THE KINEMATIC ANALYSIS AND METROLOGY OF CYLINDRICAL WORM GEARING

by

Jun HU, BEng (Hons), MEng

This thesis is submitted in fulfilment of the requirements for the Degree of

Doctor of Philosophy

Department of Mechanical, Materials and Manufacturing Engineering
University of Newcastle upon Tyne
UK

July 1997

ABSTRACT

Worm gearing is very widely used, especially in heavy industry, but due to the complexity of worm gear geometry, worm gear research has lagged behind that for spur and helical gears. In the last decade, however, the potential for significant improvement in worm gearing has dramatically increased: computers have given greater freedom to analyse worm gearing; CNC machines make it possible to aim for optimised worm gear geometries with very high accuracy and the development of synthetic lubricants has substantially improved lubrication conditions. In the UK, over the last few years, research effort in the field of worm gearing has increased considerably.

As a part of this recent activity in the UK, the author has been involved mainly in developing the analytical mechanics and metrology of worm gears. A method for the generalised 3D non-elastic worm gear mesh analysis and associated software have been developed and worm wheel metrology software has been implemented on a CNC measuring machine in the UK National Gear Metrology Laboratory, to allow, for the first time, analytical measurement of worm wheel tooth flanks. Combination of the mesh analysis software and CNC measurement of worm wheels has assisted in the design and manufacture of worm gears with modified tooth profiles.

Two methods of 3D non-elastic worm gear analysis have been developed for conjugate action and non-conjugate action respectively. The conjugate analysis determines the lines of contact, sliding and rolling velocities, limitations of the working area (the envelope of contact lines on a worm surface and singularities on a wheel surface), principal relative curvatures and the orientations of contact lines. It is based on the B-matrix method [Zhang and Hu, 1989]. The non-conjugate analysis predicts entry and exit gaps, contact ratio, wear marking on the worm flank, instantaneous contact topology on all the engaged tooth flanks, total contact area, contact pattern and transmission error. This is based on numerical simulation of the actual worm gear running process under no-load.

Although the non-elastic analysis models have been designed for any type of worm gearing, and have been used to study Cavex (ZC) wormgears and the meshing of a ZA

worm with a helical gear, most of the work has been on involute (ZI) worm gearing, since this is, by far, the most commonly produced type in the UK.

This thesis presents the work as follows:

1) The development of the B-Matrix kinematic method for conjugate analysis.

The B-Matrix method, presented in chapter 2, elegantly simplifies the derivation and calculation procedures, since the geometric parameters and the motional parameters can be arranged in separate matrices. As a result, the models can be applied to different geometries and coordinate systems with no need for further difficult derivations. The method leads to an easier way of integrating the theory of various types of worm gearing into compact generalised models. It is much more convenient and reliable to let the computer formulate and solve matrix equations numerically, treating each matrix as a simple variable, than to develop analytically the corresponding long tedious non-linear equations.

2) The development of mathematical equations to allow CNC measuring machines to measure cylindrical worm wheels with respect to their mating worms.

The measurements are 3-dimensional and absolute, in the sense that the results are the deviations from the theoretical geometries rather than comparative measurements relative to a (necessarily imperfect) master worm wheel. The measurement theory has been implemented on a particular CNC measuring machine. This is presented in chapters 3 and 5.

3) The development of the non-conjugate analysis.

The fundamental basis of the non-conjugate analysis presented in this thesis is to rotate the worm wheel to bring its tooth flank into contact with the worm flank at each given angle of worm rotation, so that the no-load transmission error and gap contours can be determined. This method is suitable for both cylindrical and globoidal worm gears, since the rotation angles of worm and wheel are used to simulate the running process directly.

The method also allows the wheel tooth flank to be obtained either by conjugate analysis of the hobbing process, or by analytical measurements or other methods (for example, when a theoretically-generated involute helical gear is used to mesh with a worm). This work is presented in chapter 4.

4) implementation of the non-elastic analysis theory.

The non-elastic analysis software has been written for personal computers. In addition, dimensional calculations specified by BS 721 and commonly used hob design methods have been added to the non-elastic analysis software to increase user-friendliness. The software has been used to investigate the effects on the worm gear contact and performance of machining errors and profile deviations or modifications. The structure of this analysis software allows for the inclusion of new modules for other types of worm gearing without in any way disturbing the integrity of the program's existing abilities. The non-elastic analysis software is user-friendly with a "Windows" graphical user interface. Software reliability and error tolerance have been of particular concern during program development. This work is presented in chapter 5.

5) The software has been thoroughly validated against other published results and/or actual production.

The software has been used extensively for both research and commercial purposes, and the user interface developed further in response to user feedback. Examples of these applications are given in chapter 7.

ACKNOWLEDGEMENTS

First of all I would like to thank Mr. D. A. Hofmann and Mr. J. A Pennell. Mr. Hofmann, Director of Design Unit, University of Newcastle upon Tyne, offered me the opportunity to undertake this project which is the basis of completing my thesis. To Mr. Pennell, my supervisor, I am grateful for his constructive technical advice and suggestions during this project. Their help has been invaluable.

