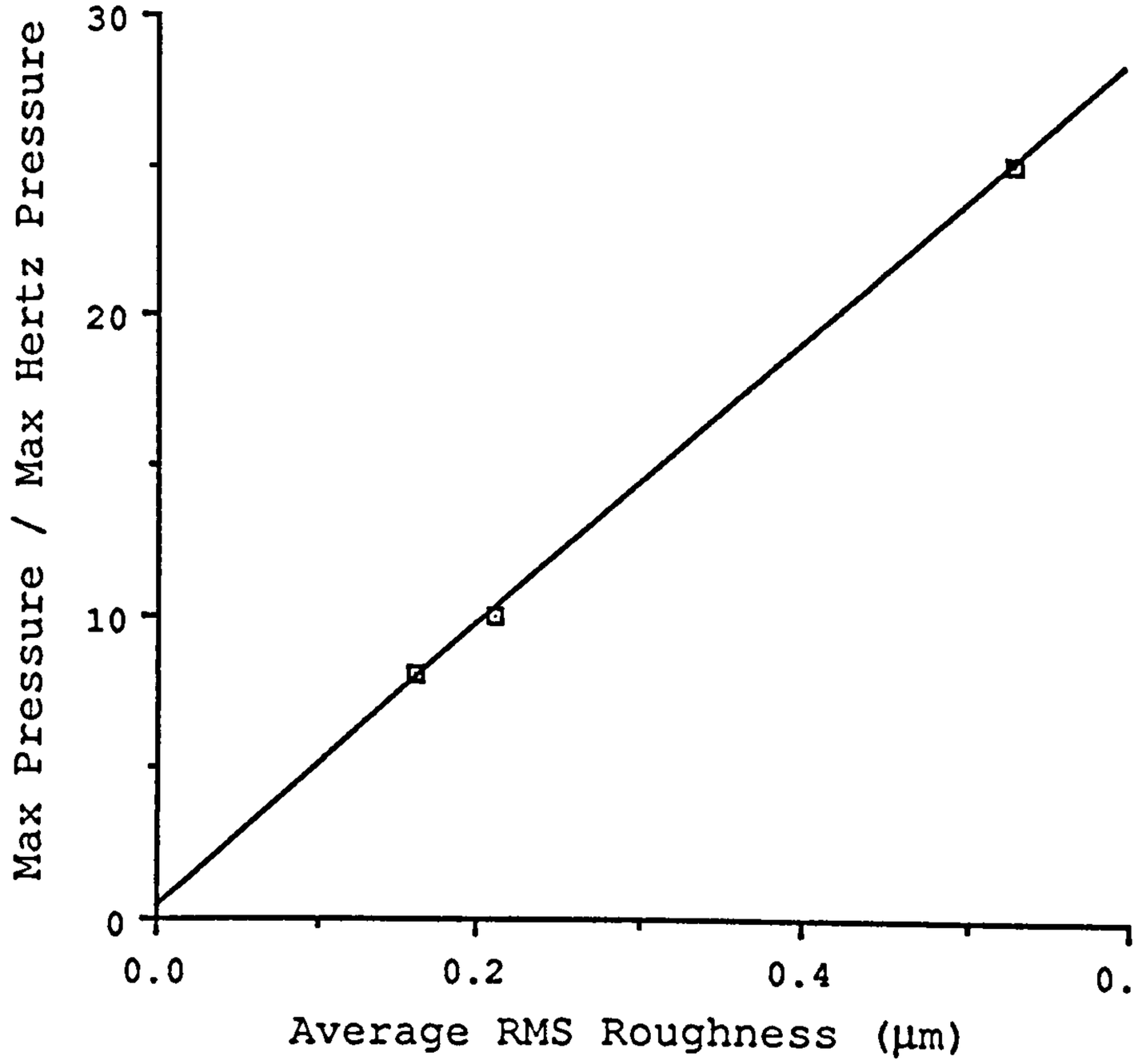


Figure 8.7 Ratio of maximum pressure to Hertzian pressure against average surface roughness for the test specimens.

From Figure 8.7 it would seem that there is a direct relationship between the maximum recorded surface pressure and the average surface RMS roughness. Thus it would seem that combining orthogonal two-dimensional surface measurements gives a reasonable approximation of three-dimensional surface parameters. More importantly it may enable a few full three-dimensional analyses to be used to

estimate the contact conditions in other surfaces where only two-dimensional parameters are known.

The fact that the three-dimensional surface contact results correlate well with the two-dimensional profile measurements is further evidence that to study the contact of surfaces, contact models or the use of profile measurement parameters show great promise. This is in marked contrast to the poor results obtained from the experimental programme utilising the optical vibration monitor.

CHAPTER 9

CONCLUSIONS

9.1 GENERAL CONCLUSIONS

This project, studying the effects of surface finish on rolling and rolling-sliding contacts, such as those found in gears and bearings, has been carried out in three basic parts:-

1) An experimental programme using a mini-disc machine to investigate the links between surface finish and vibration.

2) The development of a theoretical model of a profile structure function allowing roughness characteristics to be studied on an asperity size basis.

3) The development of a numerical technique capable of simulating the elastic contact of rough surfaces and estimating the effect of plasticity.

The conclusions reached from each of the three parts of the project are detailed in the following sections.

9.1.1 The Experimental Programme

Commissioning the mini-disc machine used for the experiments has been successfully completed. This provides a very versatile machine for many forms of test, with its speed controls giving infinitely variable slide-roll ratios to produce a whole range of contact conditions.

A prototype optical vibration monitor has been developed. Although it requires more investigation it has shown a high potential as an on line monitoring device of contact conditions. The vibrational level, wear rate and surface finish of the disc specimens can all be continuously monitored.

From the results presented, there appears to be little correlation between overall vibration levels and roughness parameters. However, this could simply reflect the complex nature of the asperity interaction problem, overall roughness parameters being inadequate to quantify the range of asperity shapes and sizes which influence the problem. Attempts at isolating this range are presented, through parameters such as σ and β^*, but more work is required if ideal bearing and gear tooth flank surface topographies are to be defined using the mini-disc machine and the optical vibration monitor.

9.1.2 Theoretical Structure Function Model

As an aid to resolving the problems mentioned in

Section 9.1.1 above, a theoretical model of a profile structure function has been developed and is presented in Chapter 5. This allows roughness characteristics to be studied on an asperity size basis to identify features of particular sizes that may be causing damage during the wear process.

The model is shown to give promising results on the tests performed. However, its application to the particular problem of bearing and gear tooth contact is limited. This is due to the fact that the asperities on these surfaces are not randomly distributed throughout the height range of the surface.

9.1.3 Three-Dimensional Surface Contact Model

A numerical technique has been developed which is capable of simulating the elastic contact of rough surfaces and estimating the effect of plasticity. The method is shown to be capable of creating the real pressure distribution together with a complete picture of local deformation and real contact area on an asperity scale and also to define the shapes of the recovered surfaces after plastic deformation.

An initial attempt has been made to apply the technique to gear contact conditions with the results suggesting that the method shows great promise. The actual asperity scale surface interactions can be studied in both a qualitative and quantitative manner.

9.2 SUGGESTIONS FOR FURTHER WORK

The failure of the experimental programme to identify links between surface roughness and contact vibration was probably due to a lack of sensitivity in the optical vibration monitor. Therefore if this approach to investigating surface finish effects is to be taken further then more work will be required to increase the sensitivity of the vibration monitoring device.

The use of the structure function model appears to be limited by the need to have information about the distribution of asperities within the surface being investigated. The same information can be obtained by other techniques, such as fourier analysis. However, the model may be useful as a method of checking the results obtained by other techniques.

The three-dimensional surface contact simulation technique appears to show the greatest promise as a general method of analysing the conditions found when surfaces come into contact with each other. To increase the capabilities of the technique and fully understand the contact conditions found in gears and bearings a number of enhancements would be useful, namely:-

1) The incorporation of an elastohydrodynamic (ehd) film model between the contacting surfaces. This would enable the effects of different lubricants to be predicted.

2) At present the contact model does not take any account of horizontal tractions and so the important case of sliding contacts with friction cannot be simulated. Although the addition of this feature is not trivial it would make the contact simulations much more general so that they could be used in many more situations.

3) The surface stress distributions calculated from the two-dimensional contact model (Webster and Sayles, 1986) have been used as input to a main-frame computer finite element program to calculate the sub-surface stress distributions. It has then been possible to identify regions within the bodies where failure by crack growth may be initiated.

Results from the three-dimensional contact simulation model could be treated in the same way, but the model would be greatly enhanced if the sub-surface stress distributions could be calculated directly within the program itself. This would allow the full contact conditions to be found in one step rather than having to transfer data between computer systems to obtain full results.

APPENDIX A

THE CONSTRUCTION OF EQUATION 5.6c

THE CONSTRUCTION OF EQUATION 5.6c

The construction of equation 5.6c from Section 5.2 of the main text is shown below for a single asperity, where the delay length τ, and asperity half width a, are within the limits $\tau<a$. The working for other values of τ and a are similar.

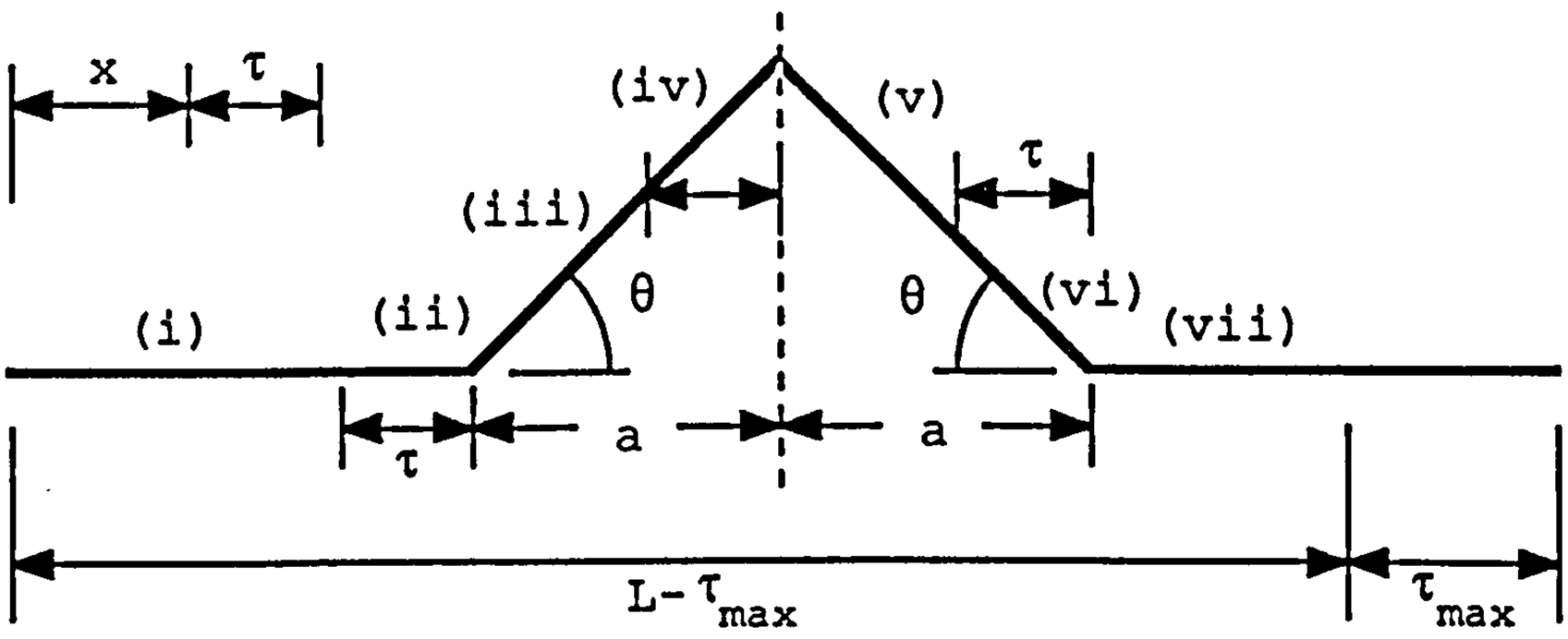


Figure A.1 The possible regimes

The structure function for the situation of Figure A.1 is built up from equations 5.1 and 5.5 from the main text, namely:-

$$S(\tau) = E\left\{\left[z(x) - z(x+\tau)\right]^2\right\} \qquad (5.1)$$

and the conditional expectation, that:-

$$E\left\{\begin{matrix}\text{some}\\\text{event}\end{matrix}\right\} = \sum_{\substack{\text{ALL STATES}\\\text{OF EVENT}}} E\left\{\begin{matrix}\text{PARTICULAR}\\\text{STATES OF}\\\text{EVENT}\end{matrix}\right\} \times \begin{matrix}\text{PROBABILITY}\\\text{OF BEING IN}\\\text{THE PARTICULAR}\\\text{STATE}\end{matrix} \qquad (5.5)$$

If the profile is broken down into the states or regimes of Figure A.1, the expectation of each regime and the probability of being in that regime can be computed. These can then be summed to form the structure function. The particular regime is dictated by the value of x as it varies from zero to $(L-\tau_{max})$. Thus the probability of being in any particular regime is just the length of that regime as a proportion of $(L-\tau_{max})$.

Taking each regime (i to vii) in turn:-

i) In this regime the expectation value is zero so it cannot contribute to the structure function.

ii) The probability of being in this region is,

$$P_{ii} = \frac{\tau}{(L-\tau_{max})}$$

and the expectation, $E_{ii} = \frac{1}{\tau}\int_0^{\tau} l^2\theta^2 dl$

$$E_{ii} = \frac{\tau^2\theta^2}{3}$$

Note: It is assumed that θ is a small angle.

iii) Probability, $P_{iii} = \frac{(a-\tau)}{(L-\tau_{max})}$

Expectation, $E_{iii} = \tau^2\theta^2$

iv) Probability, $P_{iv} = \frac{\tau}{(L-\tau_{max})}$

Expectation, $E_{iv} = \frac{1}{\tau}\int_0^{\tau} \theta^2(\tau - 2l)^2 dl$

$$E_{iv} = \frac{\tau^2\theta^2}{3}$$

v) Probability, $P_v = \frac{(a-\tau)}{(L-\tau_{max})}$

Expectation, $E_v = \tau^2\theta^2$

vi) Probability, $P_{vi} = \frac{\tau}{(L-\tau_{max})}$

Expectation, $E_{vi} = \frac{1}{\tau}\int_0^{\tau} l^2\theta^2 dl$

$$E_{vi} = \frac{\tau^2\theta^2}{3}$$

vii) As in (i) $E_{vii} = 0$ so there is no contribution.

These are combined to form the structure function:-

$$S(\tau) = \frac{\tau^2\theta^2}{3} \times \frac{\tau^2}{(L-\tau_{max})} + \frac{\tau^2\theta^2(a-\tau)}{(L-\tau_{max})} + \frac{\tau^2\theta^2}{3} \times \frac{\tau}{(L-\tau_{max})}$$

$$+ \frac{\tau^2\theta^2(a-\tau)}{(L-\tau_{max})} + \frac{\tau^2\theta^2}{3} \times \frac{\tau}{(L-\tau_{max})}$$

which simplifies to give equation 5.6c

$$S(\tau) = \frac{\theta^2\tau^2(2a-\tau)}{(L-\tau_{max})} \qquad \text{for } \tau < 2a \qquad (5.6c)$$

APPENDIX B

A Method of Three-Dimensional Topography Measurement and Analysis on Arcuate Surfaces

M.N. Webster, M.A. West and R.S. Sayles

A paper presented at:
The Third International Conference on Metrology and
Properties of Engineering surfaces,
Teesside Polytechnic,
England,
April 1985

Published in Wear, Volume 109 (1986)

- 184 -

Wear, *109* (1986) 385 - 399

A METHOD OF THREE-DIMENSIONAL TOPOGRAPHY MEASUREMENT AND ANALYSIS ON ARCUATE SURFACES*

M. N. WEBSTER†, M. A. WEST and R. S. SAYLES

Tribology Section, Department of Mechanical Engineering, Imperial College, Exhibition Road, London SW7 2BX (U.K.)

Summary

A computer-automated method of mapping small areas of curved surfaces is described. The technique is shown to provide not only a good visual image of topographical detail but also quantitative information that can be used directly for assessment and analysis.

An example of the use of the technique is presented by means of a case study on the effects of debris indentation on rolling element fatigue. An example of a debris-damaged raceway is used to demonstrate the power of filtering techniques to focus attention on the topography of interest. Results from a deterministic rough surface elastic contact model are also presented. To date, the analysis technique described has been used to investigate the contact of debris-damaged raceways and to study the mechanics of rough surface elastic contact.

1. Introduction

The aim of this paper is twofold: firstly, to describe the design features of a computer-automated mapping system that has the ability to recover surface height information from small areas of curved surfaces and secondly to demonstrate that the quantitative information gained not only supplies the user with good quality pictures of the surface but can also be used directly to study fundamental problems involving topography. The second point is illustrated by an example of how the mapping device has been used to study the problems of fatigue life reduction due to debris damage in rolling contact bearings [1].

Many of the problems associated with surface topography require quantitative information in three dimensions. There are a number of existing

*Paper presented at the Third International Conference on Metrology and Properties of Engineering Surfaces, Teesside Polytechnic, Teesside, U.K., April 10 - 12, 1985.

†Present address: Tribology Group, Shell Research Ltd., Thornton Research Centre, P.O. Box 1, Chester CM1 3SH, U.K.

0043-1648/86/$3.50

© Elsevier Sequoia/Printed in The Netherlands

techniques that can be used to study three-dimensional surfaces including optical techniques, holography and, perhaps the most popular, scanning electron microscopy (SEM).

Conventional stylus instruments are confined to measuring two-dimensional profiles, essentially measuring a surface section. As the stylus is only capable of moving in a line, a scanning technique is required to achieve a three-dimensional assessment. Most experimenters prefer a raster scan of parallel tracks although spiral and polar scans have been attempted [2 - 6]. Thus suitably adapted stylus instruments offer a practical method of measuring three-dimensional topography.

The main problems and limitations of such systems are the provision of a datum, a capability to store and process large quantities of data in a reasonable length of time and the achievement of a common origin for the individual profiles that make up the raster scan.

Following the pioneering work of Williamson [2], a fully automated technique for mapping three-dimensional flat surfaces was devised by Sayles [7]. The datum plane was provided by an aerostatic bearing. In principle this technique removes the effect of any short wavelength roughness that exists on the bearing surfaces. The experimental data collected from the rig were used to compare the statistical analysis of Longuet-Higgins [8] and Nayak [9] against the real surface results [10].

Many problems in tribology involve mechanical components that have one or two radii of curvature, *e.g.* gears and bearings. A method of measuring the three-dimensional topography of such components is useful both as a visual aid and as a method of obtaining numerical data. George and Radcliffe [11] have used an automated stepping rig to study the topography of a number of curved surfaces and, in particular, to study the volume loss in wear scars.

2. The three-dimensional mapping system

2.1. Mechanical design

The system adopted follows the conventional practice of recording closely spaced parallel profiles to achieve a three-dimensional surface representation. The main component of the system is a device for rotating the specimen a known amount at the completion of each stylus traverse whilst maintaining an accurate radial datum: The stylus re-set and start, the data collection and the specimen rotation are controlled through various interfaces by an LSI 11/23 microcomputer system.

One of the major problems associated with surface mapping is the provision of an accurate datum. In the case of nominally flat surfaces the datum is a plane, whereas in the case of curved surfaces the datum is an axis about which the specimen revolves. The system adopted here consists of a main shaft controlled by a stepping motor on the end of which is a precision chuck assembly. Figure 1 represents a cut-away section showing details of

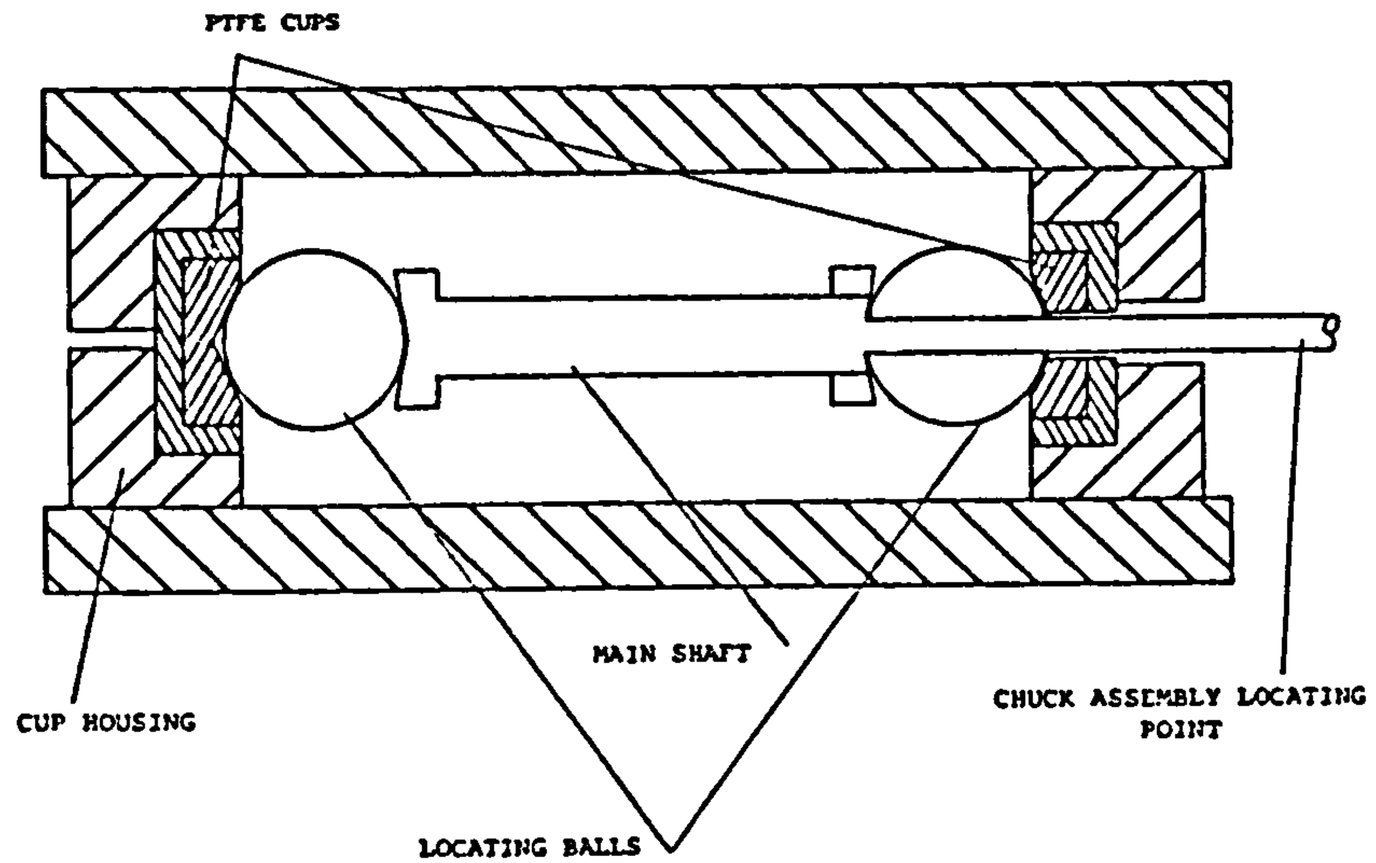


Fig. 1. A vertical section through the stepping rig assembly showing the method of locating the rotating main shaft.

the method of locating the main shaft of the stepping rig by means of two highly finished balls of $1\frac{1}{4}$ in diameter. The balls run on two polytetrafluoroethylene (PTFE) cups which are anchored in housings at either end of the rig. The shaft extends through a spark-eroded hole in one of the balls and clearance cut-outs in the housing, thus providing a locating point for the chuck assembly.