Special thanks must also be given to Mr. R. C. Frazer and Mr. E. J. Myers, with whom I worked closely and discussed technical problems on many occasions. This has been very helpful to me.

I would also like to thank DTI, and the members of the BGA Worm Gearing Research Consortium (David Brown Radicon, Holroyd, Renold, Highfield, Express Lift, University of Huddersfield, University of Wales at Cardiff and Castrol) for their technical input to my thesis and/or financial support during this project.

Last but not least, I would like to extend my gratitude to all of my colleagues in the Design Unit, particularly, the secretaries Mrs. D. Graham, Mrs. J. McLean and Miss K. Stoker, the chief technician Mr. C. Kirby, and the other co-workers of this project: Mr. N. Baty, Mr. I. Sewell, Mr. N. Carr, and Mr. J. Orrell, for their kindness, friendliness and enthusiastic co-operation, who have all contributed to this project and made my stay here a memorable part of my life.

A special word of appreciation goes to Mr. M. Fish, PhD candidate, University of Huddersfield, with whom I have had an enjoyable and mutually beneficial collaboration.

On a more personal note, I am deeply indebted to my parents, my wife and my daughter. My gratitude to them is beyond words, so I dedicate this thesis to them.

To YouFen and TingTing

DECLARATION

This thesis consists of the original work of the author except where specific reference is made to the work of others, and has not been previously submitted for any other degree or qualification.

NOTATION

A_j - contact area round an individual grid point j,

A_{tot} - total contact area over all engaged tooth flanks,

B - B matrix,

$B_{\Phi 1}$ - partial derivative of B matrix with respect to Φ_1,

b_2, b_{eff} - worm wheel face width and worm wheel effective face width respectively,

C - centre distance,

D - distance from worm wheel middle face plane to the section in which profile errors are to be measured,

f_i' - single tooth pair transmission error,

F_i' - overlapped transmission error,

$g_{minf}^{(1)}$,

$g_{maxf}^{(1)}$ - unit vectors of principal directions on the worm surface in coordinate system S_f, $g_{minf}^{(1)}$ has the minimum curvature,

h - surface: h = 1 worm/hob, h = 2 worm wheel,

i - transmission ratio, $i=\Phi_2/\Phi_1$,

j - grid point on wheel flank, $j = 1, 2, \cdots, N$,

K - $K=p_z/(2\pi)$,

k_1 - worm tooth number, $k_1 = 1,2, \cdots, z_1$,

k_2 - worm wheel tooth number, $k_2 = 1,2, \cdots, z_2$,

$k_{min}^{(1)}$,

$k_{max}^{(1)}$ - principal curvatures on the worm surface along $g_{minf}^{(1)}$, $g_{maxf}^{(1)}$ respectively, $k_{min}^{(1)}$ is the minimum curvature

$k_n^{(1,2)}$ - maximum relative normal curvature of the mating worm and wheel surfaces,

m - (subscript) for coordinate systems: m = 0, 1, 2, f, p

m_x - worm axial module

M - meshing function

$M_{\Phi 1}$ - partial derivative of meshing function with respect to Φ_1,

$N_m^{(h)}$ - normal vector to flank $\sum^{(h)}$ represented in terms of components of coordinate system $S_m (O_m, x_m, y_m, z_m)$, h=1,2

$N_{xm}^{(h)}$,

$N_{ym}^{(h)}$,

$N_{zm}^{(h)}$, - components of normal vector to flank $\sum^{(h)}$ in coordinate system $S_m(O_m, x_m, y_m, z_m)$.

p_x - axial pitch of worm thread,

p_z - lead of worm thread,

R_1, R_2 - radial positions of points on worm and wheel flank respectively,

r_1, r_2 - worm and worm wheel reference radii, respectively,

r_{f1}, r_{a1} - worm root radius and outside radii, respectively,

$r_m^{(h)}$ - position vector of a point $P(u,\theta)$ on flank $\sum^{(h)}$ represented in terms of components of coordinate system $S_m (O_m, x_m, y_m, z_m)$.

S_m - coordinate system with origin O_m and coordinates (x_m, y_m, z_m), with m = 0,1, 2, f, p. S_m can be fully represented as $S_m (O_m, x_m, y_m, z_m)$.

s - planar curve which generates the worm helicoid surface by screw motion.

$x_m^{(h)}$

$y_m^{(h)}$

$z_m^{(h)}$ - components of position vector of $r_m^{(h)}$

T_{nm} - coordinate transformation matrix from S_m to S_n, m = 0,1,2,f,p, n = 0,1,2,f,p. For example, T_{21} represents a transformation matrix from coordinate system S_1 to S_2, i.e. from coordinates specified in S_1 to coordinates specified in S_2.