The principle behind this design is that of elastic conformity [12]. Essentially the PTFE cups deform elastically around the local roughness present on the locating balls and average-out any inperfections in the ball topography. Figure 2 shows the elastic deformations that can occur in such a contact region when a rough steel surface contacts a compliant smooth cylinder. The plot is the output from a deterministic numerical contact model that can be used to study the contact mechanics of rough surfaces. The derivation of this model is contained in ref. 13. Figure 2 demonstrates the conforming nature of the PTFE surface and the consequential reductions in displacement perturbations caused by asperity contact; the effect is analogous with the output of a low pass filter.

The specimen is mounted on a magnetic chuck assembly which is located on the main shaft extension. The chuck assembly consists of an outer shell, which houses various levelling devices, and an inner assembly on which the specimen is held in place by small magnets. The inner assembly is located against three point contacts by tension springs anchored behind the back face of the outer shell. The inner assembly may be finely adjusted in two orthogonal directions and independently tilted about two orthogonal

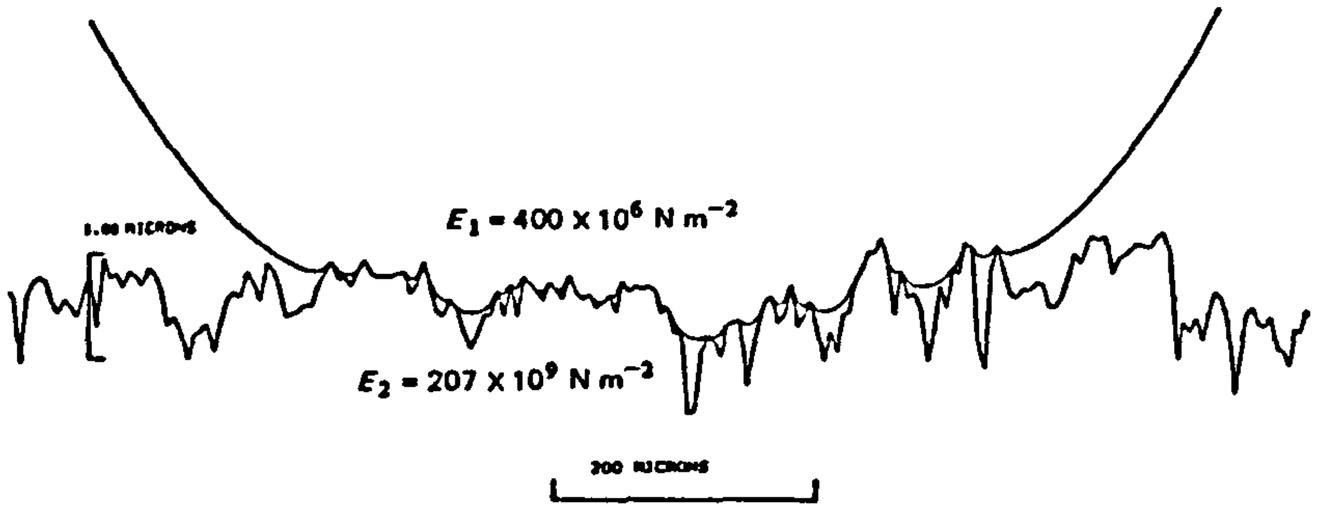


Fig. 2. The surface displacement output from the elastic contact model. The plot demonstrates how the more flexible (upper) body elastically conforms around the rough surface (radius of roller, 15.9 mm; roughness R_a of profile (unfiltered), 0.281 μm; total normal load, 1.41×10^3 N m^{-1}).

axes. Thus it is possible to set up a specimen axis to be concentric with and parallel to the axis of rotation.

The rotation of the shaft and chuck is controlled by a stepping motor with 200 steps per revolution. It is necessary to reduce this to an angular rotation, of the shaft, of 12 seconds of arc per step in order to resolve the fine-scale roughness perpendicular to the stylus tracking direction. This is achieved by means of a simple lever arm arrangement operated by a micrometer thread of 0.5 mm pitch acting as a lead screw. The system is thus limited to mapping over small areas as there is an error associated with converting the linear motion of the lead screw into a rotary motion. The error is surprisingly small, a theoretical positional accuracy of 1% being maintained over a $\pm 5°$ movement of the main shaft.

The automatic stylus re-set and start and the stylus lift operations are carried out by two small d.c. geared servomotors normally used in radio-controlled models. One servo is attached, via a spring link, to the stylus return lever and is used to start the stylus traversing mechanism. The other is used, through a reducing lever and a fine wire, to raise the stylus a fraction of a millimetre clear of the specimen surface to avoid surface damage and lateral stylus forces whilst the specimen is being rotated. The servos are mounted on Perspex plates attached to the Talysurf gearbox assembly.

2.2. Computer interfaces

The entire mapping sequence is controlled by an LSI 11/23 microcomputer. The data logging is achieved using the computer's analogue-to-digital converter. The communication to the stepping motor and d.c. servo is through a parallel input–output board that is directly connected to the computer's data bus. This device allows the input and output of a TTL 0 or 1 on up to 16 independent channels under direct program control. The channels in use are connected to an interface panel which is located on top of the computer and which conditions the signals to the correct form for the

- 186 -

various motor interfaces. A further interface, to control the stepping motor, and a power supply is located with the Talysurf equipment on a seismic table.

The problem of providing a common origin for the profiles has been solved by using a linear-variable differential transducer. The transducer works on the inductive principle; a primary winding is supplied with an oscillating voltage which induces a voltage in two secondary windings. The ratio of the voltages induced in the two secondary windings is governed by the permeability of the gap between them and the primary winding. The permeability is changed by moving an iron core between the primary and secondary windings; thus the difference in voltage between the secondary windings is a function of the position of the iron core. The iron core is located on a yoke which is attached to the Talysurf stylus nose piece. The output signal from the linear-variable differential transducer is demodulated and at a pre-set level a signal is sent to the computer.

To obtain coarse adjustment of the triggering position the coil body can be moved relative to the iron core; a fine electrical adjustment is provided on the interface panel. The state of the command line to the computer is indicated by a small light-emitting diode and thus the triggering point can be set with a high degree of accuracy. The theoretical repeatability of the triggering mechanism is better than ± 0.01 μm, although in practice mechanical considerations, such as thermal drift, mechanical vibration and electrical noise, reduce this overall repeatability (the present authors' estimate of the accuracy is ± 0.1 μm). Figure 3 shows a set of repeated traverses across the Talysurf calibration block with the stepping mechanism disengaged.

The operation of the servo and stepping motors is controlled by one program which contains a mixture of FORTRAN and MACRO assembler language subroutines. MACRO assembler language is used to access the physical addresses of the data buffer registers and the control and status registers of the interface boards that communicate with the peripheral devices. The

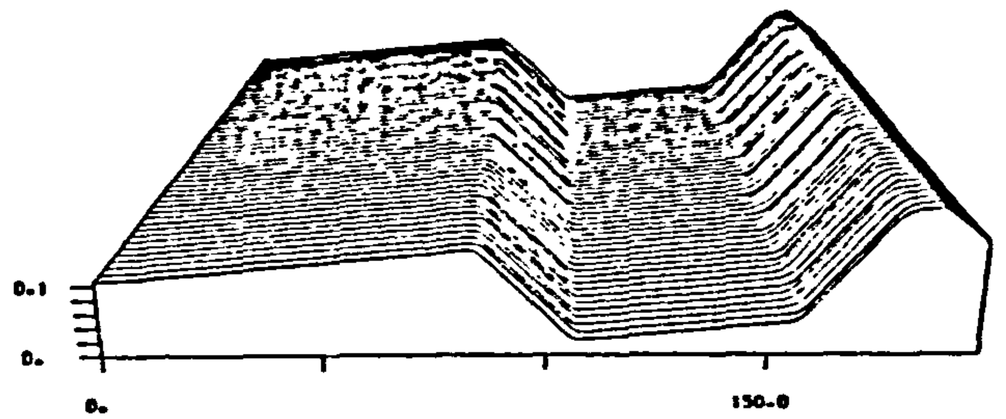


Fig. 3. Results from 50 repeated traverses of the Talysurf calibration groove (length of surface, 198.00 μm; width of surface, 492.02 μm; height range of data, 0.18 μm; filter length (low pass), 0.02 mm). The measurement was taken with the stepping mechanism disengaged and thus tests the triggering device that provides a common origin for the parallel profiles. The results are drawn as a three-dimensional surface plot including perspective.

FORTRAN routines are used to collect input parameters from the visual display unit, to provide information to the operator and to set various constants and counters.

The MACRO subroutines have been restricted to perform simple tasks. The sequence in which these are performed is governed by a FORTRAN subroutine which calls the MACRO routines in the correct order. The operating sequence is lift stylus, step specimen, re-set stylus to start position, lower stylus, start stylus tracking, start data collection at trigger position and write data array to permanent store. This sequence is repeated until the required number of profiles has been recorded.

3. Data analysis

A flow chart of the available processing facilities for the recorded series of Talysurf profiles is shown in Fig. 4. Essentially, the data are first subjected to filtering to remove any unwanted wavelength features and then displayed pictorially as either a three-dimensional view or a contour map. Further analysis can then be carried out on particular topographical features of interest selected from the surface.

3.1. Three-dimensional surface plotting data storage

The number of data points recorded from the surface is only limited by the storage area available on the LSI-11 Winchester disc system. However, for practical purposes, the number is normally limited to a maximum of about 125 000 (250 profiles of 500 points each). The advantage of recording as many points as possible is that a larger area is mapped. This makes the inclusion of features of interest less sensitive to the accurate setting up of the measuring equipment, although there is a penalty to be paid in the time taken for the measurements to be made.

The main plotting program (3DPLOT) has a limited array size of 32 750 data points (corresponding to 131 profiles of 250 points each). In practical terms, the use of any more data than this would make the program unacceptably slow. When the extent of stored data is large it is desirable, and possible, to input only a specified number of points and profiles to the graphics software. From this initial crude representation, areas of interest can be identified and mapped in detail.

3.2. Three-dimensional filtering

Once the data have been read into the program a least-squares mean plane is fitted to remove any general trends. The option is then available to filter the data on an area basis, if required. If the data are plotted without filtering, the picture can often be confused and features of topographical interest obscured (Fig. 5).

The filter routine employed is similar to the moving-average filter commonly used in two-dimensional profilometry, *i.e.* to obtain a low pass

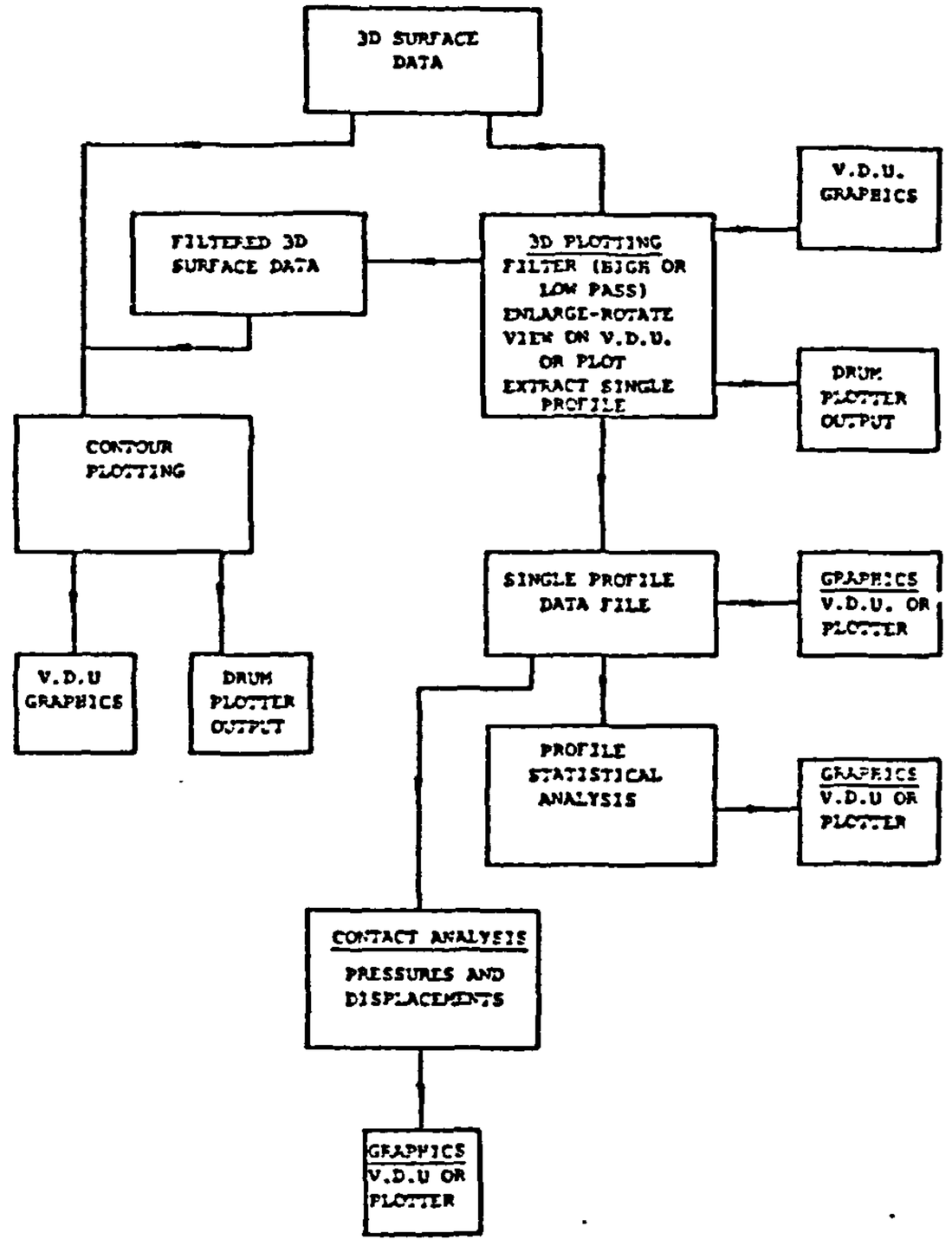


Fig. 4. A flow chart for processing the data files created after the surface measurement. The chart shows how profiles may be removed from the three-dimensional data set for either statistical analysis or contact analysis.

filter, each data point height is replaced by the average height of all the points falling within a filter length centred on that point. To obtain a high pass filter, the low pass result is subtracted from the original height at that point. When this is extended to three dimensions, the filter length becomes a filter box, centred on the point being considered, with sides equal to the filter length. Any point which lies within this box is included in the average used to find the low pass filtered output at the central point. Again, the high pass filter is obtained by subtracting the low pass output from the original data.

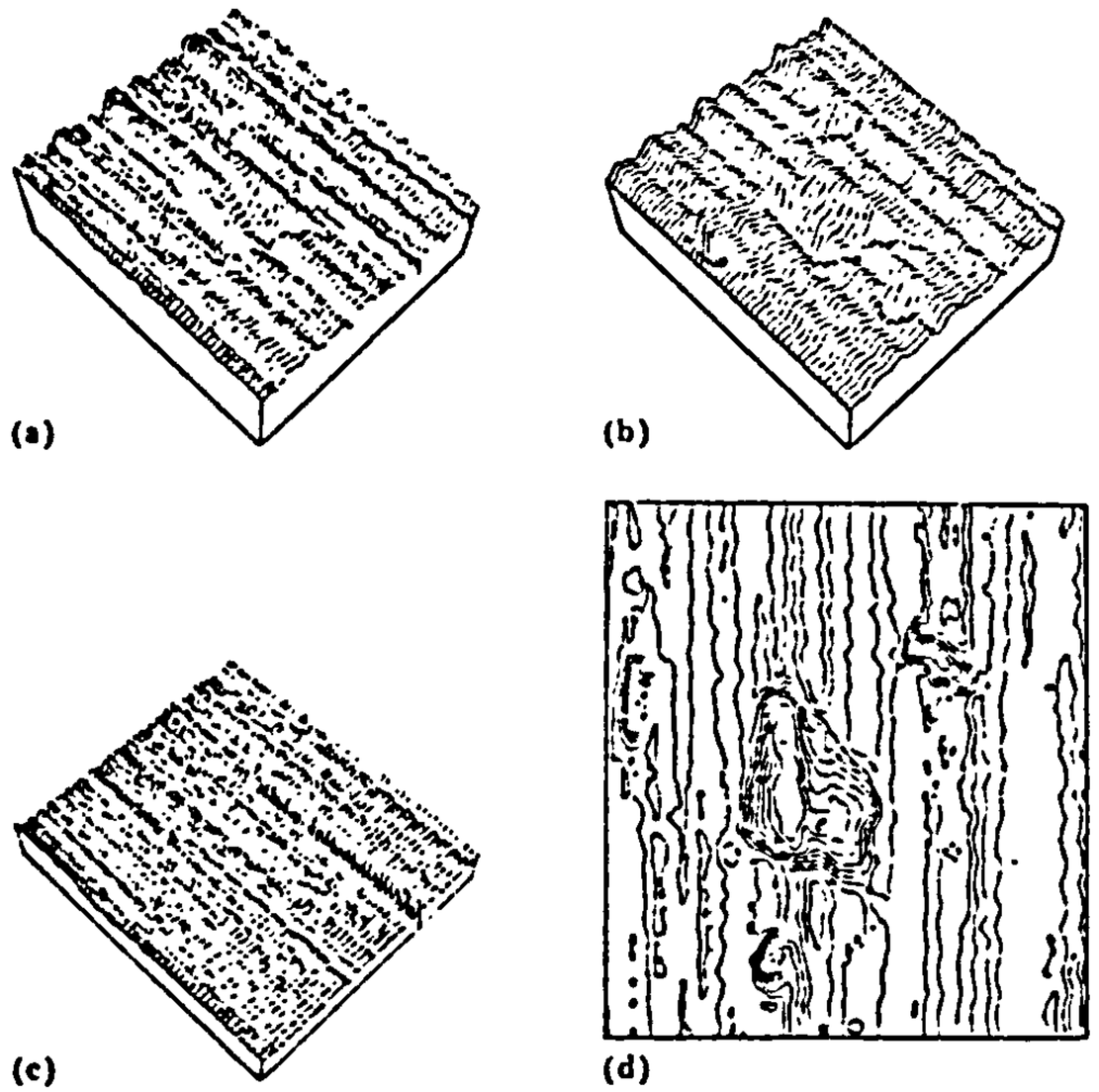


Fig. 5. (a)-(c) A set of three-dimensional views and (d) a contour plot of a debris-damaged bearing outer raceway. The three-dimensional views are the same data plotted (a) unfiltered and (b), (c) after filtering using (b) a 40 μm low pass filter and (c) a 40 μm high pass filter. The contour plot is of the same data showing the planform shape of the debris dents.

The effectiveness of the filtering is clearly demonstrated in Fig. 5, where the unfiltered, high pass filtered and low pass filtered surface of a damaged bearing raceway is shown. In the low pass filtered view the damage caused by debris indentation is enhanced, whilst the high pass filtered view shows the original grinding marks.

The program can also produce pass band filtered data by successively high and low pass filtering the data with appropriate filter lengths.

3.3. Three-dimensional viewing with perspective

The view of the surface, on the monitor screen or Benson drum plotter, can be defined by two angles and a distance. The two angles are the rotation of the surface about the horizontal and vertical axes through its central point and the distance is the range from the viewing position to the centre of the surface. This distance is used to enhance or reduce the effects of perspective. It should be noted that in general because engineering surfaces are relatively flat a true three-dimensional picture would not show any detail at all. To

overcome this the program allows the vertical scale to be adjusted relative to the horizontal scales. For bearing surfaces a vertical scale of approximately 75 times greater than the horizontal usually gives good results. The view obtained is therefore a true three-dimensional picture of a surface 75 times rougher than the actual surface.

The picture is formed by joining the data points as profiles in either or both of the grid directions, with any hidden detail removed. To save time and to aid clarity not all the profiles have to be drawn.

Once a picture has been produced on the monitor screen several options are available.

(i) Another viewing position can be selected and this new view drawn on the screen or plotter.

(ii) Areas of the surface can be isolated interactively so that important topographical features can be enlarged for more detailed analysis.

(iii) Individual profiles (in either of the grid directions) can be isolated to be stored as separate data files. These files can then be used as the input data to the contact model described later, used to assess the size and shape of surface features or used as input data for a separate suite of programs to calculate the statistical parameters of the profile.

3.4. Contour mapping

A separate program with the same data input and filtering routines as 3DPLOT was written to draw contour maps of the surface data (CTMAP). This program can draw up to 20 constant height contours at height intervals supplied by the user.

The contours are generated by searching the height data for adjacent points straddling a particular contour height and then using linear interpolation to find the actual position of the contour. Once a contour has been defined by its positions between data points, the actual contour path is drawn by fitting a smooth cubic spline through the contour points. The output can again be directed to the monitor screen or Benson drum plotter, as required by the operator. A typical example of a contour map is shown in Fig. 5(d).