$V_m^{(h)}$ - velocity vector of a contact point on flank $\sum^{(h)}$, represented in terms of components of coordinate system $S_m (O_m, x_m, y_m, z_m)$

$V_{s1}^{(12)}$ - relative (sliding) velocity in coordinate system S_1, $V_{s1}^{(12)}=V_1^{(1)}-V_1^{(2)}$,

$V_{r1}^{(12)}$ - rolling velocity in coordinate system S_1, $V_{r1}^{(12)}=V_1^{(1)}+V_1^{(2)}$,

$V_e^{(12)}$ - entrainment velocity,

z_1 - number of worm starts,

z_2 - number of worm wheel teeth,

[V, -V] - lead measurement range,

W_{nm} - rotation matrix from from S_m to S_n,

δ_a - single tooth pair axial gap,

Δ_a - overlapped axial gap.

δ_n - single tooth normal axial gap.

Δ_n - overlapped normal gap.

u - curvilinear parameter of the worm/hob flank surface,

$\sum^{(1)}$ - worm/hob flank,

$\sum^{(2)}$ - wheel flank,

Φ_1 - angle of worm/hob rotation,

Φ_2 - angle of wheel rotation,

ω_1 - speed of worm/hob rotation,

ω_2 - speed of wheel rotation,

η - hob swing angle,

λ - angle of generator of convolute worm helicoid,

γ - worm/hob lead angle,

γ_b - base lead angle,

Ω - angle between the tangent to a contact line and worm surface principal direction $g_{minf}^{(1)}$.

θ - curvilinear parameter of the worm/hob flank surface,

Θ - angle between the normal to a contact line and worm surface principal direction $g_{minf}^{(1)}$.

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Background: worm gear applications and developments

1.1.1 Worm gear applications, characteristics and families

The worm gear set is a primary transmission device used to transmit torque and motion between crossed shafts. The angle between the two shafts is usually 90^0 but other angles may be used to satisfy particular requirements. A typical worm gear pair is shown in Figure 1.1.