4. A practical example of the use of the system

4.1. Analysis of debris-damaged rolling element bearings

The mapping system described was principally designed to provide quantitative information on the size and shape of damage caused to bearing and gear surfaces because of particulate contamination in the lubricant supplies. The system has found many other uses over the 2 years that it has been in operation. However, for the purposes of providing an example of its use, we will concentrate on the problem of debris damage.

The work of Bhachu [1] and Sayles and MacPherson [14] provided clear evidence of the correlation between fatigue life and the size of debris

particles allowed into the elastohydrodynamic (EHD) contact zone. Broadly speaking, the finer the filter used to remove particles from the lubricant, the longer is the fatigue life of the bearing. These carefully controlled experiments confirmed what had been observed in practical situations. One of the important conclusions from the work was that the reduction in fatigue life associated with the presence of debris did not rely on its continued presence. Indeed fatigue tests that were started with a relatively coarse lubricant supply filter, which was subsequently swapped for a finer filter, showed that the bearings continued to failure as if no change in filtration had occurred. The implication of such a result is that it is the surface damage caused by debris indentation that promotes early failure.

Although the use of optical and SEM methods can show the presence of debris-induced dents, it is difficult to obtain quantitative measures of the true shapes. Figures 5(a) - 5(c) show three-dimensional views of the plastic deformation caused by debris in the lubricant supply to a bearing outer raceway. The measurement consisted of 50 profiles containing 400 points at a sampling interval of 5 μm. The profiles were spaced 28.5 μm apart. The three views, in which only the first 250 points in each profile are plotted, serve to demonstrate how the choice of filter type and filter area can focus attention on a particular size range of surface features.

The unfiltered data (Fig. 5(a)) show evidence of both the debris damage and the original surface finish. The low pass results (Fig. 5(b)), using a filter area of 0.04 mm × 0.04 mm, show the debris damage with much of the fine surface detail removed. The overall shape of the damaged areas is made much clearer by removing such features. Experience has shown that low pass filtering between 0.01 and 0.05 mm gives the best results for viewing such features. The high pass results (Fig. 5(c)) show the fine surface detail with little evidence of the debris damage associated with longer wavelengths.

One point worthy of note is the continued presence of the fine-scale roughness even in the centre of the plastically deformed region. This is clearly brought to light in the high pass results which show little evidence of any change in the visible topography that remains after filtering. This a good example of what has been termed "asperity persistence" [15 - 17]. The reason asperity persistence occurs appears to involve a complex interaction of a number of effects. Local work hardening of the material around asperities suggests that the maximum load supported by the contacting asperities is somewhat higher than one might expect. Secondly, as the number of contact spots increases, thus becoming more dense, it becomes more difficult to extrude plastically material between the contacts, thus constraining local plastic flow. Thirdly, once the contact load has been removed, elastic recovery may account for some of the topography recovering its original shape. In the case of debris indentation, further effects, due to the presence of a lubricant and an enveloping EHD lubrication pressure distribution, render the problem even more difficult.

A general observation from the dents measured, which are almost exclusively taken from the damaged surfaces produced by the experiments

- 189 -

of Bhachu [1] and Sayles and MacPherson [14], is that dent depth does not vary a great deal. The size range is 10 - 200 μm width across the shoulders and 0.5 - 3.5 μm in depth. A crude calculation of the size range of damaging particles is contained in Sayles and Webster [18].

Figure 5(d) shows the same surface data again presented in the form of contours of equal height. This type of surface presentation is particularly useful if the planform shape of surface features is important. From this plot the four debris dents are clearly visible.

5. Simulation of elastic contact

One of the analysis techniques developed by the present authors is that of the simulation of elastic contact using a numerical method (see Fig. 2 and Section 2.1). The surface input data can be either an individual profile measurement or a profile data file extracted from a three-dimensional measurement. The output results give the surface pressures and displacements for a given total load.

The contact technique has been used to investigate the influence of debris dents on the contact pressures in an EHD zone. A typical set of output plots is demonstrated in Fig. 6. The plots show the "smooth case" and computed results for a smooth roller loaded against a profile containing a number of debris dents. The effect of the debris dents is to introduce pressure peaks local to the dent shoulders. Depending on the size and shape of the dent, the pressure peaks range from 1.5 to 3.5 times the smooth case maximum pressure.

A subsurface stress analysis using a finite element technique reveals that the high pressure peaks give rise to high local shear stresses just below the bearing surface (Fig. 7). Experimental evidence of significantly reduced bearing fatigue lives [1, 14], due to the presence of debris dents, suggests that these local stress concentrations offer a competing fatigue crack initiation site to the subsurface model generally associated with roller bearing failure. Early results from current research appear to justify this hypothesis.

A further example of the use of the elastic contact model is represented in Fig. 8. The model can be used to calculate the area of contact for each load step taken during the solution. If these results are plotted we can obtain graphs of load *versus* area for a particular surface. The results shown in Fig. 8, which are for a smooth roller loaded against a rough abraded surface, demonstrate that the area of contact due to elastic deformation is proportional to the load.

The load *versus* area graph confirms that the real area of elastic contact is only a small proportion of the apparent area. The load area proportionally is in broad agreement with existing statistical models and endorses the hypothesis that Amonton's laws of friction are a consequence of the nature of rough surface contact.

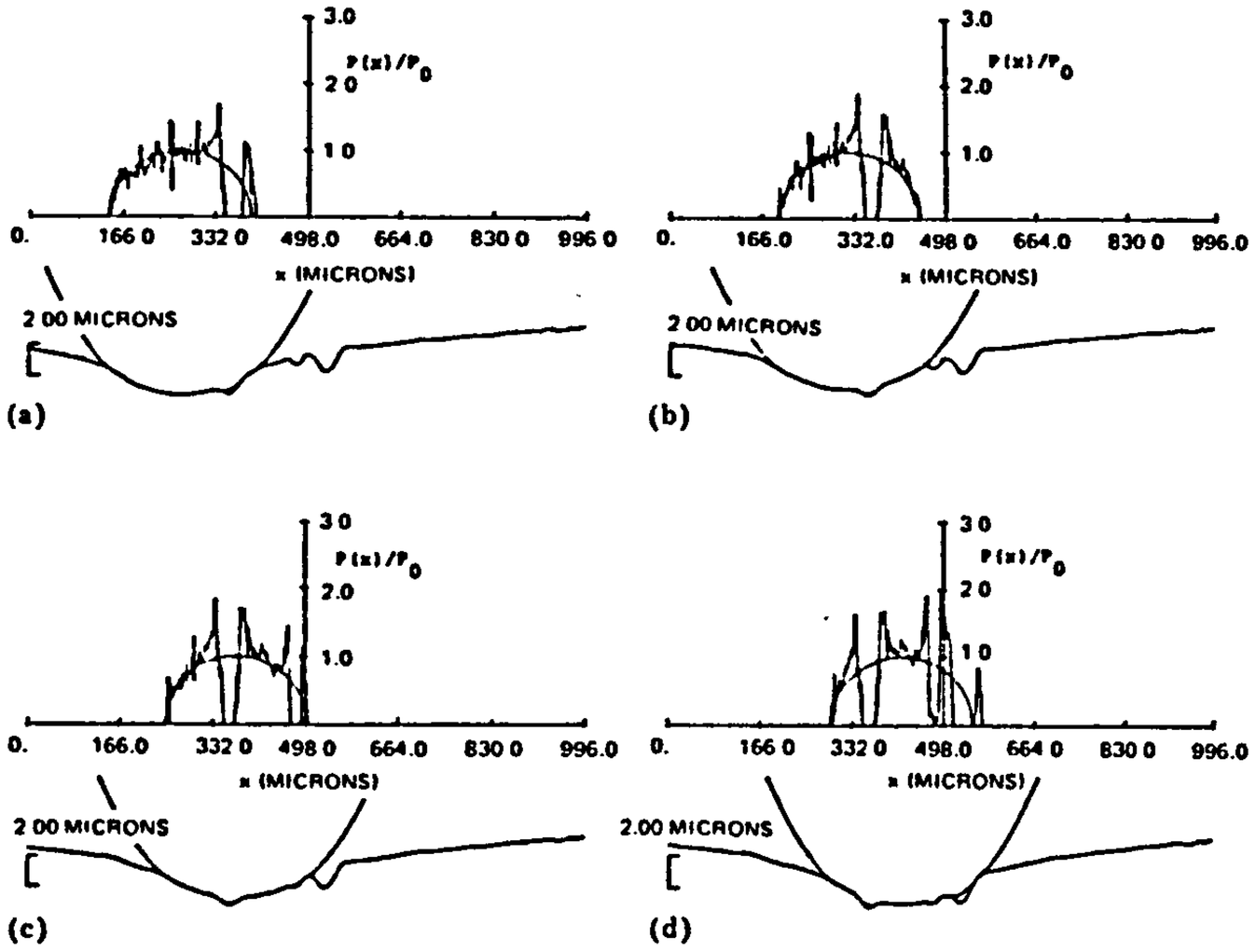


Fig. 6. Pressure and displacement results from the elastic contact model showing the effect of debris dents on the contact between a roller and raceway. The plots show the results for four different roller locations with the smooth hertzian results plotted for comparison. The load and geometry are typical for rolling element bearings (radius of roller, 2.85 mm; $P_0 = 2.55 \times 10^9$ N m^{-2}).

6. Conclusions

A computer-automated method of mapping small areas of curved surfaces, using the parallel tracing technique, is described. The system provides good quality output that may be previewed on a graphics screen prior to plotting. The advantage of such a system is not only in its ability to provide visual images of a surface but also in that it is straightforward to obtain numerical data on local surface features and to use this data for further analysis.

The system has been used to study a wide range of problems associated with surface topography. As an example of its many uses, results are presented on debris indentation in rolling element bearings. These results are then used to study the influence of such damage on bearing fatigue via an elastic contact analysis technique, which we have developed as a general analytical tool to study the contact mechanics of rough surfaces. The

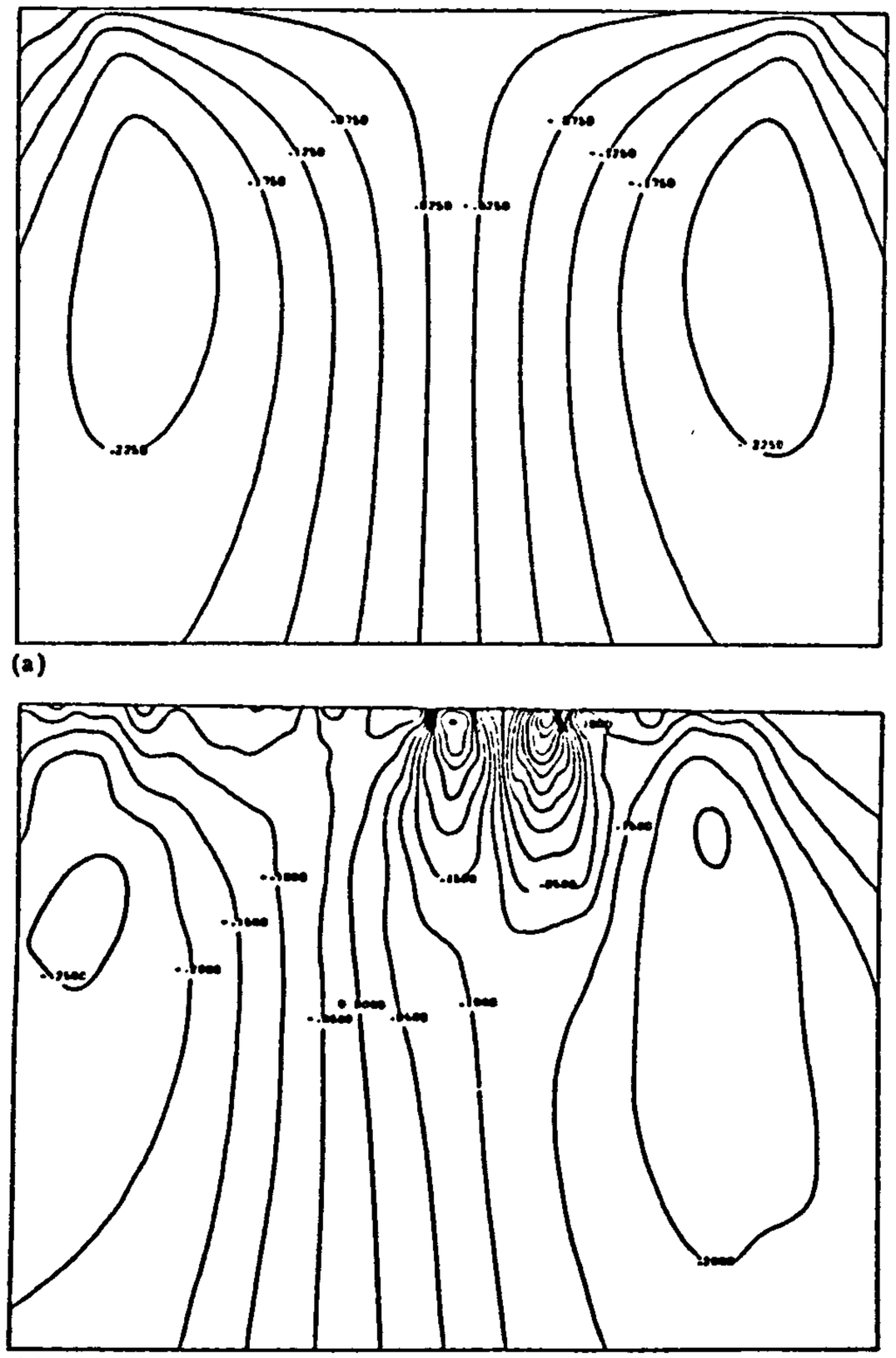


Fig. 7. Contour plots of the subsurface orthogonal shear stress normalized by the theoretical maximum surface pressure (*i.e.* τ_{xy}/P_0). (a) The smooth case contours show the subsurface maximum shear associated with subsurface fatigue crack initiation. (b) The debris dent case results correspond to the load case shown in Fig. 6(b). The higher shear stress concentration near the surface corresponding to the dent shoulders should be noted.

example presented here demonstrates how real height data gathered from digitized surface measurements can be used to study a range of problems associated with the elastic contact of real rough surfaces.

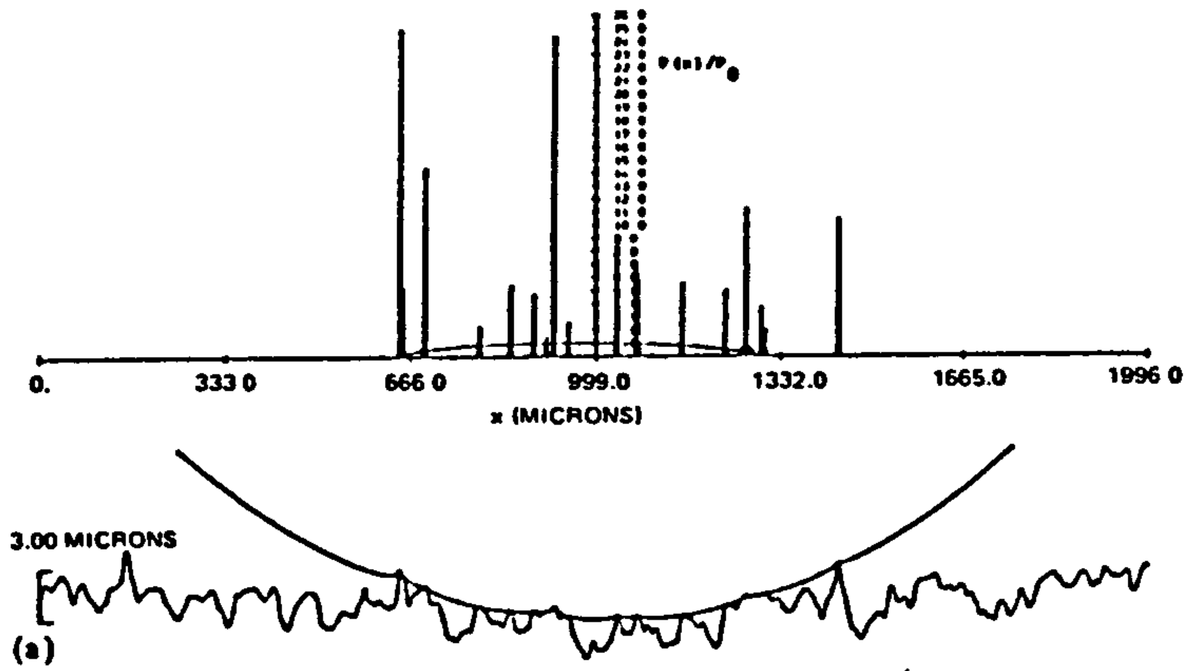


Fig. 8. (a) Results from an elastic contact analysis of a smooth roller loaded against a rough surface (radius of roller, 25 mm; roughness R_a of profile (unfiltered), 0.742 μm; $E_1 = E_2 = 207 \times 10^9$ N m^{-2}; total normal load, 352×10^3 N m^{-1}; $P_0 = 712 \times 10^6$ N m^{-2}). (b) The load *vs.* area curve shows a least-squares line plotted through the calculated points (radius of roller, 25 mm; $E_1 = E_2 = 207 \times 10^9$ N m^{-2}; maximum $P_0 = 712 \times 10^6$ N m^{-2}; slope of line, 0.243×10^3).

Acknowledgments

The authors are gratefully indebted to the U.S. Navy's Office of Naval Research and Westland Helicopters Ltd. for supporting this work.

References

1 R. S. Bhachu, The influence of debris on rolling contact fatigue life, *Ph.D. Thesis*, University of London, August 1980.
2 J. B. P. Williamson, Microtopography of surfaces, *Proc., Inst. Mech. Eng., London, Part 3K, 182* (1967 - 1968) 27 - 30.
3 J. Peklenik and M. Kubo, A basic study of a three-dimensional assessment of the surface generated in a manufacturing process, *Ann. CIRP, 16* (1968) 257 - 265.
4 M. Kubo and J. Peklenik, An analysis of micro-geometrical isotropy for random surface structures, *Ann. CIRP, 16* (1968) 235 - 242.
5 J. J. Grieve, H. Kaliszer and G. W. Rowe, A normal wear process examined by measurements of surface topography, *Ann. CIRP, 18* (1970) 585 - 592.
6 M. J. Edmonds, A. M. Jones, P. W. O'Callaghan and S. D. Probert, A three-dimensional relocation profilometer stage, *Wear, 43* (1977) 329 - 340.
7 R. S. Sayles, The topography of surfaces, *Ph.D. Thesis*, Teesside Polytechnic, 1977.
8 M. S. Longuet-Higgins, The statistical analysis of a random moving surface, *Philos. Trans. R. Soc., Ser. A, 249* (1957) 321 - 387.
9 P. R. Nayak, Random process model of rough surface, *J. Lubr. Technol., 93* (1971) 378.
10 R. S. Sayles and T. R. Thomas, Measurement of the statistical microgeometry of engineering surfaces, *J. Lubr. Technol., 101* (1979) 409 - 418.
11 A. F. George and S. J. Radcliffe, Automated wear measurement on a computerized profilometer, *Proc. 2nd Int. Conf. on the Metrology and Properties of Engineering Surfaces, Leicester Polytechnic, Leicester, April 14 - 16, 1982*, in *Wear, 83* (1982) 327 - 337.
12 R. S. Sayles, G. M. S. de Silva, J. A. Leather, J. C. Anderson and P. B. MacPherson, Elastic conformity in hertzian contacts, *Tribol. Int., 14* (December 1981) 315.
13 M. N. Webster and R. S. Sayles, A numerical model for the elastic frictionless contact of real rough surfaces, *J. Tribol.*, to be published.
14 R. S. Sayles and P. B. MacPherson, Influence of wear debris on rolling contact fatigue. In J. Hoo (ed.), *ASTM Spec. Tech. Publ. 771*, 1982, pp. 255 - 278 (ASTM, Philadelphia, PA).
15 J. Pullen and J. B. P. Williamson, On the plastic contact of rough surfaces, *Proc. R. Soc. London, Ser. A, 327* (1972) 159 - 173.
16 T. H. C. Childs, The persistence of asperities in indentation experiments, *Wear, 25* (1973) 3 - 16.
17 T. H. C. Childs, The persistence of roughness between surfaces in static contact, *Proc. R. Soc. London, Ser. A, 353* (1977) 35 - 53.
18 R. S. Sayles and M. N. Webster, The characteristics of surface roughness important to gear and rolling bearing problems, *Proc. AGARD Conf. on Gear and Power Transmission Systems for Helicopters and Turboprops, 1985*, Publication 369, Advisory Group for Aerospace Research and Development, Neuilly sur Seine, 1985.

APPENDIX C

A New Method for Rough Surface Contact Analysis

M.A. West, M.N. Webster and R.S. Sayles

A paper presented at:

'Tribology - Friction, Lubrication and Wear,
Fifty Years On',
Institution of Mechanical Engineers,
London, England,
July 1987

Published in the Proc. Inst. Mech. Engrs, (1987-5)

C161/87

A new method for rough surface contact analysis

M A WEST, MA, M N WEBSTER, MSc, PhD and R S SAYLES, BSc, PhD
Department of Mechanical Engineering, Imperial College of Science and Technology, London

SYNOPSIS. A numerical technique capable of performing elastic contact analysis of measured rough surface profiles is presented. Results are shown to compare well with the theoretical method of Bush and his colleagues for a set of abraided surface specimens produced to cover a range of σ and β^* in a controlled manner. The technique is shown to reveal the real pressure distribution within a cylindrical Hertzian contact together with information on individual contact spots and how this changes as deformation proceeds. An interesting result of this approach is seen in the way neighbouring asperities interact, and a degree of local conformity obtained. Such results are unobtainable through the application of statistical models.

1. INTRODUCTION

Traditionally stochastic techniques have been used to analyse rough surface contact problems. Generally the measured topography is reduced to its probability density functions, and their moments, which are then used to model the real surface by a population of idealised asperity shapes (e.g. spherical or parabolic asperity caps) for which the contact behaviour is well known (1, 2 and 3). Such an approach has a number of shortcomings:

1) It requires a number of assumptions regarding the probability distribution of asperity heights, slopes, curvatures etc.

2) It requires the use of simple asperity shapes.

3) Each discrete contact spot is assumed to deform independently.

4) It cannot predict the deformed shape of the surfaces, particularly for the non-contacting areas.

5) Prior knowledge is required about the size range of topography that is important to the problem.