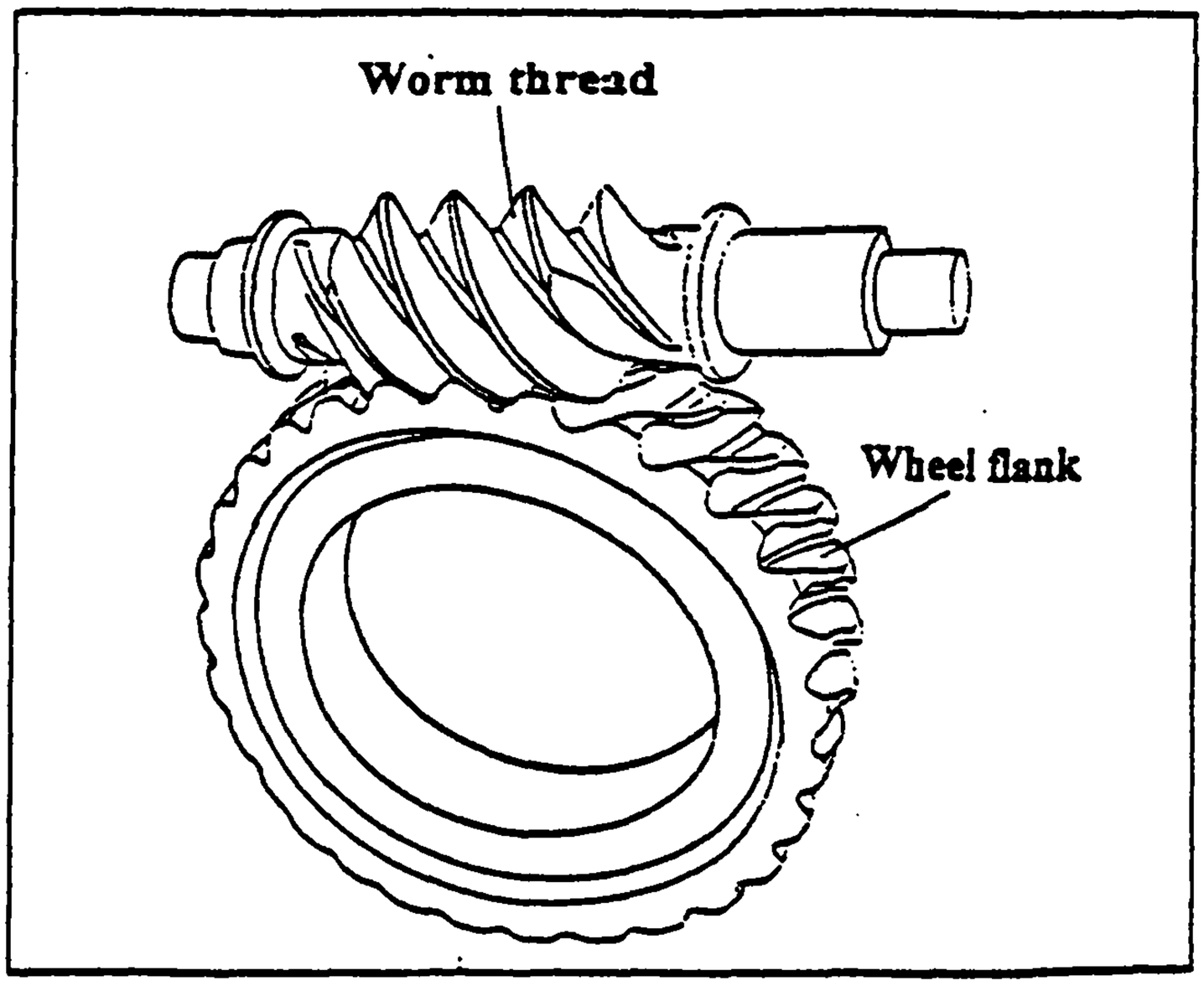


Figure 1.1 A Typical Worm Gear Pair

Worm gearing is used in applications to transmit powers from a few Watts up to 775 KW and torques from 1 Nm up to 350 KNm. Worm speeds can vary from 0 to 10,000 rpm,

depending on size, and the centre distance, from as little as 7.5 mm to 2.5 metres or more. Reduction ratios in the range from 1:1 to 400:1 are feasible. Use of worm gearing as a speed-increasing drive is possible but unusual. A large percentage of worms are single-start, but ten or more starts on the worm is not uncommon.

The worm gears have found a wide variety of applications and in 1984 [Anon-2, 1984] accounted for more than 28% of the total value of industrial gears sold in the UK. This is because of their special characteristics - good vibration damping characteristics, the ability to produce a high ratio in a single stage, simple construction, damage tolerance, compact angle drive layout, possible self-locking behaviour and, above all, low noise [Pennell, 1990].

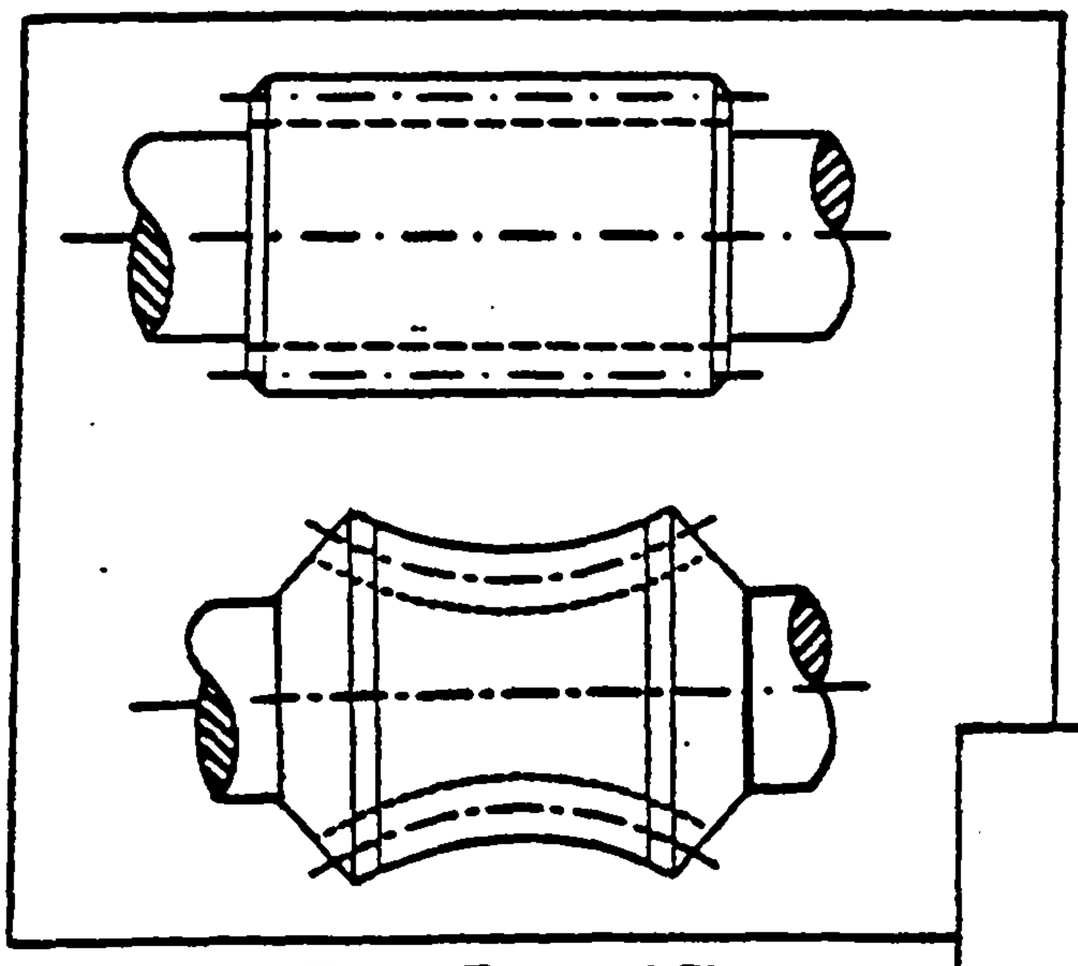


Figure 1.2 Worm External Shapes

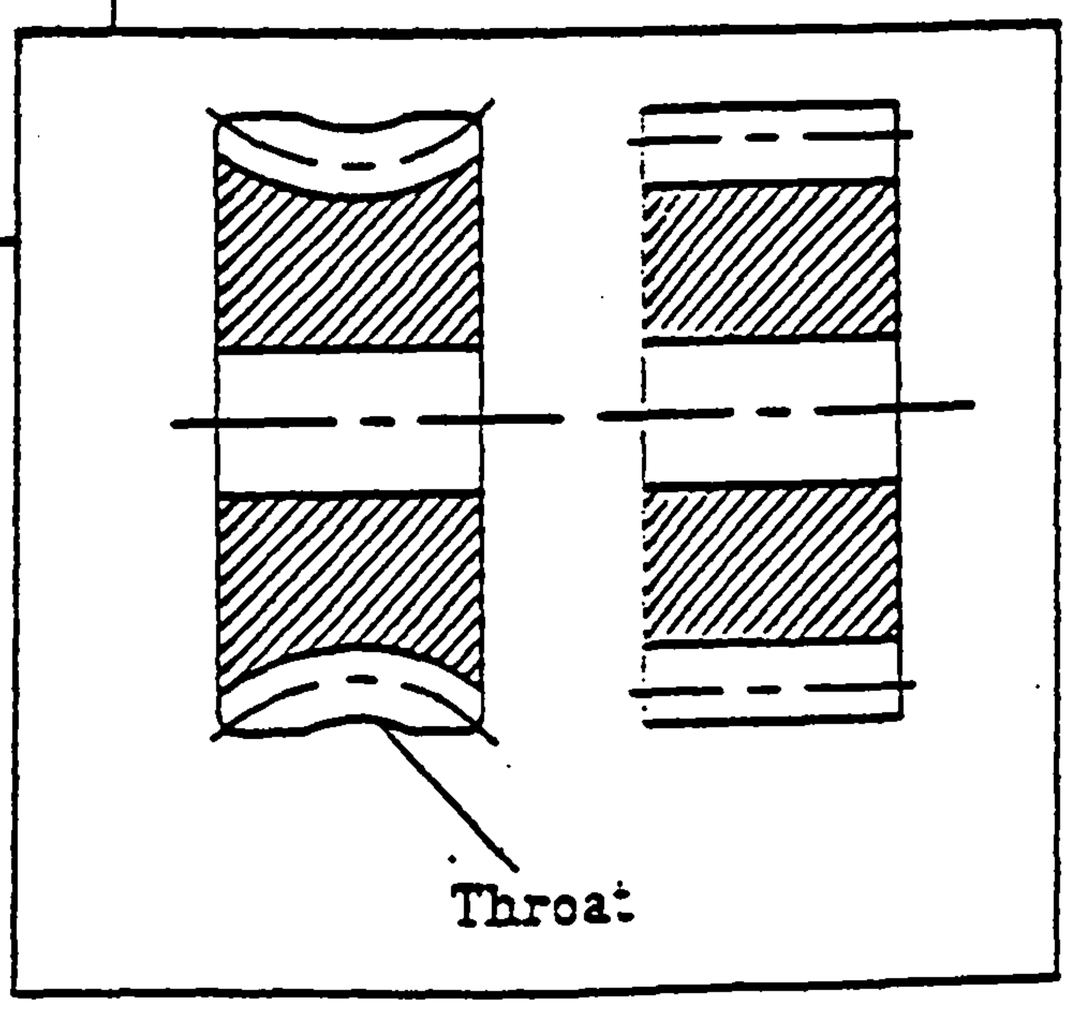


Figure 1.3 Wheel External Shapes

A worm's external shape may be a cylinder or it may be hourglass-like (Figure 1.2). The external shape of the meshing worm wheels may also be a cylinder or one with a (globoidal) throat recess (Figure 1.3). As a result of this, four options can be chosen to

combine the external shapes. However, worm gearing today is normally divided into two classes by the worm's external shape since commonly used worm wheels are almost exclusively globoidal (with the throat recess). The two classes are cylindrical (or single-enveloping) worm gear drives and globoidal (or double-enveloping) worm gear drives, as shown in Figures 1.4 and 1.5. Use of a globoidal worm encourages higher contact ratio and higher load capacity since more teeth are engaged.