The advent of small but powerful micro-computers linked to stylus based surface topography measuring instruments has led to the development of deterministic solutions to rough surface contact problems (4). The deterministic solution provided by numerical techniques can offer some advantages over probabilistic models:

1) Since the measured topography is used directly no probabilistic assumptions need to be made and an asperity model is not required.

2) The matrix methods used take full account of the interaction of deformations from all the contact points.

3) It is possible to predict the elastic strains at all points and thus calculate the deformed geometry for both bodies.

4) Individual asperity pressures are calculated values not expected values.

Despite these advantages, a number of difficulties still remain. A decision about the sampling interval and sample length is required, thus a prior knowledge of the spacial range of topography of interest is still necessary. Restrictions associated with the original measurement still apply e.g. finite sampling interval and stylus size, response of the instrument, vertical resolution etc. In an effort to establish the differences between stochastic and numerical rough surface contact models a comparison between our own model and that of Bush et al (5) for strongly anisotropic rough surfaces, forms part of our paper.

2 NUMERICAL MODEL AND ITS PREVIOUS APPLICATION

A full description of the numerical model has been previously reported (6). Essentially a previously recorded rough surface is represented by a number of surface contact elements. Over each element a uniform pressure distribution is assumed. Similar models have been reported which have used contact elements having a triangular pressure distribution (7). For a given overlap the individual elemental pressures are calculated from the matrix of influence coefficients and the estimated surface displacements. Iteration is required until a solution is found in which all elemental pressures are positive.

Previous Applications

The technique can be used as a general contact mechanics tool, but to date our work has concentrated on its application to surface and sub-surface stress predictions within nominally

C161/87 © IMechE 1987

Hertzian contacts in the presence of rough surfaces and surfaces with indentations due to rolled-in wear debris.

Figure 1a shows a typical debris indentation measured by a Talysurf scanning method (4). The numerical contact technique is then applied to a profile selected from such a dent to produce a quantitative picture of the contact geometry and the surface stress distribution (Figure 1b). Figure 2 shows an extension of the technique to predict sub-surface stresses using the surface stress distribution as the boundary condition in a finite element analysis. Note the high shear stress regions caused by the shoulders of the indentations, and how they approach the surface more closely than for a smooth contact. Such results have been used in modelling early life fatigue failures experienced in rolling element bearings (8 and 9).

3 APPLICATION TO A SET OF REFERENCE SURFACES

Following the introduction of the plasticity index concept by Greenwood and Williamson (1), interest in contact mechanics has broadened in recent years to take account of both height and spatial parameters of roughness. The significance of such parameters has been the subject of a considerable amount of theoretical work, but very little experimental confirmation has been attempted, or achieved. One exception to this is available through the work of Hirst and Hollander (10) who related the transition from safe to unsafe sliding, under boundary lubrication, to the rms roughness σ and the correlation length β^*. Their results were explained in terms of the ratio of σ/β^*, the topography parameters within the plasticity index.

To investigate rough surface contact behaviour with varying topography in an ordered way, it is necessary to choose from a confusingly large number of available surface parameters, each of which could be varied. On the basis of tribological applications research, such as that of Hirst and Hollander, it was decided to use a set of surface specimens which had a regular and controlled variation in the parameters σ and β^* to test the model.

Production of surface specimens

A series of 10 reference surfaces were produced using varying grades of silicon carbide abrasive papers in a similar way to the method described by Hirst and Hollander (10). The specimens were produced with a one-dimensional lay having regular variations in σ and β^* around the values typically found on engine cylinder bores. The production of these specimens and the many problems involved is described elsewhere (11).

The nominal values of σ and β^* for each of the 10 specimens, numbered 1 to 10 are shown in Table 1. The triangular appearance of the σ and β^* range was due to an interesting physical restriction of using abrasive methods. It was not possible to produce surfaces of high σ to β^* ratio, physically a spikey surface. This is not surprising as to produce surfaces of low β^*, papers of short grit spacing (high grade number) are required. The greatest roughness that can be obtained with a virgin paper of high grade number falls well below the higher values of σ indicated in the Table. The limiting value of σ/β^* appears to be between about 0.08 and 0.1, higher values being unobtainable by abrasive paper methods.

					σ
			10		1.2μm
		8	9		0.9μm
	5	6	7		0.6μm
1	2	3	4		0.3μm
5μm	10μm	15μm	20μm	β^*	

Table 1

Measurement of the reference surfaces

The reference surfaces were measured using a Talysurf 4 with standard stylus, the recorded profiles being analysed on an LSI 11/23 based micro-computer. Each profile consisted of 1001 digitised points at a spacing of 2μm, thus giving a profile length of 2mm. 10 such profiles were recorded from each reference surface and the σ and β^* values shown in Table 2 were based on an ensemble average of the 10 measurements.

To obtain the required parameters the profiles were subjected to a high-pass, moving average spline filter with a 0.2mm filter length cut-off to remove any waviness. Figure 3 shows the profile, its height distribution, slope distribution and correlation function for specimen number 6, which in character is typical of all the specimens.

© IMechE 1987 C161/87

Table 2 AVERAGE σ AND β* RESULTS FOR REFERENCE SURFACES

Reference surface	Average σ (μm)	Standard Deviation of σ Measurements (%)	Average β* (μm)	Standard Deviation of β* Measurements (%)
1	0.294	6	6.3	19
2	0.327	12	11.4	18
3	0.286	12	14.1	33
4	0.296	14	21.2	18
5	0.555	7	10.8	22
6	0.583	11	16.4	22
7	0.611	17	21.8	23
8	1.001	10	16.1	20
9	1.058	14	24.5	26
10	1.111	13	21.0	34

Table 3 Surface Parameters used in Contact Analysis

Reference Surface	RMS Roughness σ(μm)	Correlation Length β* (μm)	Moment m_2
1	0.308	7	0.0204
2	0.363	13	0.0132
3	0.320	12	0.0104
4	0.506	32	0.0054
5	0.644	15	0.0339
6	0.885	35	0.0272
7	0.667	39	0.0216
8	0.967	14	0.0543
9	1.558	44	0.0493
10	1.328	24	0.0492

C161/87 © IMechE 1987

4 CONTACT ANALYSIS

The surfaces described in section 3 were used as input data for the numerical technique in a comparison with existing theories of the elastic contact of plane surfaces. The numerical technique simulates a cylinder on plane contact but at a large cylinder radius the results should not differ appreciably from plane/plane contact, Lo (12).

The numerical technique is two-dimensional and is therefore applicable to a highly anisotropic surface. An existing theory for this situation is available through the work of Bush and his colleagues (5). His overall elastic contact model is involved and cannot be applied easily, but an asymptotic approximation valid for high separations is tractable and was used to compare the numerical and theoretical techniques. The main surface parameter of this model is the second moment of the power spectral density function, m_2 measured across the lay. Another theory of surface contact is that due to Onions and Archard (13). This uses the surface profile parameters σ and β^* but unfortunately only applies to isotropic surfaces. However, a comparison with the numerical technique was attempted.

Before performing the contact analysis each surface described in section 3 was remeasured at a 1µm sampling interval with a high resolution stylus to include as many of the small scale features as possible in the numerical technique and its comprison with theory. The profiles were subjected to a high-pass, moving average spline filter with a 0.9mm filter length (the Hertzian contact width). Table 3 gives the σ, β^* and m_2 values for the ten profiles.

The following surface contact analyses were carried out:-

1 Simulation of a 100mm diameter smooth cylinder being loaded onto the centre of each measured profile, with the contact load and real contact area being recorded in approximately 60 steps up to a total load of 0.36N/µm.

2 A repeatability test was carried out by simulating the contact of a 100mm diameter cylinder being loaded onto the profile of surface 6 at five different positions separated by 500µm each. The range of the results should give some idea of the repeatability of the tests carried out in 1 above.

3 A simulation of 12 different sized cylinders, ranging from 50µm to 400mm in diameter being loaded on to the centre of the profile of surface 6. Thus indicating the effect of cylinder size on the results.

4 By only using some of the data points from the profile of surface 6, the effect of the use of different sampling intervals on the contact results for a cylinder of 100mm diameter was studied. Sixteen analyses with sampling intervals in the range of 1µm to 200µm were carried out.

5 DISCUSSION OF RESULTS

Figure 4 shows the results of the contact analyses for the ten surfaces. Approximate linearity of the normal load against real contact area is achieved over all surfaces. Figure 5 shows the pressure distribution and deformed cylinder and profile shapes for the analysis of surface 6 at its final load of $0.36N/\mu m^2$.

The gradients of the normal load against real contact area of Figure 4 give the numerical force/contact area results to be compared to those using the theoretical models.

In order to apply the Bush model to a 3-dimensional contact it is necessary to also define the parameter λ which is the ratio of the second spectral moment parallel to and across the lay. The value of λ for the surfaces used was 0.3. However as the numerical contact model is 2-dimensional a value of $\lambda = 0.01$, indicating a very high degree of anisotropy, was used in the comparison (approximately the practical limit of the Bush model). Thus for the conditions described, the theory of Bush can be stated as, $F/A = 0.214\ E^*\sqrt{m_2}$. Figure 6 shows the force/contact area results from the numerical model plotted against $E^*\sqrt{m_2}$. The straight line gives the result from Bush's analysis. Surprisingly good agreement is evident for all but surface 10, which is the roughest surface having the least number of contact points.

Figure 7 shows the force/contact area results plotted against $E^*\sigma/\beta^*$. The straight line gives the result from the Onions and Archard theory for isotropic surfaces, ie $F/A = 0.3E^*\sigma/\beta^*$. Exact agreement would not be expected due to the anisotropic nature of the numerical model.

The repeatability tests gave good results, with the variation of force/contact area, with differing position of the cylinder, showing a standard deviation of only 3%. The results of Figures 6 and 7 would thus seem to be reliable.

Figure 8 shows the results of the analyses of different sized cylinders. As can be seen, for cylinders of diameter 50mm and above the results are fairly constant. Only with smaller cylinder diameters do the results differ. The 100mm diameter cylinder chosen for the comparison with the theoretical models would seem to be sufficiently large to give agreement with results which are based on a nominally flat contact geometry.

Figure 9 shows the effect of sampling interval on the force/contact area results. As the sampling interval increases the value of the force/contact area for both the numerical model and Bush's model decrease in a similar manner. Thus it would seem possible to carry out the comparison at any sampling interval, as long as the calculation of the spectral moment m_2 and contact analysis are carried out at the same sampling interval. This is in keeping with previous work on applying real surface measurements within the spectral moment techniques, Sayles and Thomas ASME (14).

© IMechE 1987 C161/87

6 CONCLUSIONS

A numerical technique is described which is capable of simulating the elastic contact of rough surfaces. The method is shown to be capable of creating the real pressure distribution together with a complete picture of local deformation and real contact area on an asperity scale. The technique is illustrated by its application to define local surface and sub-surface stresses created around debris indentations on rolling element bearings.

Results are also presented on the force to real-contact-area relationships produced in a nominally Hertzian cylindrical contact on a series of surface specimens manufactured to give a regular and controlled variation in the rms roughness σ and the correlation length β^*. Linearity in load-v-area is demonstrated on all specimens over a wide range of loads, with the proportionality changing in keeping with the values of σ and β^*. For constant σ, an increase in β^* creates a corresponding increase in the real contact area and a subsequent reduction in local asperity stresses. Keeping β^* constant and reducing σ is seen to have a similar effect.

Numerical results of load-v-area proportionality are compared with the random process model for strongly anisotropic surfaces due to Bush and his colleagues. Agreement is surprisingly good.

It is evident from the elastic contact analyses we have presented on abraded surface specimens, that with real materials, where non-linear stress-strain relationships must be imposed, significant local asperity deformation can take place. We have refrained from counting the proportions of elastic and plastic contact we can observe as this would obviously depend on the material involved and the way in which the surface geometry changes during the plastic and elasto-plastic phases of local asperity deformation. These aspects are the subject of our present work with this contact technique.

REFERENCES

1 Greenwood, J.A., and Williamson, J.B.P. "Contact of Nominally Flat Surfaces". Proc. Roy. Soc., Vol A295, pp.300-319, (1966).

2 Nayak, P.R. "Random Process Model of Rough Surfaces". Trans, ASME, J. Lub. Tech. 93F, 398-407 (1971).

3 Bush, A.W. and Gibson, R.D. "The Elastic Contact of a Rough Surface". Wear Vol. 35, pp87-111 (1975).

4 Webster, M.N., West, M.A. and Sayles, R.S. "A Method of 3-Dimensional Topography Measurement and Analysis on Arcunate Surfaces". 3rd. Int. Conference, "Metrology and Properties of Engineering Surfaces", Teeside Polytechnic, England, 1985.

5 Bush, A.W., Gibson, R.D. and Keogh, G.P. "Strongly Anisotropic Rough Surfaces". ASME Journal of Lub. Tech., Jan 1979, Vol. 101, pp 15-20.

6 Webster, M.N. and Sayles, R.S. "A Numerical Model for the Elastic Frictionless Contact of Real Rough Surfaces". Trans ASME Journal of Tribology, Vol. 108 No.3 (1986) pp 314-320.

7 Francis, H.A. "A Finite Surface Element Model for Plane-Strain Elastic Contact", Wear Vol 76, pp 221-245, 1982.

8 Sayles, R.S. and Webster, M.N. "The Characteristics of Surface Roughness Important to Gear and Rolling Bearing Problems", AGARD Conf. Proc. No. 369, Gear and Power Transmission Systems for Helicopters and Turboprops.

9 Webster, M.N., Ioannides, E. and Sayles, R.S. "The Effect of Topographical Defects on the Contact Stress and Fatigue Life in Rolling Element Bearings". Leeds/Lyon Symposium 1985.

10 Hirst, W. and Hollander, A.E. "Surface Finish and Damage in Sliding". Proc. Roy. Soc., Vol A337 (1974).

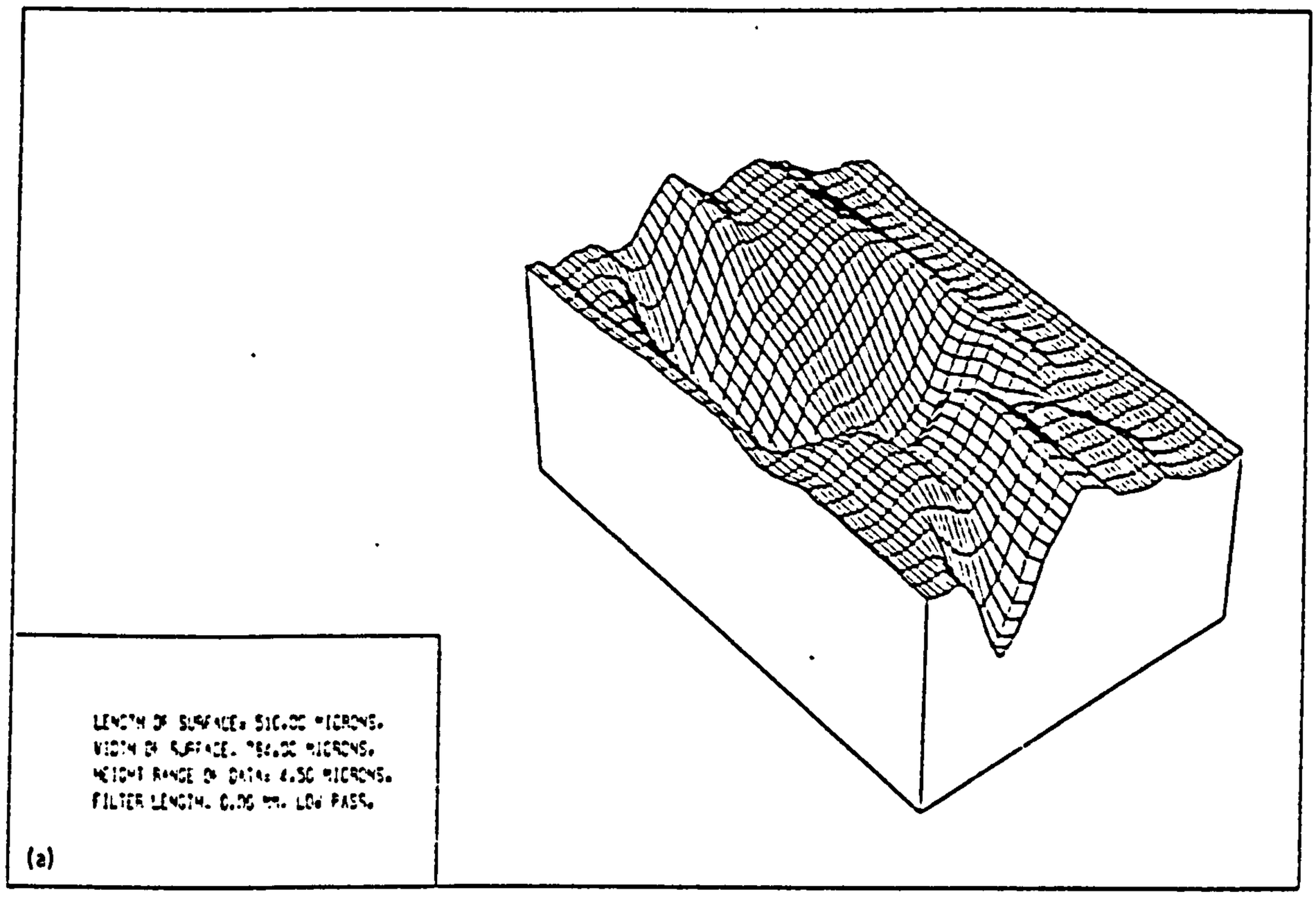
11 Hamer, J.C., West, M.A., and Sayles, R.S. "The Production and Topographical Definition of Reference Surfaces for Calibration of the J.R.C. Flat Optical Profilometer". Imperial College, Tribology Section Report, TS011/85. (1985).

12 Lo, C.C. "Elastic Contact of Rough Cylinders". Int. J. Mech. Sci., Vol 11, pp105-115. (1969)

13 Onions, R.A. and Archard, J.F. "The Contact of Surfaces having a Random Structure", J. Phys. D: Applied Phys., Vol. 6, pp289-304 (1973).

14 Sayles, R.S. and Thomas, T.R. "Measurements of the Statistical Microgeometry of Engineering Surfaces", Trans, ASME Journal of Lub. Tech. Paper 78-Lub-15.

C161/87 © IMechE 1987



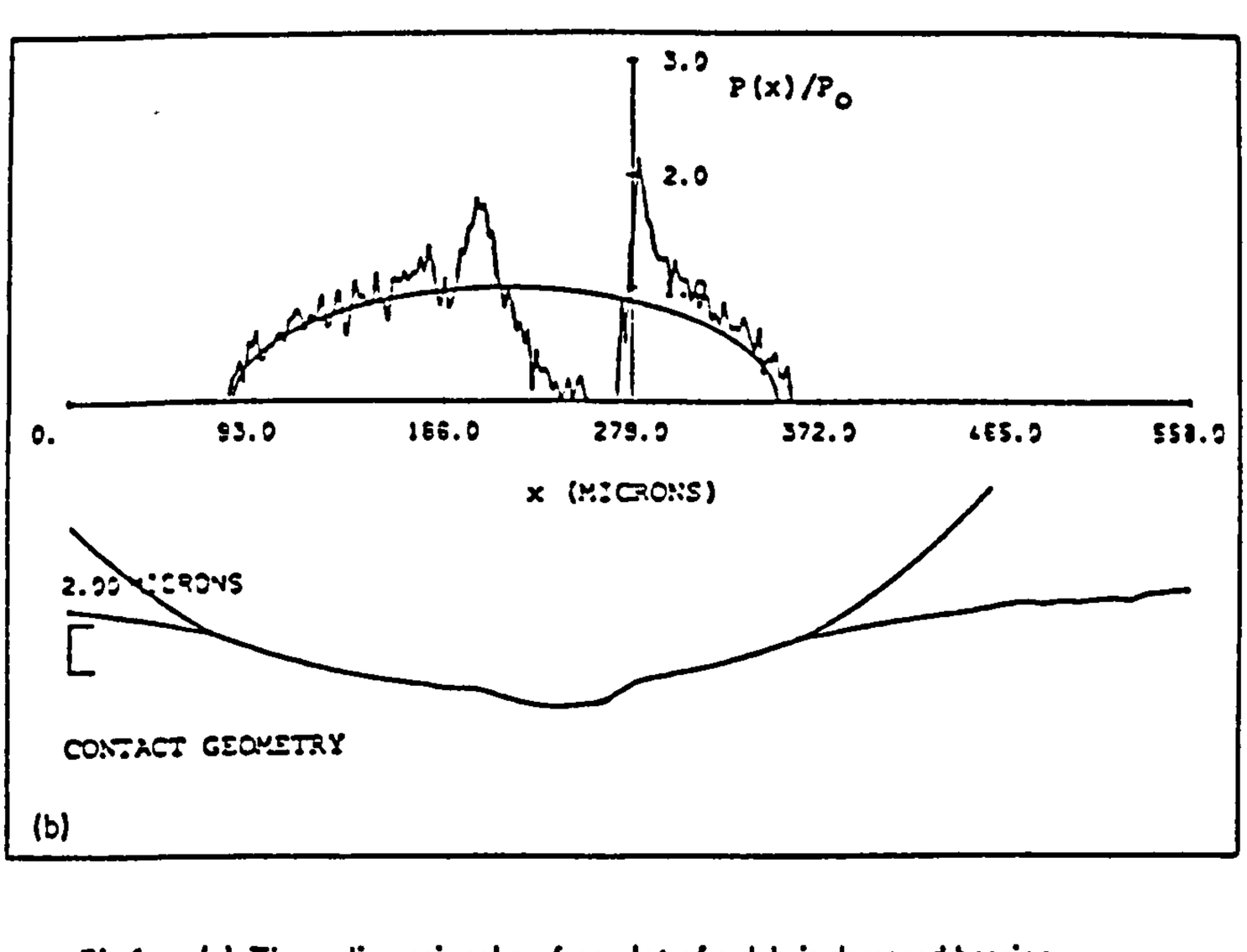


Fig 1 (a) Three-dimensional surface plot of a debris damaged bearing raceway. The dent is approximately 3 μm deep
(b) Numerically calculated pressure distribution and deformed geometry for a smooth cylinder loaded onto a selected profile. The load = 0.7/μm, reduced radius = 2.35 mm, $E^* = 110$ GN/m^2 and $\nu = 0.3$