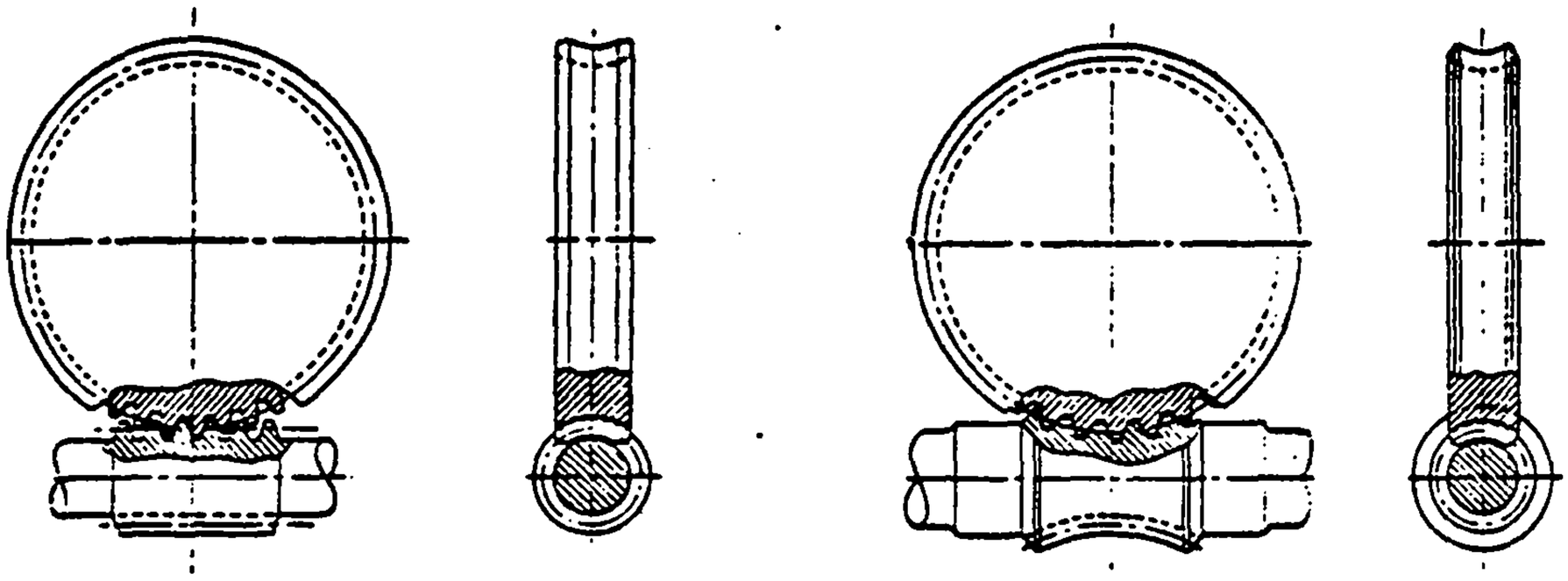


Figure 1.4 Schematic of Cylindrical and Globoidal Worm Gear Pairs

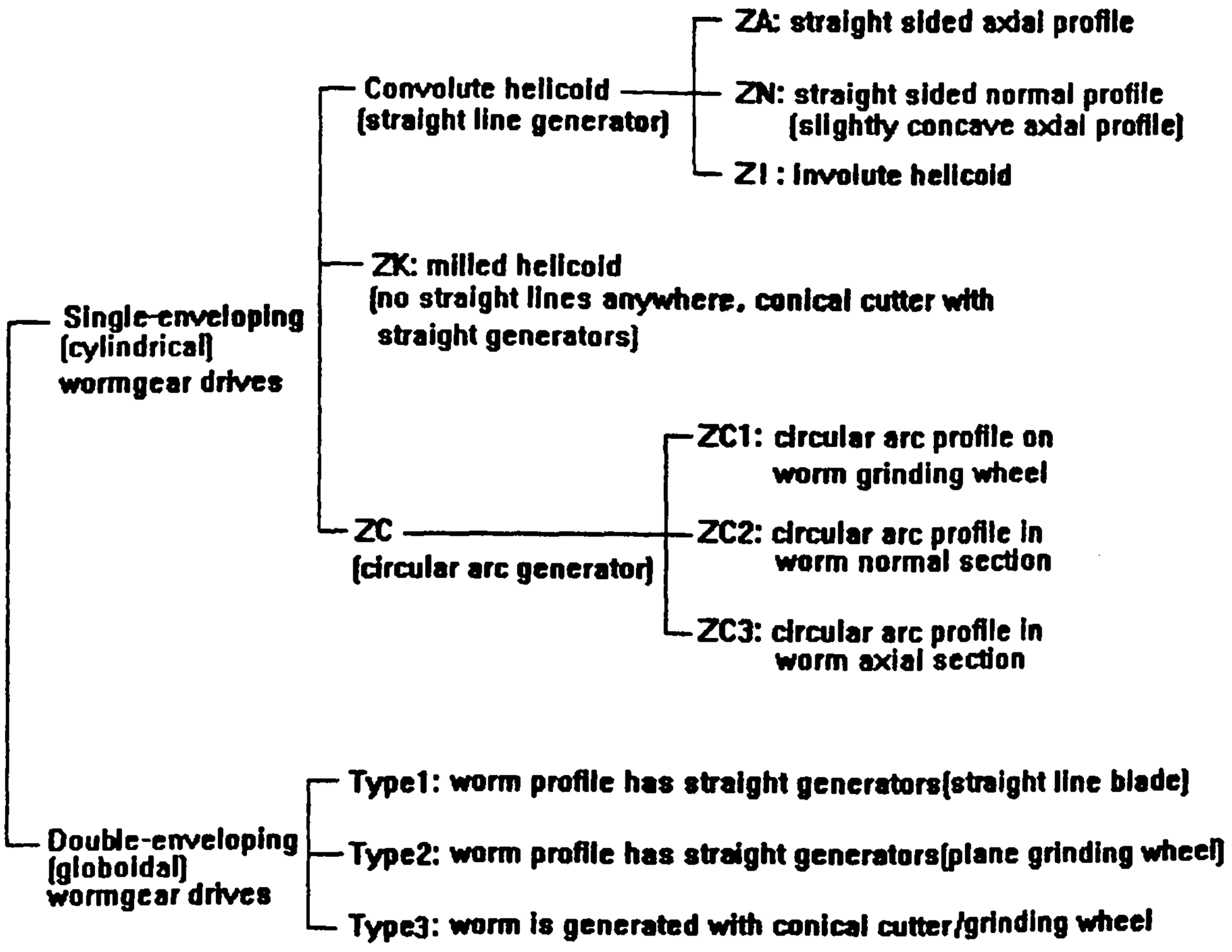


Figure 1.5 The Family of Common Worm gear Drives

Each class of worm drive falls into various types distinguished by the worm tooth profiles. Work on searching for new types of worm gears is still going on [Simon, 1994]. According to AGMA6022-C93, DIN3975 and also draft ISO standards, cylindrical worm tooth profiles can have ZA (straight sided axial profile), ZI (involute helicoid), ZN (straight sided normal profile), ZK (milled helicoid, or Klingelnberg helicoid), and ZC (or ZH, ZF, circular arc profile) tooth forms. ZA, ZI and ZN forms can all be produced with straight-line generators, and are categorized as convolute helicoids by Buckingham [Buckingham, 1949]. ZC forms are generated with circular arcs, although at least 3 different types of ZC worms are in use. ZK forms result from the use of conical cutters and nominally there is no straight line on a ZK helicoid. The family of common worm gears is given in Figure 1.5.

1.1.2 History of worm gear development

Worm gearing is of great antiquity, going back about 2100 years to Archimedes, who is generally acknowledged as its inventor. Archimedes used an (Archimedean) screw to rotate a toothed wheel. Subsequent development of the worm gearing principle progressed very slowly along conventional lines until about 500 years ago when Leonardo DaVinci evolved the double enveloping gear concept [Loveless, 1980]. The modern worm gearing industry and systematic research into worm gears started in the early 1920's (David Brown patented worm gears in 1910 and 1912) and, since then, many different types of worm gear have been developed.

In the UK, worm gear production is almost exclusively of ZI cylindrical worm gears. In the postwar period and through into the 1950's, the two major UK companies (David Brown Radicon and Renold-Holroyd-Crofts) dominated this field in Europe, with much practical research and product development [Hofmann,1993]. One result of this active industrial research and development was the British Standard for the rating of ZI worm gears, BS 721, first published in 1937, then revised in 1963 and 1983.

In the war and postwar period, up to the 1960's there was significant research in Germany by Niemann and his colleagues in this field [Hofmann,1993]. In the late 1940's, Niemann developed and patented cylindrical worm gears with concave flanks with which

he obtained excellent theoretical and experimental results. In the early 1950's, gear producer Flender A.G. acquired the patent rights and began commercial production of these drives, bringing worm gearing for the first time, into a competitive position in mainland Europe. At the same time the name CAVEX was coined for this type of gearing [Vos, 1996]. CAVEX worms are the main form of ZC worms.

During the same period, globoidal (double enveloping) worm gear drives were also developed. Historical records indicate that more than 200 years ago, in York, the famous clockmaker Henry Hindley made the first set of Hindley globoidal worm gears. In the early 1920's, Mr. Samuel L. Cone of Portsmouth, Virginia, significantly improved Hindley's globoidal worm gearing to produce the drive which presently carries his name - Cone Drive Double Enveloping Worm Gear Drives [Tang, 1991][South, 1995]. Since the 1950's several different types of globoidal worm gears have also been manufactured in Japan and China respectively.

Worm wheels were originally cut with a hob that had the same basic geometry as the mating worm. In this case it was expected that the worm gear would have conjugate action with line contact between the tooth flanks. Later on, however, deliberate mismatch between the hob and the worm was introduced into worm gear production to improve the worm gear performance. This is mainly because the mismatch makes it easier to get proper contact and lubrication conditions in the presence of unavoidable manufacturing errors and elastic and thermal deflections. In this case, worm gears are of non-conjugate action and have initially point contacts between the tooth flanks under no-load.