© IMechE 1987 C161/87

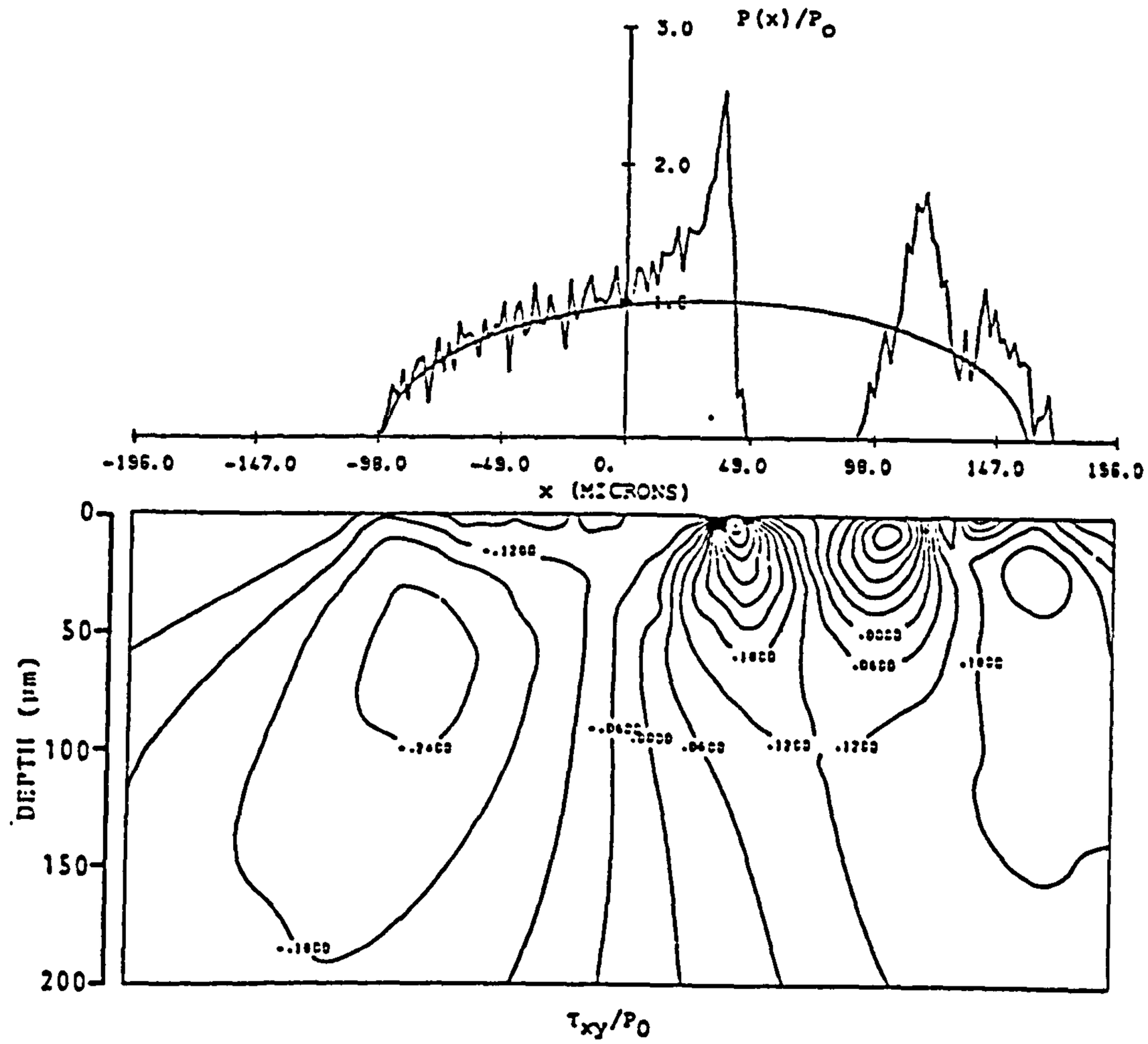


Fig 2 Surface and sub-surface stress distributions for a simulation of a smooth roller passing over a debris dent

C161/87 © IMechE 1987

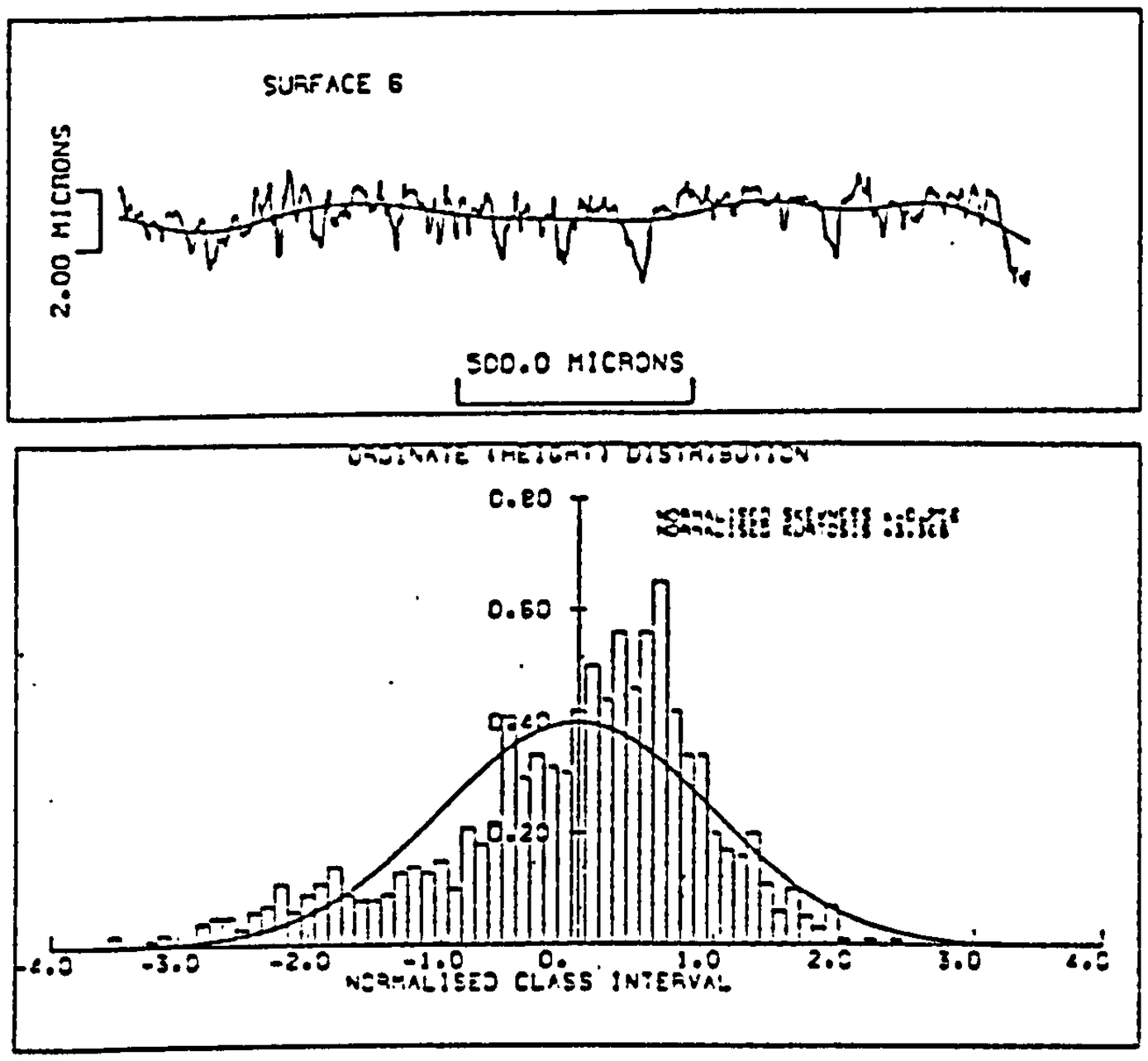


Fig 3 The profile with filter mean line, height distribution, slope distribution and correlation function of a typical profile from reference surface 6

© IMechE 1987 C161/87

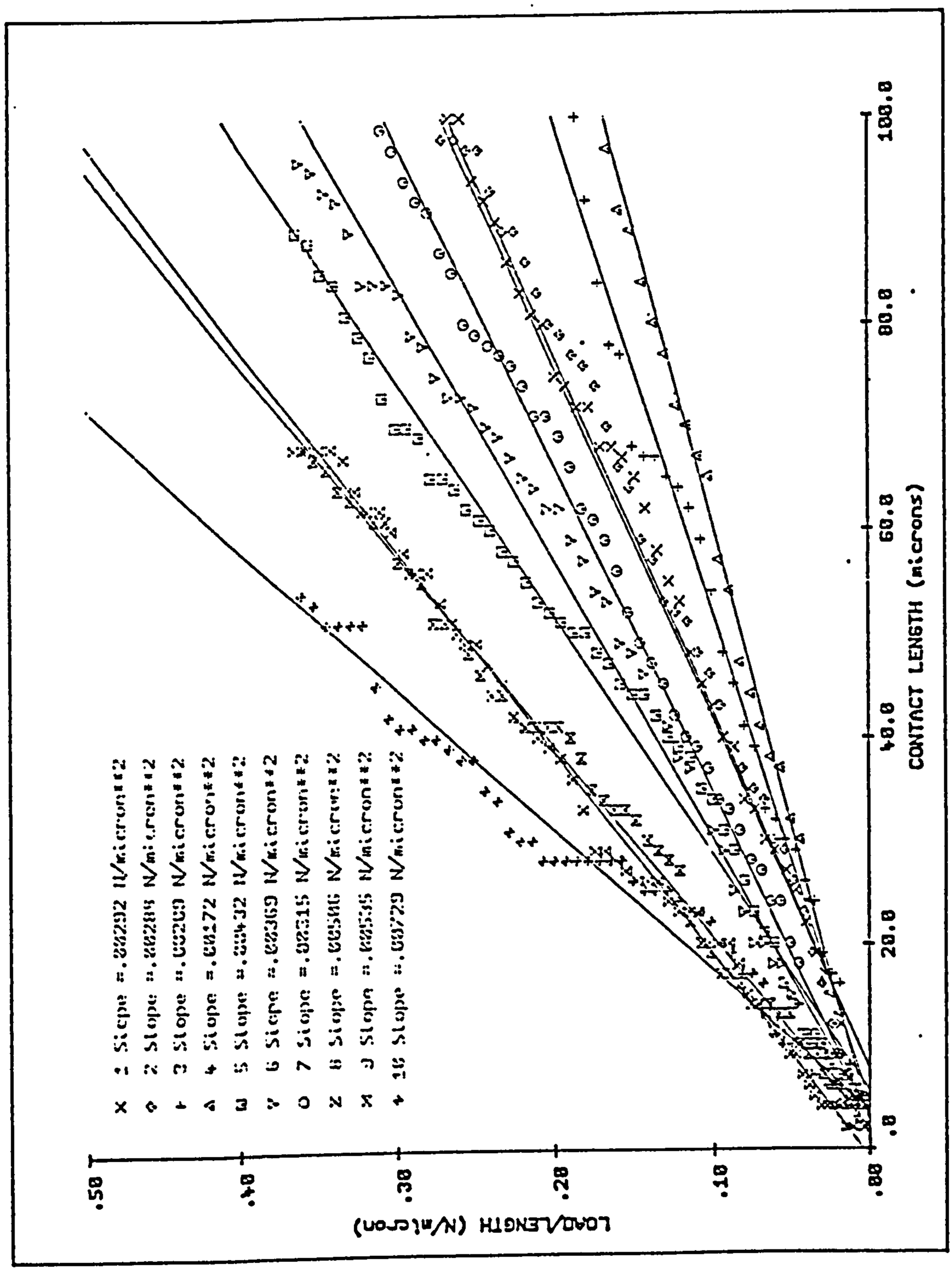


Fig 4 The normal force plotted against real contact area from the two-dimensional numerical contact model, for the ten surfaces analysed

C161/87 © IMechE 1987

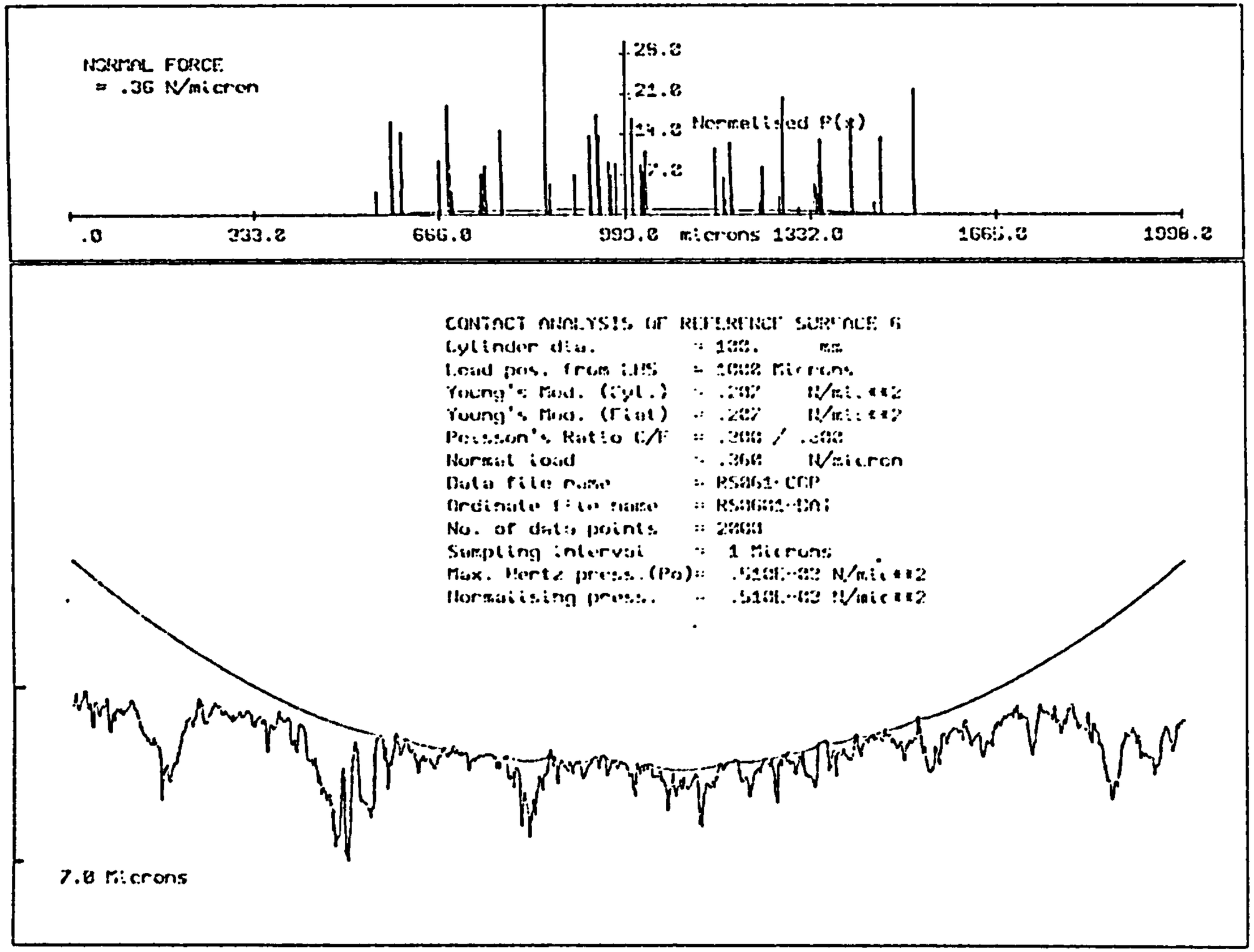


Fig 5 The numerically calculated pressure distribution and contact geometry for a smooth 100 mm diameter cylinder loaded onto a profile of reference surface 6 (assuming constants as for steel)

© IMechE 1987 C161/87

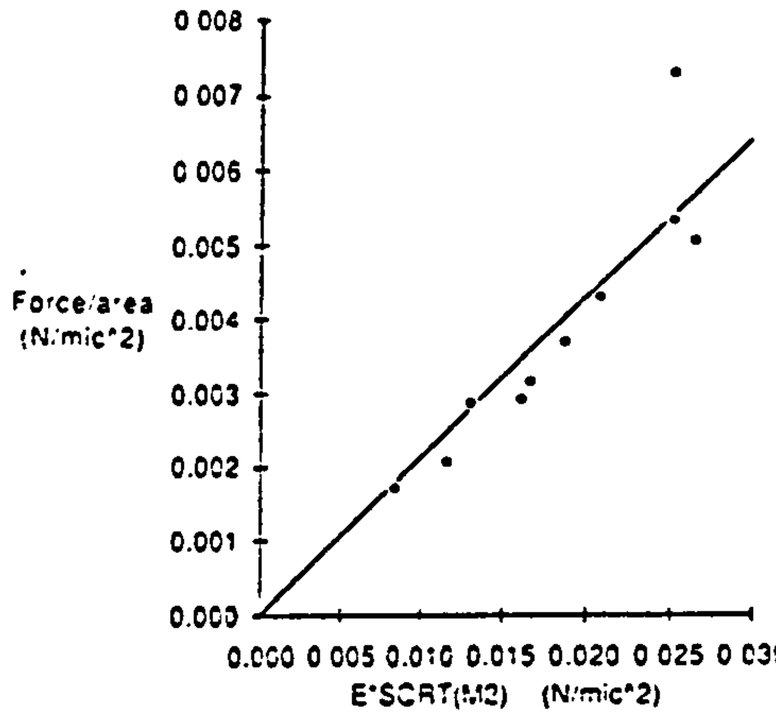


Fig 6 Force/real contact area results from the numerical model plotted against $E^*\sqrt{m_2}$. The straight line gives the result from Bush's theory for anisotropic surfaces

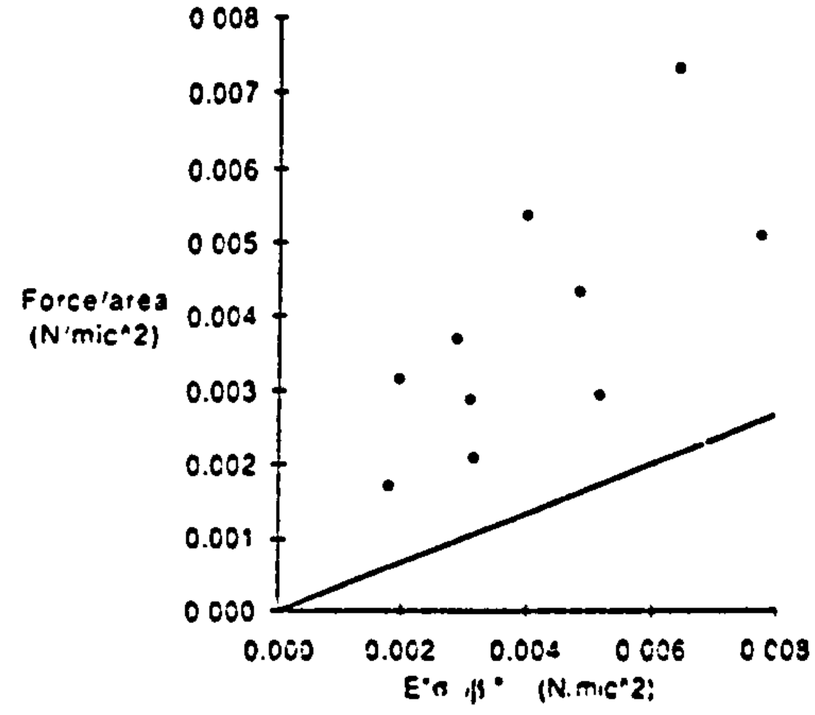


Fig 7 Force/real contact area results from the numerical model plotted against $E^*\sigma/\beta^*$. The straight line gives the result from Onion's and Archard's theory for isotropic surfaces

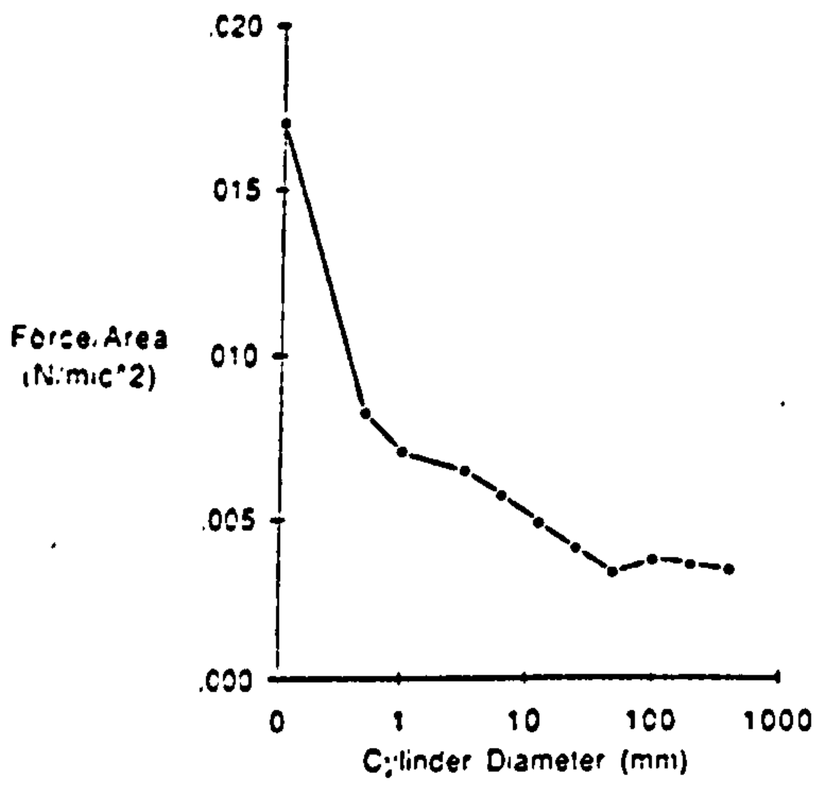


Fig 8 Variation in force/real contact area with cylinder diameter using the numerical model on reference surface 6

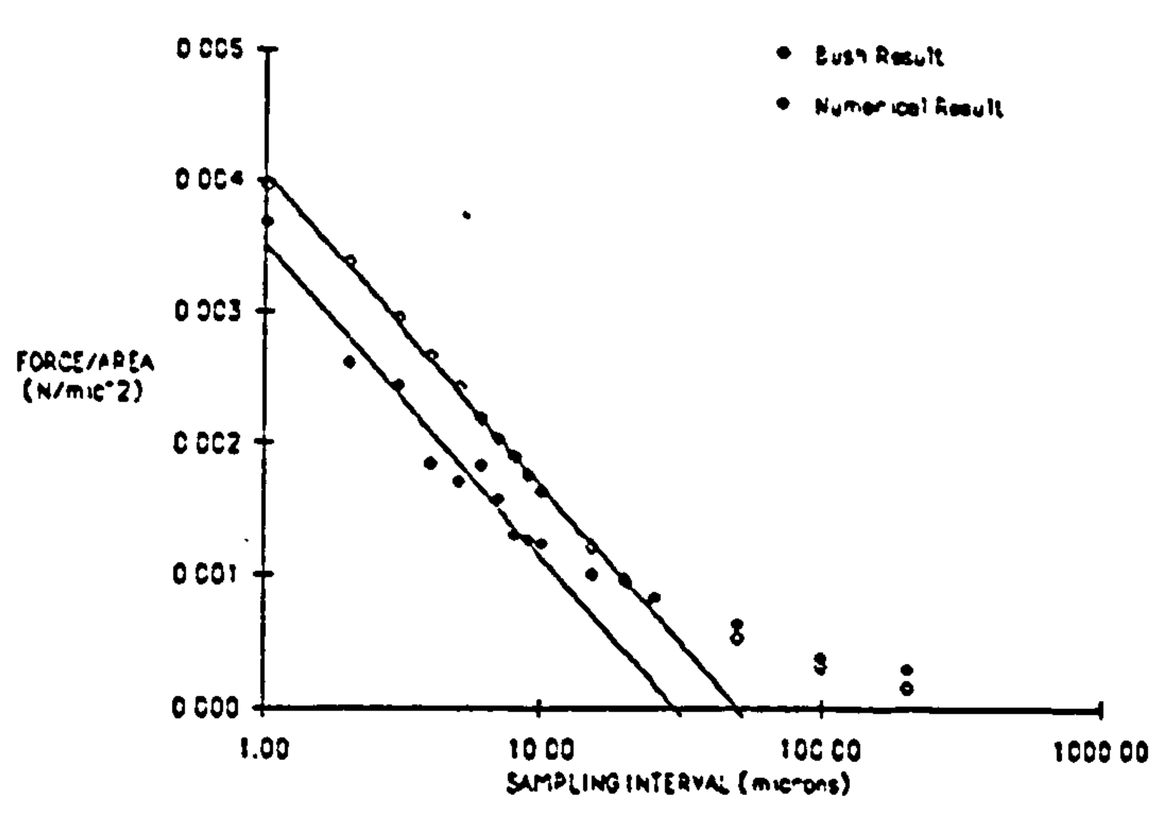


Fig 9 Variation in force/real contact area with profile sampling interval using the numerical and Bush models. The two lines do not coincide as the comparison was carried out using reference surface 6 which did not give perfect agreement between the two models

C161/87 © IMechE 1987

APPENDIX D

A 3-Dimensional Method of Studying 3-Body Contact Geometry and Stress on Real Rough Surfaces

M.A. West and R.S. Sayles

A paper presented at:
The 14th Leeds Lyon Symposium,
'Interface Dynamics',
Lyons,
France,
September 1987

Published in the conference proceedings

A 3-DIMENSIONAL METHOD OF STUDYING 3-BODY CONTACT GEOMETRY AND STRESS ON REAL ROUGH SURFACES

SYNOPSIS

A technique for three-dimensional surface contact analysis is described in detail with examples of results obtained on contact area and stress distribution presented for several typical engineering surface finishes. Other aspects such as the influence of third bodies within the contact and plasticity are examined in terms of their significance to the Hertzian type non-conforming contacts existing in rolling element bearings and gears.

1 INTRODUCTION

The traditional statistical models of rough surface contact using assumed shapes and distributions for the asperities, such as those by Greenwood and Williamson (1) or Nayak (2), have the disadvantage that they do not give the actual deformed geometries of the bodies, especially away from the contact zones. They also tend to become inaccurate for significant levels of penetration between the surfaces.

The availability of powerful micro computers has led to the development of surface measurement and analysis systems where real measured data can be used for contact analysis. One such three-dimensional mapping system (3) uses a Talysurf stylus instrument to record a matrix of surface heights from flat or curved surfaces. Two-dimensional plane strain contact analysis can then be carried out on selected profiles (4 and 5). This paper describes work carried out to extend this method to a full three-dimensional contact model.

2 THE CONTACT MODEL

The aim of the contact model is to predict the surface pressure distribution and deformed contact geometry of a real measured rough surface topography in contact with a smooth geometrically shaped surface. The second body can be smooth because the contact of two uncorrelated rough surfaces can be modelled by the contact of a smooth surface and a rough surface with a variance equal to the sum of the variances of the two original surfaces (6).

The geometry of the second body is that of a surface with it's principal radii of curvature in orthogonal directions. This allows the simulation of most gear and bearing contacts to be attempted.

The model is principly based on elastic contact but, so that the contact conditions can be made more realistic, a simple model of plasticity has been incorporated.

2.1 Constraints of the Model

Even though the model has been kept as general as possible some assumptions have been made, namely:-

1) Strains are small compared to the overall surface geometries, leading to the usual linear elastic theory equations.

2) The elastic contact area is contained within the geometric overlap region and there is no contact outside the region defined by the model.

3) Inside the contact region the penetration is just equal to zero and the surface pressures are greater than or equal to zero.

4) Outside the contact region there is no penetration and the normal surface pressures are zero.

5) The surface slopes are reasonably small so that forces act normally to the surfaces.

6) Tangential stresses and displacements are ignored, ie. it is a frictionless contact.

2.2 Contact Equations

The classical approach to finding the stresses and displacements in an elastic half-space due to surface forces is due to Boussinesq (7) and Cerruti (8), who made use of the theory of potential. The approach is well presented by Love (9). The Boussinesq theory gives the normal surface deflection at a general surface point due to a concentrated contact at some other point. If the deflection of a surface point is required due to a number of concentrated contact points over an area S, then the total deflection in polar coordinates (s,ϕ) is:-

$$u_z = \frac{(1-\nu^2)}{\pi E} \iint_S p(s,\phi)\, ds\, d\phi$$

As the surface data to be used with the contact model is in the form of a rectangular matrix of height data, it would seem sensible to use each known data point to be the centre of a rectangular area of sides 2a and 2b over which the normal pressure is constant. Note that the sampling intervals in the x and y directions are 2a and 2b respectively (Figure 1). Pressure elements similar to these have been used in other contact models (10).

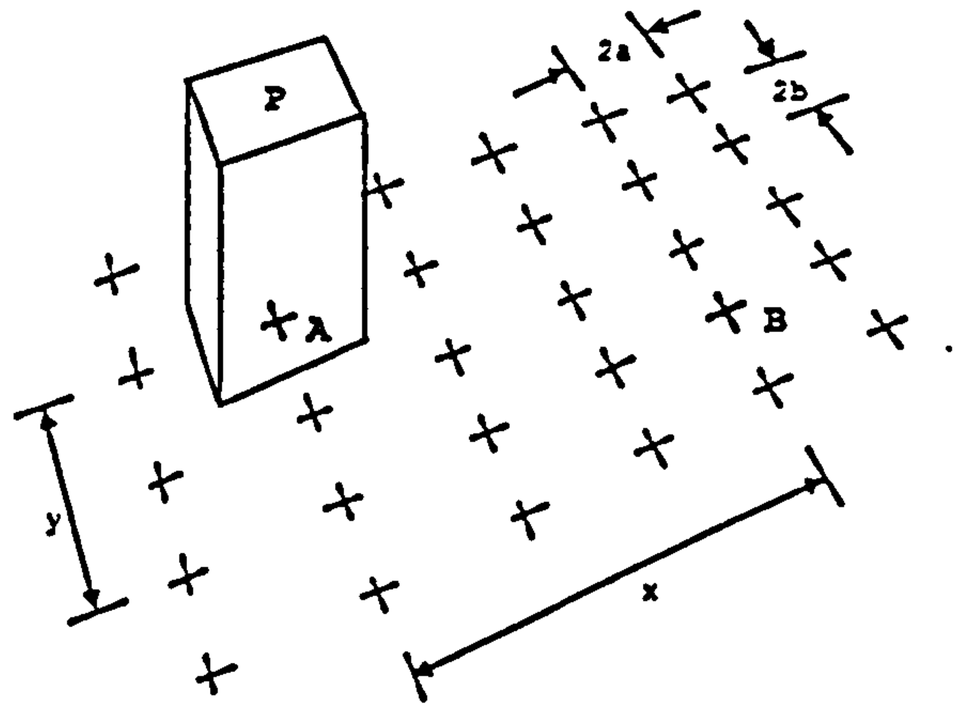


Fig 1 DETAIL OF PRESSURE ELEMENT USED:-
Each cross represents a data point and P is the normal pressure on the element.

The disadvantage of this method is that there is a discontinuity in the pressure at the boundary of each element. This could be overcome by making the elements into triangular pyramids where the pressure varies linearly between its value at each corner. However, the increase in accuracy achieved by doing this does not justify the added complications to the calculation.

The effect of a uniform pressure acting on a rectangular area of 2ax2b has been analysed in connection with rectangular foundations by Love (11). The deflection of a general point (x,y) on the surface due to a uniform pressure p over a rectangle centred at the origin is:-

$$u_z = \frac{(1-\nu^2)}{\pi E} p \left\{ (x+a) \ln \left[\frac{(y+b)+\left\{ (y+b)^2+(x+a)^2 \right\}^{1/2}}{(y-b)+\left\{ (y-b)^2+(x+a)^2 \right\}^{1/2}} \right] \right.$$

$$+ (y+b) \ln \left[\frac{(x+a)+\left\{ (y+b)^2+(x+a)^2 \right\}^{1/2}}{(x-a)+\left\{ (y+b)^2+(x-a)^2 \right\}^{1/2}} \right]$$

$$+ (x-a) \ln \left[\frac{(y-b)+\left\{ (y-b)^2+(x-a)^2 \right\}^{1/2}}{(y+b)+\left\{ (y+b)^2+(x-a)^2 \right\}^{1/2}} \right]$$

$$+ (y-b) \ln \left[\frac{(x-a)+\left\{ (y-b)^2+(x-a)^2 \right\}^{1/2}}{(x+a)+\left\{ (y-b)^2+(x+a)^2 \right\}^{1/2}} \right] \left. \right\}$$

This was used as the basis for the equations in the contact model. This equation, adjusted by putting the pressure as unity, is used to form the influence coefficient of any contacting point on any other point on the surface. The total deflection at any point can therefore be calculated by summing each element of pressure times it's influence coefficient.

$$\text{Total deflection of a point} = \sum \text{elemental pressures} \times \text{influence coefficients}$$

Thus, if the pressures on each contacting element are known, the deflection of any point on the surface and hence the deformed geometry can easily be found. However, the elemental pressures are not known. This can be overcome by giving the bodies a known bulk overlap and noting that the sum of the strains in the two bodies at any contacting point is just equal to the geometric overlap.

The strain in each body can then be found using the relationship:-

$$\frac{u_{z_1} E_1}{(1-\nu_1^2)} = \frac{u_{z_2} E_2}{(1-\nu_2^2)}$$

By assuming some of the points to be in contact, then from the strains at these points and their influence coefficients on all other points of the surface, a set of linear equations can be set up to find the unknown pressures at the contacting points. In matrix terms:-

$$(u_z) = [C] \ (p)$$

Where (u_z) is the vector of surface displacements at the contacting points, [C] is the matrix of influence coefficients and (p) is the vector of pressures at the contacting points.

The matrix [C] is symmetric and the method of solution used is that of Choleski Factorisation. This makes use of the symmetry so that only just over half of the full matrix has to be stored at any one time, thus making efficient use of the computer's memory.

Once the solution for the elemental pressures have been found then any elements where the pressures are negative are removed from the contact (tension is not permitted as adhesive forces are assumed to be relatively small). The deformations are then found everywhere to check that no contact could occur outside the assumed contact region. If so these points are added to the assumed contact region. The system of linear equations is then solved again for the new assumed contact points.

This process is repeated until convergence on a particular set of contact points is achieved. The deformed shapes of the bodies can then be found and the total normal load calculated by integrating the pressure distribution. The general principles of this method of solution by "Matrix Inversion" are outlined in Johnson (12).

2.3 The Plasticity Model

Using a linear elastic contact model, such as the one described, will produce large pressure spikes on most topographies even at very low normal loads due to the finite sampling and infinite pressure gradients between adjacent points of the surface data. These spikes, often well above the elastic limit of the materials being simulated, are obviously not realistic.

One partial solution to this problem is by the use of "Functional Filtering" (13). Both the highest and lowest frequencies present in the data can be filtered out just leaving the bandwidth of interest to the problem being solved. The method of three-dimensional filtering employed on the data is that of a moving average box filter (3). However, if within the scale of roughness being studied plasticity may play a part, it must be included.

The plasticity criterion incorporated into the contact model is that of a perfectly elastic - perfectly plastic material. Assuming the contact to be that between solids of revolution, the point of plastic yield is reached at a value of pressure, p_y such that:-

$$p_y = 1.6Y$$

Where Y is the yield stress of the material in simple tension.

The plastic zone is contained within the interior of the body and the plastic yield manifests itself as a sub-surface movement of the material axi-symmetrically away from the point of loading (this is in agreement with the Johnson "cavity" model of mixed elastic-plastic deformation (12)).

This criterion is added to the calculation by modifying the shape of the original undeformed surfaces at any point at which the pressure is above p_y so that the strain, and hence the pressure, is reduced. The elastic equations are then solved again for this new shape.

This process is repeated until convergence on a solution is reached with no pressures above p_y, and all points where plastic deformation has occurred are subjected to a constant pressure of p_y. The shapes of the modified original bodies then define the elastically recovered surfaces after contact.

3 IMPLEMENTATION OF THE MODEL

The contact simulation program is a set of FORTRAN routines implemented on an Acorn A443 Workstation micro computer. This fast 32 bit machine with 4 megabytes of RAM, although not as powerful or fast as a main-frame computer, was chosen so that it could form part of a stand alone surface data acquisition and analysis system.

The computer chosen effectively limits the surfaces to be analysed to approximately 20,000 data points with a maximum of 1250 data points in contact.

4 VALIDATION OF THE MODEL

The model was tested against theoretical Hertzian results by simulations using a smooth plane instead of a rough surface. For all shapes of smooth curved bodies used the contact patches and pressure distributions were as predicted by theory within the accuracy of the finite data.

5 EXAMPLES USING REAL DATA

5.1 Worn bearing

A 0.5 mm square area of worn bearing surface with some debris damage was used for a series of contact analyses (Figure 2). In all cases the normal load and theoretical contact area were the same but the ratio of the smooth body major to minor radius of curvature was varied.

At the load employed it was found that the actual contact area in all cases was approximately 57% of the theoretical smooth surface value and the maximum pressure spikes were approximately 4.5 times the theoretical Hertzian maximum. Assuming the material to be gear steel with a yield stress in simple tension of 2.6 GPa then the maximum pressure (4.5 GPa) would not be enough to cause any significant plastic deformation. Note that the surface used was that of a worn bearing where any plastic deformation should already have occurred.

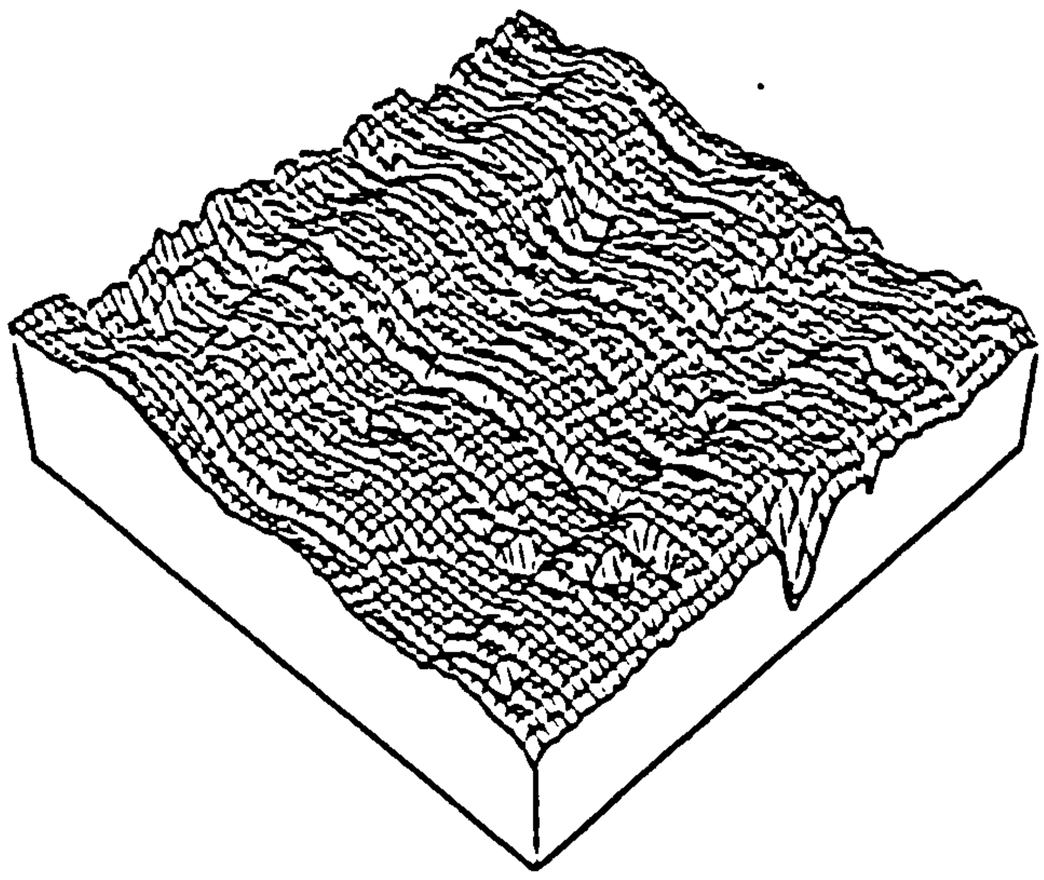


Fig 2 WORN BEARING SURFACE:-
Area = 0.5 mm square
Height range of data = 2.6 µm
Rolling direction = downwards to right

Figures 3,4 and 5 show the actual contact patch, deformed shape of the smooth body and pressure distribution for one of the test cases.

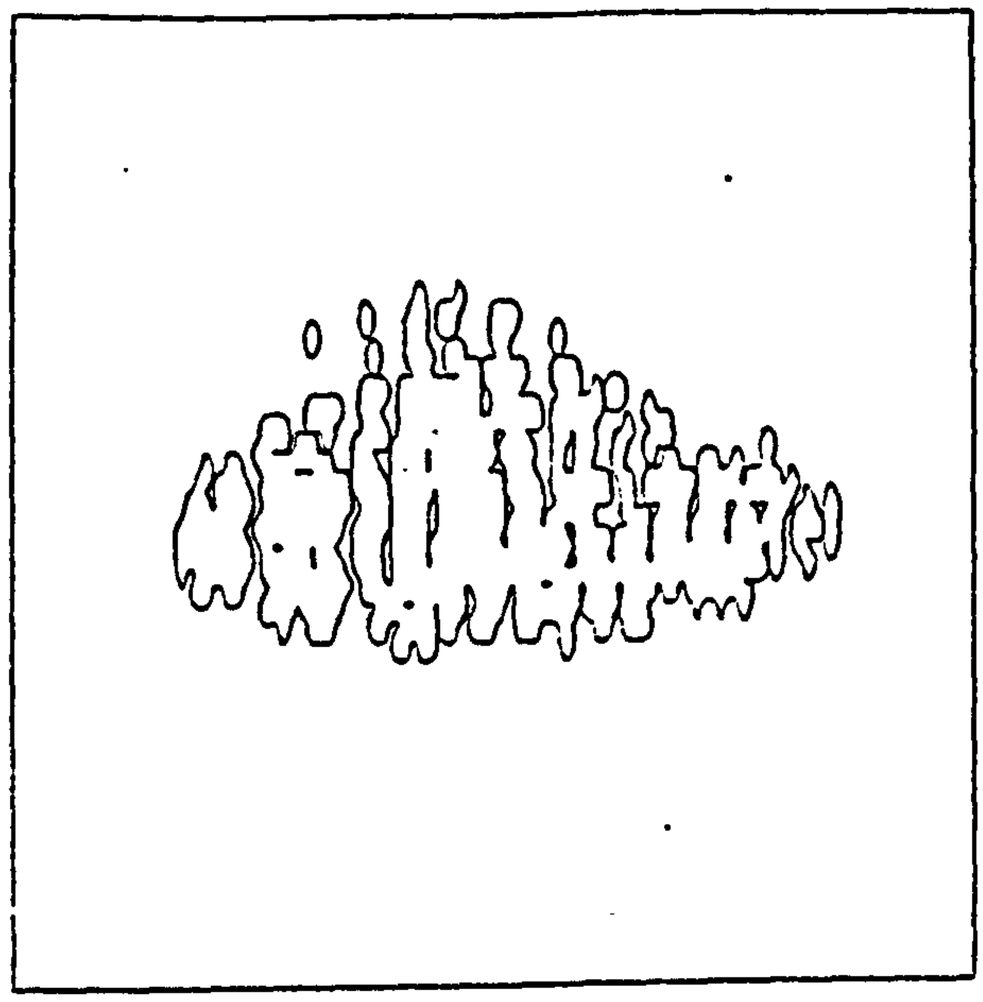


Fig 3 SMOOTH BODY IN CONTACT WITH WORN BEARING:-
CONTACT PATCH
Area shown = 0.5 mm square
Load = 25 N, Hertz P_0 = 1 GPa
Smooth Body Radii of curvature = 15 mm/ 5 mm
Hertz semi-contact widths = 162 µm / 78 µm
Ratio of true to apparent contact area = 0.57

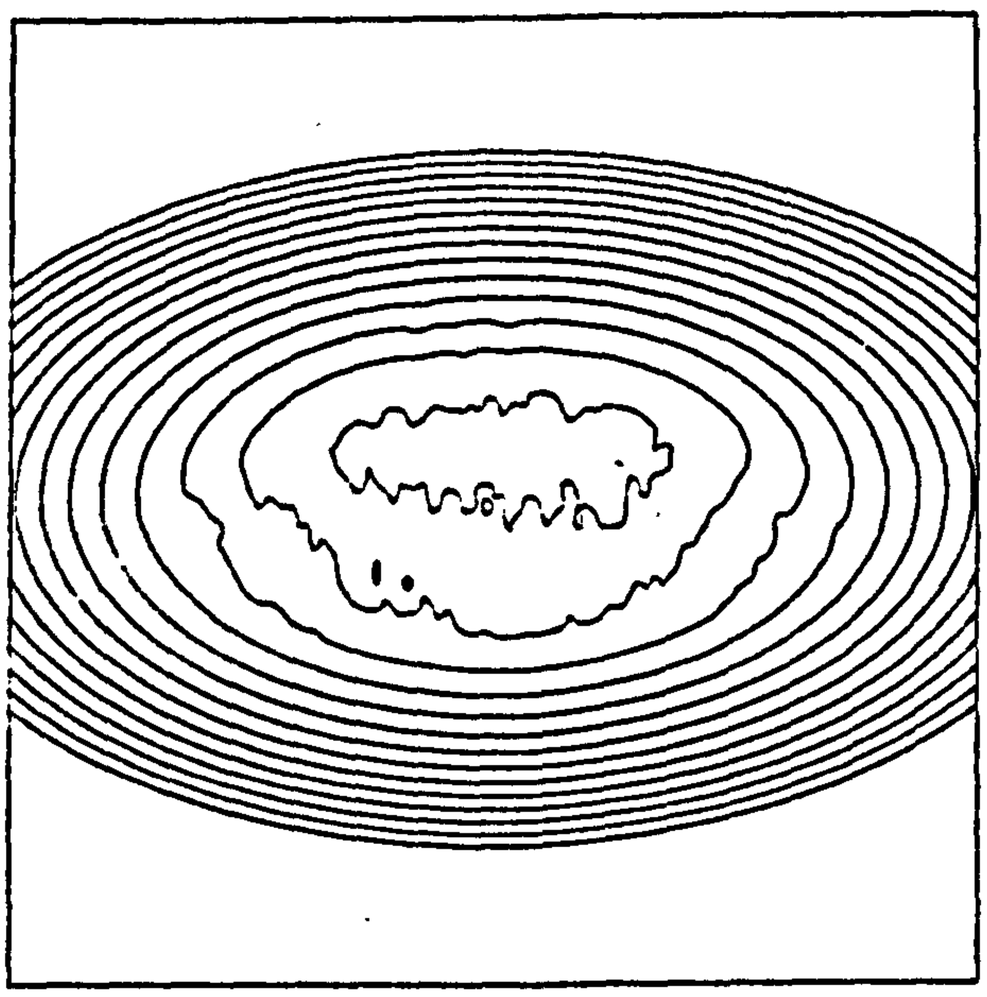


Fig 4 SMOOTH BODY IN CONTACT WITH WORN BEARING:-
DEFORMED SHAPE OF SMOOTH BODY
Load = 25 N, Hertz P_0 = 1 GPa
Smooth Body Radii of curvature = 15 mm/ 5 mm

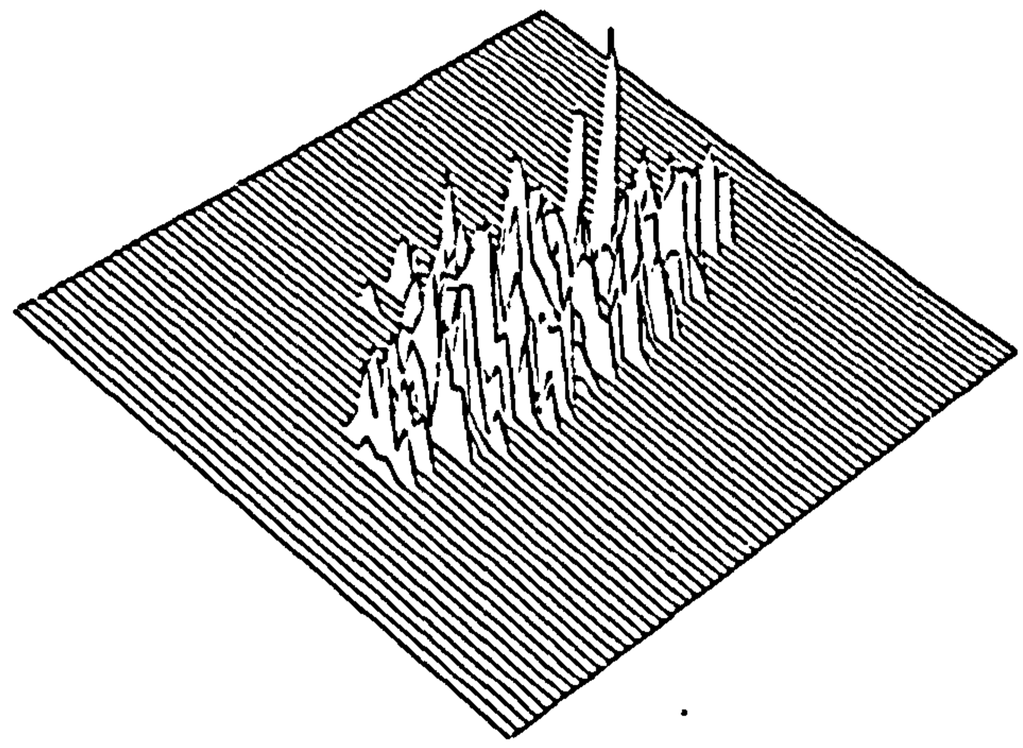


Fig 5 SMOOTH BODY IN CONTACT WITH WORN BEARING:-
SURFACE PRESSURE DISTRIBUTION
Load = 25 N, Hertz P_0 = 1 GPa
Smooth Body Radii of curvature = 15 mm/ 5 mm
Hertz semi-contact widths = 162 µm / 78 µm
Maximum pressure spike = 4.5 GPa

5.2 Abraded Steel Surface

Similar contact analyses were carried out using a 0.5 mm square area of abraded steel surface Figure 6. It was found that the actual contact area in all cases was approximately 40% of the theoretical smooth surface value and the maximum pressure spikes were approximately 7.0 times the theoretical Hertzian maximum. Figures 7,8 and 9 show the actual contact patch, deformed shape of the smooth body and pressure distribution for one of the test cases.

Assuming the same material as the bearing tests the maximum pressures recorded of 7.0 GPa would cause plastic deformation. The tests were therefore repeated using the plasticity criterion and it was found that approximately 5% of the contact area would deform plastically. However the size of the contact area was not increased significantly.

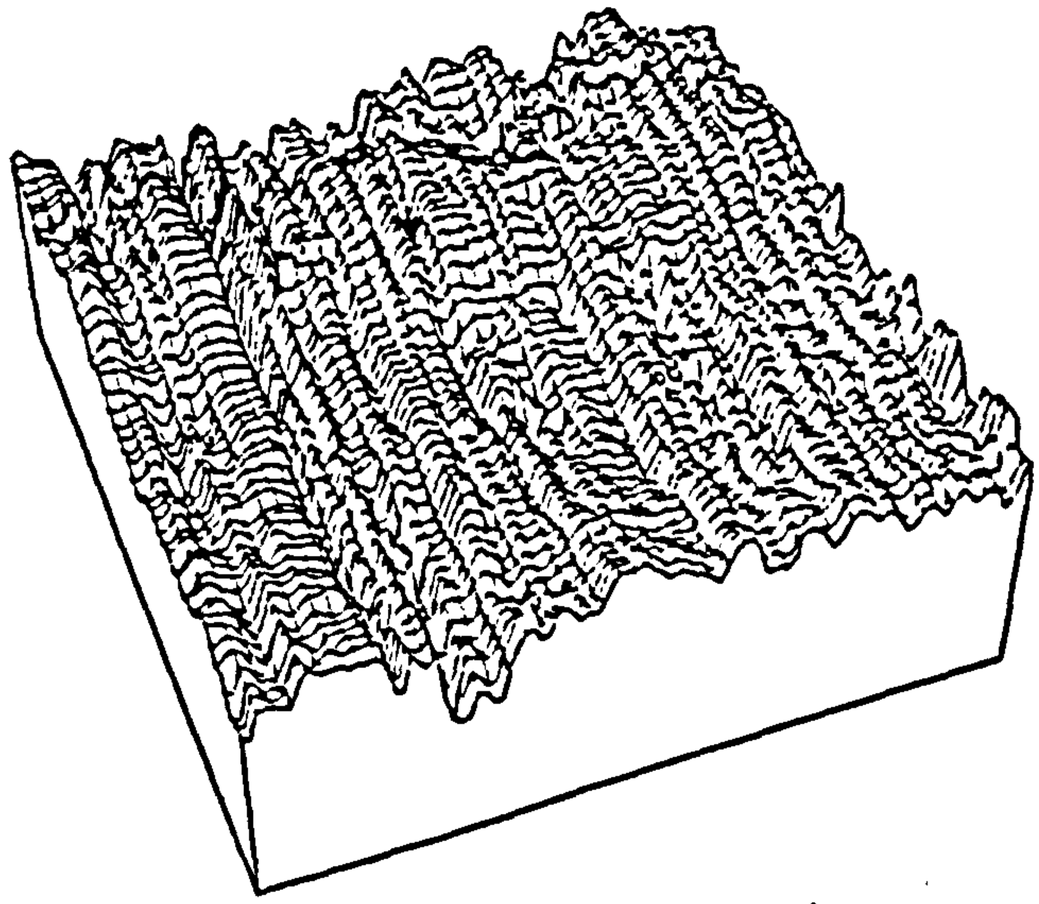


Fig 6 ABRADED STEEL SURFACE:-
Area = 0.5 mm square
Height range of data = 4.4 µm

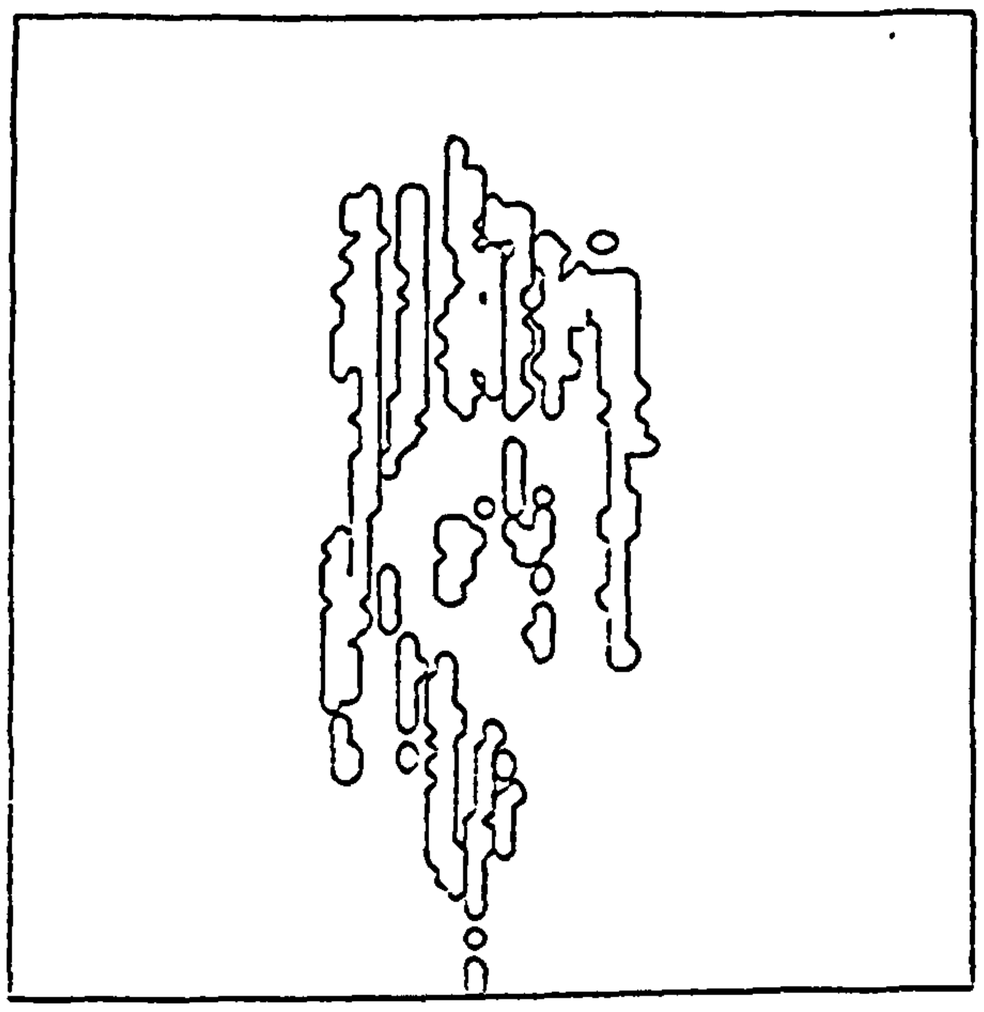


Fig 7 SMOOTH BODY IN CONTACT WITH ABRADED STEEL:-
CONTACT PATCH
Load = 25 N, Hertz P_0 = 1 GPa
Smooth Body Radii of curvature = 5 mm/ 15 mm
Hertz semi-contact widths = 78 µm / 162 µm
Ratio of true to apparent contact area = 0.39

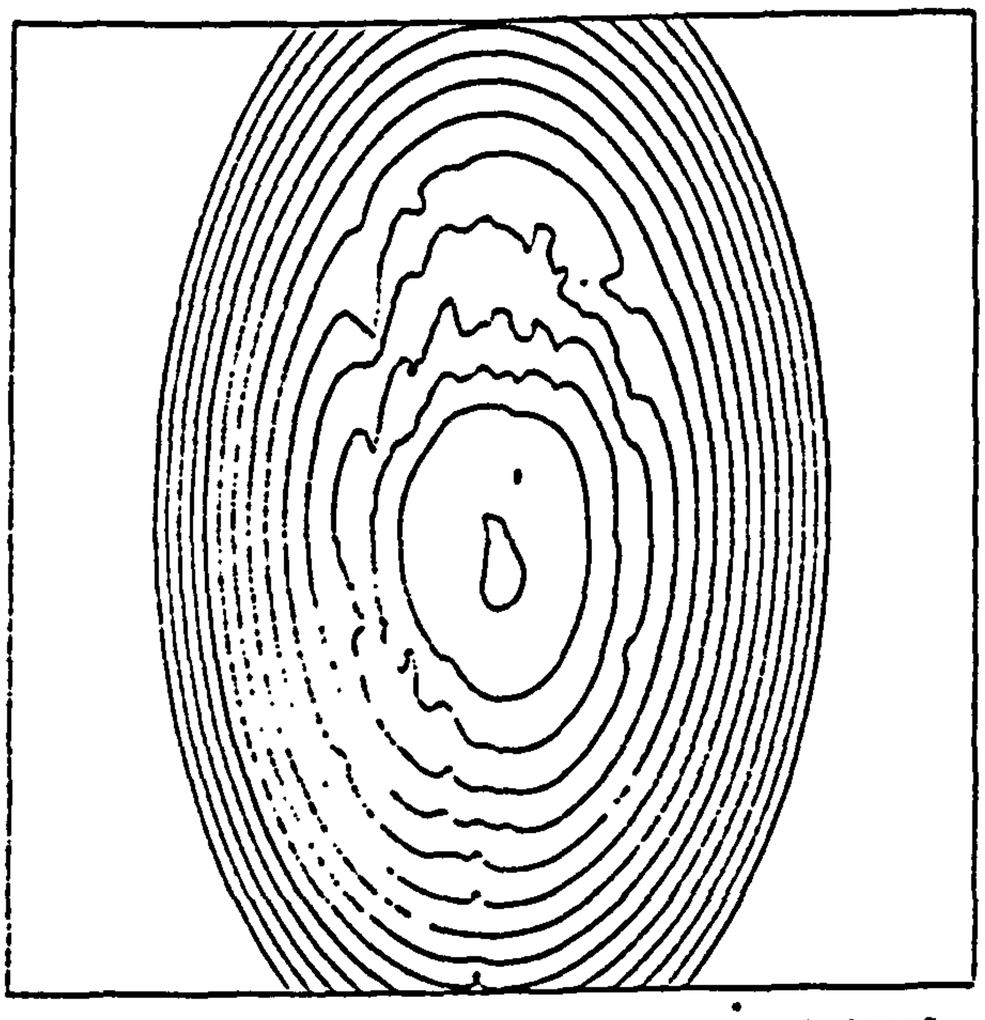


Fig 8 SMOOTH BODY IN CONTACT WITH ABRADED STEEL:-
DEFORMED SHAPE OF SMOOTH BODY
Load = 25 N, Hertz P_0 = 1 GPa
Smooth Body Radii of curvature = 5 mm/ 15 mm

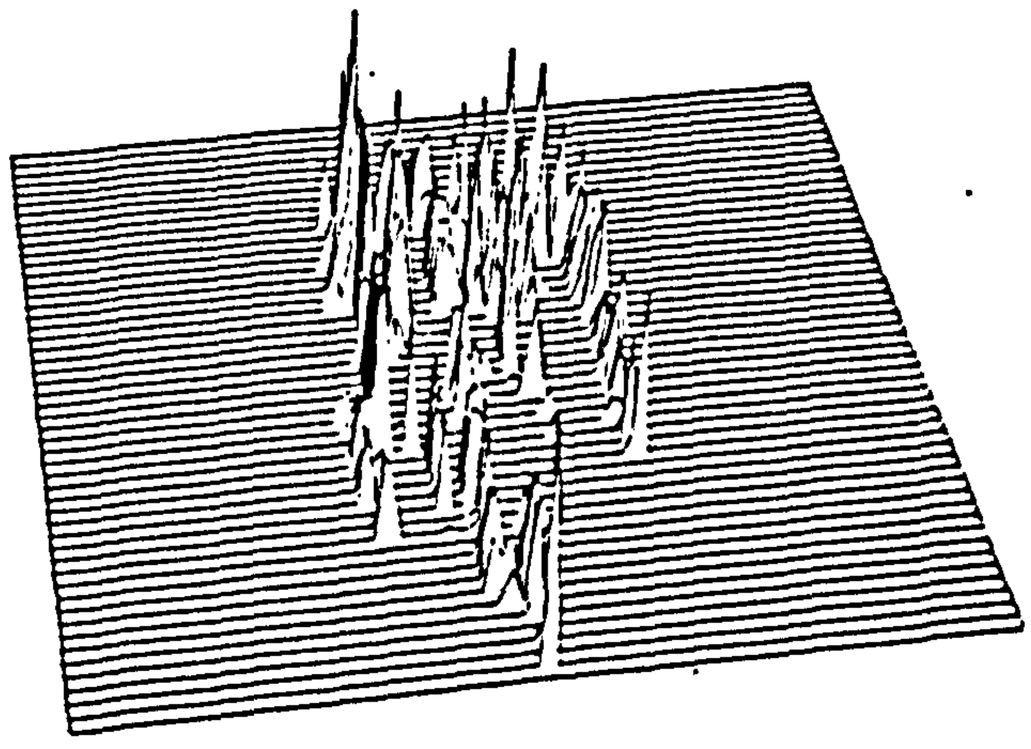


Fig 9 SMOOTH BODY IN CONTACT WITH ABRADED STEEL:-
SURFACE PRESSURE DISTRIBUTION
Load = 25 N, Hertz P_0 = 1 GPa
Smooth Body Radii of curvature = 5 mm/ 15 mm
Hertz semi-contact widths = 78 µm / 162 µm
Ratio of true to apparent contact area = 0.39
Maximum pressure spike = 7.0 GPa

6 THE EFFECT OF DEBRIS

To study the effect of debris a simulation of the contact of a smooth 0.1 mm radius sphere on a smooth steel surface was carried out. The normal load was taken as 25 N as in the earlier tests so that this could represent a debris particle becoming trapped in the contact.

Taking the material to be the same as before virtually the whole contact undergoes plastic deformation. Figure 10 shows the 4 µm deep dent left in the surface after the load has been taken away. For comparison Figure 11 shows some detail of the debris damage on the worn bearing surface drawn to the same scale.

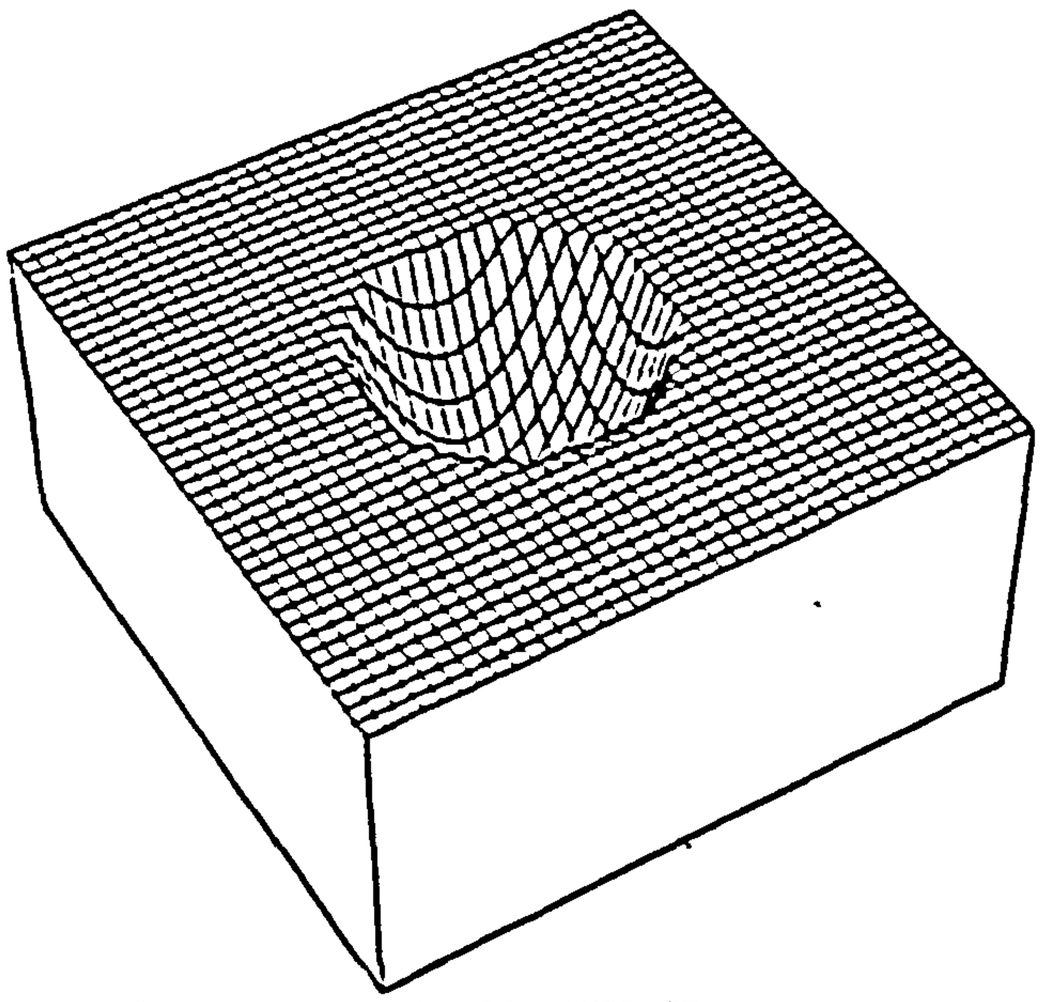


Fig 10 PLASTICALLY DENTED SURFACE:-
Dent caused by simulation of 0.1 mm radius sphere being pressed into a smooth surface
Area = 0.2 mm square
Height range of data = 4.0 µm
Load = 25 N, Hertz P_0 = 18.4 GPa
Yield Stress = 2.6 GPa

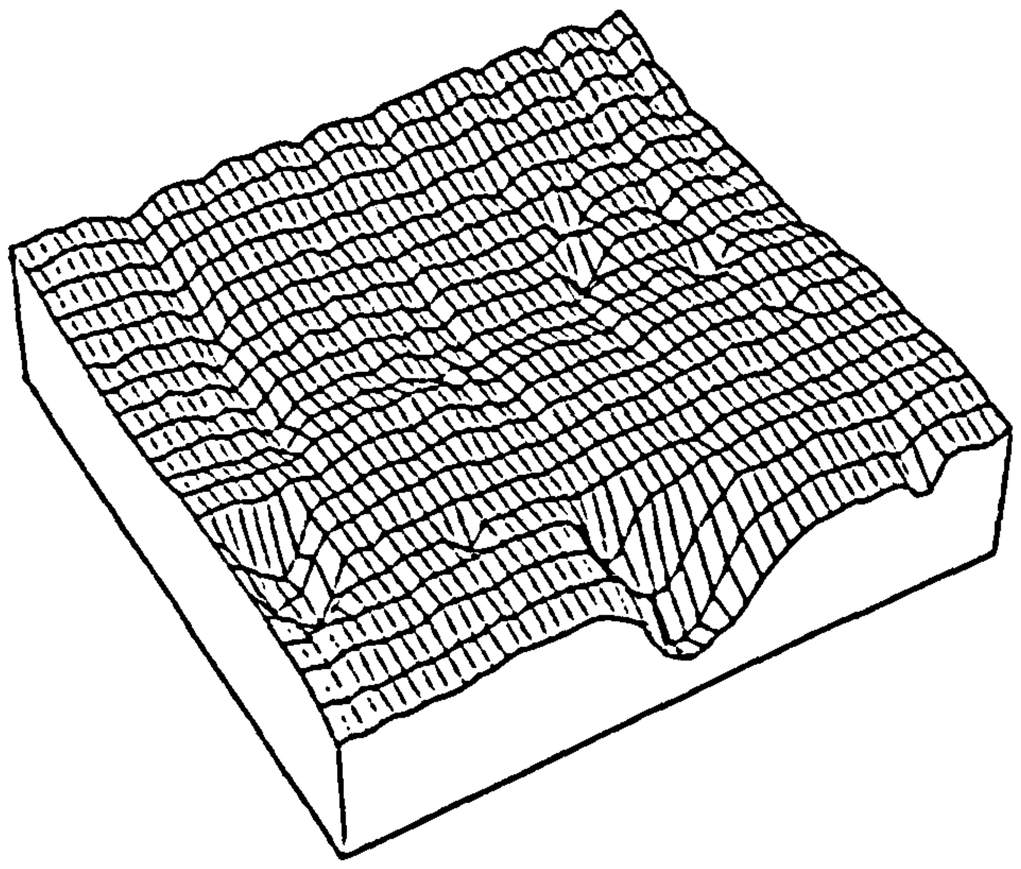


Fig 11 DETAIL OF WORN BEARING SURFACE:-
Area = 0.2 mm square
Height range of data = 2.6 μm
Rolling direction = downwards to right

7 CONCLUSIONS

A numerical technique is described which is capable of simulating the elastic contact of rough surfaces and estimating the effect of plasticity. The method is shown to be capable of creating the real pressure distribution together with a complete picture of local deformation and real contact area on an asperity scale and also to define the shapes of the recovered surfaces after plastic deformation. The technique is illustrated by its application to the contact in several typical engineering surfaces.

Results are also presented showing the effect on the surfaces of debris entering the contact zone.

REFERENCES

(1) GREENWOOD, J.A. and WILLIAMSON, J.B.P. "Contact of Nominally Flat Surfaces", Proc. Roy. Soc. Lond., 1966, A295, 300-319.

(2) NAYAK, P.R. "Random Process Model of Rough Surfaces", Trans. ASME J. Lub. Tech., 1971, 93F, 398-407.

(3) WEBSTER, M.N., WEST, M.A. and SAYLES, R.S. " A Method of 3-Dimensional Topography Measurement and Analysis on Arcuate Surfaces", Wear, 1986, 109, 385-399.

(4) WEBSTER, M.N. and SAYLES, R.S. "A Numerical Model for the Elastic Frictionless Contact of Real Rough Surfaces", Trans. ASME J. of Tribology, 1986, 108, 314-320.

(5) WEST, M.A., WEBSTER, M.N. and Sayles, R.S. "A New Method for Rough Surface Contact Analysis", Proc Instn. Mech. Engrs, 'Tribology - Friction, Lubrication and Wear fifty years on', 1987, 545-955.

(6) THOMAS, T.R.(Ed.) "Rough Surfaces", 1982 (Longman Group, London).

(7) BOUSSINESQ, J. "Application des potentials a l'etude de l'equilibre et du mouvement des solides elastiques", 1885, Paris: Gauthier-Villars.

(8) CERRUTI, V. 1882, Roma,Acc. Lincei, Mem. fis. mat.

(9) LOVE, A.E.H. "A Treatise on the Mathematical Theory of Elasticity", 4th Edition 1952 (Cambridge University Press).

(10) DE MUL, J.M., KALKER, J.J. and FREDRIKSSON, B. "The Contact Between Arbitrarily Curved Bodies of Finite Dimensions", Trans. ASME J. of Tribology, 1986, 108, 140-148.

(11) LOVE, A.E.H. "The Stress Produced in a Semi-Infinite Solid by Pressure on Part of the Boundary", Proc. Roy. Soc. Lond., 1929, A228, 377.

(12) JOHNSON, K.L. "Contact Mechanics", 1985 (Cambridge University Press).

(13) THOMAS, T.R. and SAYLES, R.S. Discussion to Radhakrishnan, V., "Analysis of Some of the Reference Lines Used for Measuring Surface Roughness", Proc. Inst. Mech. Engrs, 1973, 187, 575-582.

REFERENCES

ABBOTT, E.J. and **FIRESTONE, F.A.** (1933) "Specifying Surface Quality", Mechanical Engineering, 55, 569-572.

ARCHARD, J.F. (1953) "Elastic Deformation and Contact of Surfaces", Nature, 172, 918-919.

ARCHARD, J.F. (1957) "Elastic Deformation and the Laws of Friction", Proc. Roy. Soc. Lond., A243, 190-205.

BAY, N., WANHEIM, T. and **PETERSEN, A.S.** (1975) "R_a and the Average Effective Strain of Surface Asperities Deformed in Metal-Working Processes", Wear, 34, 77-84.

BENDAT, J.S. (1958) "Principles and Applications of Random Noise Theory", (Wiley, New York).

BOUSSINESQ, J. (1885) "Application des potentials a l'etude de l'equilibre et du mouvement des solides elastiques", (Gauthier-Villars, Paris).

BOWDEN, F.P. and **TABOR, D.** (1954) "The Friction and Lubrication of Solids - Part 1", (Oxford University Press).

BOWDEN, F.P. and **TABOR, D.** (1964) "The Friction and Lubrication of Solids - Part 2", (Oxford University Press).

BUSH, A.W., GIBSON, R.D. and KEOGH, G.P. (1979) "Strongly Anisotropic Rough Surfaces", Trans. ASME, J. Lubr. Tech., 101, 15-20.

BUSH, A.W., GIBSON, R.D. and THOMAS, T.R. (1975) "The Elastic Contact of a Rough Surface", Wear, 1975, 35, 87-111.

CARSON, R.M. and JOHNSON, K.L. (1971) "Surface Corrugations Spontaneously Generated in a Rolling Contact Disc Machine", Wear, 17, 59-72.

CERRUTI, V. (1882), Roma, Acc. Lincei, Mem. fis. mat.

CHILDS, T.C.H. (1973) "The Persistence of Asperities in Indentation Experiments", Wear, 25, 3-16.

CHILDS, T.C.H. (1977) "The Persistence of Roughness Between Surfaces in Static Contact", Proc. Roy. Soc. Lond., A353, 35-53.

CHIVERS, T.C., MITCHELL, L.A. and ROWE, M.D. (1974) "The Variation of Real Contact Area Between Surfaces with Contact Pressure and Material Hardness", Wear, 28, 171-185.

CHIU, Y.P. and HARTNETT, M.J. (1983) "A Numerical Solution for Layered Solid Contact Problems With Application to Bearings", Trans. ASME J. Lub. Tech., 105, 585-590.

CHIU, Y.P. and **HARTNETT, M.J.** (1987) "A Numerical Solution for the Contact Problem Involving Bodies with Cylindrical Surface Considering Cylinder Effect", Trans. ASME J. of Tribology, 109, 479-486.

CONRY, T.F. and **SEIREG, A.** (1971) "A Mathematical Programming Method for Design of Elastic Bodies in Contact", Trans. ASME J. Appl. Mech., 38, 387-392.

DE MUL, J.M., KALKER, J.J. and **FREDRIKSSON, B.** (1986) "The Contact Between Arbitrarily Curved Bodies of Finite Dimensions", Trans. ASME J. of Tribology, 108, 140-148.

DE MUL, J.M., VREE, J.M. and **KUYPERS, J.C.** (1987) "The Influence of Certain Raceway Dent Geometries (3-D) on Contact Stresses and Rating Fatigue Life of Rolling Bearings", Trans. ASME J. of Tribology, 109, 452-461.

DUVAUT, G. and **LIONS, J.L.** (1972) "Les Inéquations en Mécanique et en Physique", (Dunod, Paris).

EDWARDS, C.M. and **HALLING, J.** (1968) "An Analysis of the Plastic Interaction of Surface Asperities and its Relevance to the Value of the Coefficient of Friction", J. Mech. Eng. Sci., 10, 101-110.

FRANCIS, H.A. (1982) "A Finite Surface Element Model for Plane-Strain Elastic Contact", Wear, 76, 221-245.

GRAHAM, R. (1979) "Pitting of Hard Steel", PhD Thesis, University of London.

GREENWOOD, J.A. (1967) "The Area of Contact Between Rough Surfaces and Flats", Trans. ASME, J. Lubr. Tech., 89, 81-91.

GREENWOOD, J.A. (1984) "A Unified Theory of Surface Roughness", Proc. Roy. Soc. Lond., A393, 133-157.

GREENWOOD, J.A. and TRIPP, J.H. (1967) "The Elastic Contact of Rough Spheres", Trans. ASME, J. Appl. Mech., 34, 153-159.

GREENWOOD, J.A. and TRIPP, J.H. (1970) "The Contact of Two Nominally Flat Rough Surfaces", Proc. Inst. Mech. Engrs, 185, 625-633.

GREENWOOD, J.A. and WILLIAMSON, J.B.P. (1966) "Contact of Nominally Flat Surfaces", Proc. Roy. Soc. Lond., A295, 300-319.

HARTNETT, M.J. (1979) "The Analysis of Contact Stresses in Rolling Element Bearings", Trans. ASME J. Lub. Tech., 101, 105-109.

HARTNETT, M.J. (1980) "A General Numerical Solution for Elastic Bodies Contact Problem", in "Solid Contact and Lubrication", edited by Cheng, H.S. and Keer, L.M., ASME Publication No. AMD-39, 51-66.

HERTZ, H. (1896) "On the Contact of Rigid Elastic Solids, and on Hardness", Miscellaneous Papers, (Macmillan, London).

HIRST, W. and **HOLLANDER, A.E.** (1974) "Surface Finish and Damage in Sliding", Proc. Roy. Soc. Lond., A337.

HISAKADO, T. (1969) "On the Mechanism of Contact Between Solid Surfaces, (3rd Report, the Number and the Distribution of radii of Contact Points)", Bull. J.S.M.E., 12, 1537-1545.

HISAKADO, T. (1975) "Surface Roughness and Deformation of Contact Asperities Between a Rough and a Flat Surface", Wear, 35, 53-61.

HOLM, R. (1967) "Electric Contacts, Theory and Application", 4th edition, (Springer-Verlag, New York).

HUFFINGTON, J.D. (1962) "Method for the Measurement of Surface Microtopography; Experiments with Gelatin Models", J. Appl. Phys., 39, 38-40.

ISHIBASHI, A. and **MATSUMOTO, S.** (1972) "Effects of Surface Roughness and Form Factor on Rolling Contact Fatigue", Bulletin of the JSME, 15, 1434-1447.

JESSOP, D.A.C. (1984) "Imperial College Mini-disc Scuffing Rig:- Operating Manual for Torque Measuring System", WHL Research and Technology Division.

JOHNSON, K.L. (1982) "One Hundred Years of Hertz Contact", Proc. Inst. Mech. Engrs, 196, 363-378.

JOHNSON, K.L. (1985) "Contact Mechanics", (Cambridge University Press).

JOHNSON, K.L., GREENWOOD, J.A. and HIGGINSON, J.G. (1985) "The Contact of Elastic Regular Wavy Surfaces", Int. J. Mech. Sci., 27, 383-396.

KALKER, J.J. (1974) "Aspects of Contact Mechanics", Proc. Conf. "The Mechanics of the Contact between Deformable Bodies", Int. Union of Theoretical and Applied Mechanics, Netherlands, 1-25.

KALKER, J.J. (1977) "Variational Principles of Contact Elastostatics", J. Instn. Math. and Appl., 20, 199-219.

KALKER, J.J. and VAN RANDEN, Y. (1972) "A Minimum Principle for Frictionless Elastic Contact", J. of Eng. Math., 6, 193-206.

KRAGELSKY, I.V. and DEMKIN, N.B. (1960) "Contact Area of Rough Surfaces", Wear, 3, 170-187.

LEAVER, R.H., SAYLES R.S. and THOMAS, T.R. (1974) "Mixed Lubrication and Surface Topography of Rolling Contacts", Proc. Inst. Mech. Engrs, 188, 461-469.

LI, D.F., KAUZLARICH, J.J. and JAMISON, W.E. (1976) "Surface Roughness Effects on Fatigue in Partial EHD Lubrication", Trans. ASME J. Lub. Tech..

LINCOLN, B. (1953) "Elastic deformation and the Laws of Friction", Nature, 172, 169-170.

LING, F.F. (1958*a*) "On Asperity Distribution of Metallic Surfaces", J. Appl. Phys., 29, 1168-1174.

LING, F.F. (1958*b*) "Some Factors Influencing the Area-Load Characteristics for Semismooth Contiguous Surfaces under 'Static Loading'", Trans. ASME, 80, 1113-1120.

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

LONGUET-HIGGINS, M.S. (1957*a*) "The Statistical Analysis of a Random, Moving Surface", Phil. Trans. Roy. Soc., A249, 321-387.

LONGUET-HIGGINS, M.S. (1957*b*) "Statistical Properties of an Isotropic Random Surface", Phil. Trans. Roy. Soc. Lond., A250, 157-174.

LOVE, A.E.H. (1929) "The Stress Produced in a Semi-Infinite Solid by Pressure on Part of the Boundary", Proc. Roy. Soc. Lond., A228, 377.

LOVE, A.E.H. (1952) "A Treatise on the Mathematical Theory of Elasticity", 4th Edition, (Cambridge University Press).

MANDELBROT, B.B. (1977) "Fractals: Form, Chance and Dimension", (W.H. Freeman and Co., New York).

MANDLEBROT, B.B. (1983) "The Fractal Geometry of Nature", (W.H. Freeman and Co., New York).

MITCHELL, L.A. and **ROWE, M.D.** (1969) "Influence of Asperity Deformation Mode on Gas Leakage between Contacting Surfaces", J. Mech. Eng. Sci., 11, 534-545.

NAYAK, P.R. (1971) "Random Process Model of Rough Surfaces", Trans. ASME J. Lubr. Tech., 93F, 398-407.

NAYAK, P.R. (1973) "Some Aspects of Surface Roughness Measurement", Wear, 26, 165-174.

OLVER, A.V. (1983) "Micropitting and Asperity Deformation", Proceedings of the 10th Leeds-Lyon Symposium, September 1983.

OLVER, A.V., SPIKES, H.A., BOWER, A.F. and **JOHNSON, K.L.** (1986) "The Residual Stress Distribution in a Plastically Deformed Model Asperity", Wear, 107, 151-174.

OLVER, A.V., SPIKES, H.A. and **MACPHERSON, P.B.** (1985) "Wear in Rolling Contacts", Proceedings ASME Wear of Materials Conference, Vancouver, 254-272.

ONIONS, R.A. and **ARCHARD, J.F.** (1973) "The Contact of Surfaces having a Random Structure", J. Phys. D: Appl. Phys., 6, 289-304.

PULLEN, J. and **WILLIAMSON, J.B.P.** (1972) "On the Plastic Contact of Rough Surfaces", Proc. Roy. Soc. Lond., A327, 159-173.

SAYLES, R.S. (1978) "Surface Topography and Bearing Vibration", Ransome Hoffmann Pollard Ltd., Bearing Centre Report Number 0242/78/2.

SAYLES, R.S. and **THOMAS, T.R.** (1976) "Thermal Conductance of a Rough Elastic Contact", J. Applied Energy, 2, 249-267.

SAYLES, R.S. and **THOMAS, T.R.** (1977) "The Spatial Representation of Surface Roughness by Means of the Structure Function: A Practical Alternative to Correlation", Wear, 42, 263-276.

SAYLES R.S. and **WEBSTER, M.N.** (1985) "The Characteristics of Surface Roughness Important to Gear and Rolling Bearing Problems", AGARD Conf. Proc. 369, "Gear and Power Transmission Systems for Helicopters and Turboprops".

SMITH, J.D. (1983) "Gears and their Vibration", (Macmillan Press).

STUCKY, M.J. (1985) "Vibration Health Monitoring Research Simulation of Gear Tooth Pitting", WHL Mechanical Research Report MRR 20090.

SUGIMURA, J. and KIMURA, Y. (1984) "Characterisation of Topographical Changes During Lubricated Wear", Wear, 98, 101-116.

THOMAS, T.R. (1975) "Recent Advances in the Measurement and Analysis of Surface Microgeometry", Wear, 33, 205-233.

THOMAS, T.R. (Ed.) (1982) "Rough Surfaces", (Longman Group, London).

THOMAS, T.R. and PROBERT, S.D. (1970) "Establishment of Contact Parameters from Surface Profiles", J. Phys. D: Appl. Phys., 3, 277-289.

THOMAS, T.R. and SAYLES, R.S. (1973) Discussion to Radhakrishnan, V., "Analysis of Some of the Reference Lines Used for Measuring Surface Roughness", Proc. Inst. Mech. Engrs, 187, 575-582.

THOMAS, T.R., UPPAL, A.H. and PROBERT, S.D. (1971) "Hardness of Rough Surfaces", Nature Physical Sci., 229, 86-87.

TSUKIZOE, T. and HISAKADO, T. (1968) "On the Mechanism of Contact Between Metal Surfaces: Part 2 - the Real Area and Number of the Contact Points", Trans. ASME J. Lubr. Tech., 90, 81-88.

UPPAL, A.H. and PROBERT, S.D. (1972a) "Deformation of Single and Multiple asperities on Metal Surfaces", Wear, 20, 381-400.

UPPAL, A.H. and **PROBERT, S.D.** (1972*b*) "Considerations Governing the Contact Between a Rough and a Flat Surface", Wear, 22, 215-234.

UPPAL, A.H. and **PROBERT, S.D.** (1973) "Deformation of Single and Multiple asperity Models of Modelling Clay", Wear, 23, 367-375.

WEBSTER, M.N. (1986) "Measurement and Contact Analysis of Engineering Surfaces", PhD Thesis, London University.

WEBSTER, M.N., IOANNIDES, E. and **SAYLES, R.S.** (1985) "The Effect of Topographical Defects on the Contact Stress and Fatigue Life in Rolling Element Bearings", Proceedings of the 12th Leeds/Lyon Symposium.

WEBSTER, M.N. and **SAYLES, R.S.** (1986) "A Numerical Model for the Elastic Frictionless Contact of Real Rough Surfaces", Trans. ASME J. of Tribology, 108, 314-320.

WEBSTER, M.N., WEST, M.A. and **SAYLES, R.S.** (1986) "A Method of 3-Dimensional Topography Measurement and Analysis on Arcuate Surfaces", Wear, 109, 385-399.
- *Reproduced as Appendix B*

WEST, M.A. and **SAYLES, R.S.** (1987) "A 3-Dimensional Method of Studying 3-Body Contact Geometry and Stress on Real Rough Surfaces", Proceedings of the 14th Leeds/Lyon Symposium.
- *Reproduced as Appendix D*

WEST, M.A., WEBSTER, M.N. and **Sayles, R.S.** (1987) "A New Method for Rough Surface Contact Analysis", Proc Instn. Mech. Engrs, 'Tribology - Friction, Lubrication and Wear fifty years on', 945-955.

- Reproduced as Appendix C

WHITEHOUSE, D.J. and **ARCHARD, J.F.** (1970) "The Properties of Random Surfaces of Significance in their Contact", Proc. Roy. Soc. Lond., A316, 97-121.

WHITEHOUSE, D.J. and **PHILLIPS, M.J.** (1978) "Discrete Properties of Random Surfaces", Phil. Trans. Roy. Soc. Lond., A290, 267-298.

WHITEHOUSE, D.J. and **PHILLIPS, M.J.** (1982) "Two-Dimensional Discrete Properties of Random Surfaces", Phil. Trans. Roy. Soc. Lond., A305, 441-468.

WILLIAMSON, J.B.P. and **HUNT, R.T.** (1972) "Asperity Persistence and the Real Area of Contact Between Rough Surfaces", Proc. Roy. Soc. Lond., A327, 147-157.