However, as is well known, these non-conjugate tooth flanks bed in during running under load, as a result of the relatively poor lubrication condition and localized contact, which gives rise to appreciable wear and/or plastic deformation of the softer worm wheel flanks. As a result, the worm parts of the worm wheel flanks become more nearly conjugate to the (relatively unworn) worm flanks, with an extended contact area, lower contact stresses and a greatly reduced wear rate. On some worm gears, running under constant load, the wear may cease altogether once the teeth are properly bedded in.

1.2 Review of previous worm gear research

1.2.1 Move from 2D to 3D modelling

Some early research work on worm gears was reported by Dudley & Poritsky in 1943. They showed how the worm tooth shape produced by a given cutter could be computed and conversely how to determine the cutter shape to produce a desired worm shape [Narayan and Corporation, 1995]. In 1949, systematic mathematical models of worm gearing mechanics were published in a book [Buckingham, 1949], which arose out of the investigations of ASME Special Research Committee on Worm Gears. The geometries and mechanics of ZA, ZN, ZI and ZK cylindrical worm gear drives were well defined by Buckingham at that time. However, Buckingham's worm gearing analysis is built on the concept that the gear tooth action between a worm and a wheel is conjugate and identical to that of a spur gear and a rack in each transverse plane. It is thus restricted to the 2D conjugate contact problems associated with worms and wheels whose axes are perpendicular. In the 1950's through into the 1970's there was a steady output of research into the theory of gearing, from specialised institutes of former USSR and other eastern European countries. Most of the analytical techniques developed were published in Litvin's book [Litvin, 1968] and further developed by him in the USA [Litvin, 1989][Litvin, 1994] . The significant advance presented was that the equation of meshing was introduced, which made it possible to apply the theory of gearing presented in fully 3D meshes, so that it can be applied to any type of worm gear.

1.2.2 Move towards more realistic analysis

Until the late 1980's, all worm gearing analysis was exclusively based on the assumption of conjugate action and line contacts between the worm and wheel tooth flanks. However, as mentioned before, in most cases, worm gears are of non-conjugate tooth form with point contacts under no-load before bedding-in. It was pointed out by Bagci(1987) that the contact stresses in the case of point contacts may be substantially higher [Litvin and Kin, 1992] so that it would be beneficial to consider more realistic models. In 1988 and 1989, Janninck and Colbourne published independently their

computerized simulations of non-conjugate cylindrical worm gear meshing. These two authors used different methods to determine the separation contours between the worm and wheel surfaces but also studied the effect of hob oversize and swivel angles on the separation contours [Narayan & Corporation, 1995]. In this decade, a few publications introduced concept of TCA (Tooth Contact Analysis] into cylindrical worm gearing to calculate bearing contact [Kin, 1990][Litvin, 1992][Fang, 1996]. Narayan and Corporation (1995) made use of Finite Element Analysis (FEA) to study the effect of hob geometry on load distribution, and the effect of load on the transmission error of cylindrical worm gearing. Similar progress was made for double-enveloping worm gears [Hiltcher & Guingand, 1994]. Tang (1991) applied Hertzian stress analysis and FEA to double-enveloping worm gears to predict the load distribution.

The above survey shows that researchers are now approaching the analysis of "real" worm gears. However, actual "real" worm gears are produced with errors and the "real" wheel flanks are finally re-generated by the worms during the bedding-in process. To produce optimised wheels accurately and obtain the shapes of "real" worm wheels, analytical measurements of actual worm wheels are crucial since there is, at present, no theoretical method for predicting the wheel wear that occurs during this bedding-in process. Only thus, when the actual geometry of real worms and wheels is known, can realistic prediction of meshing conditions be undertaken.

1.3 The objectives of this research project

1.3.1 The state of worm gear development

Worm gear research and development has not progressed as rapidly as that of spur and helical gears. As understanding of spur and helical gears has improved, there has been a move away from the use of relatively 'soft' gears, which bedded-in during initial running, to surface-hardened gears which are manufactured with extreme precision to the ideal running profile. This has resulted in a four-fold to five-fold increase in load capacity for a given size. Worm gears, in contrast, have not developed significantly over the same period [Hofmann, 1993]. The major reason for this slow development has been the

complex geometry of worm gearing and the attendant difficulties of analysing the mesh.

Although measurement of involute gear profiles is possible within ±1 μm [Beyer, 1996], worm wheels are still tested in industry using "blue" contact marking patterns. Thus, while modifications can be precisely applied to involute gears to improve their performance dramatically, worm gears are still heavily reliant on the result of "bedding-in". No standard for profile and lead accuracy has ever been produced for worm wheel quality control since the accuracy of worm wheel manufacturing could not be measured absolutely.

Worm gear technology, with a case-hardened worm and a phosphor bronze wheel, dates back to an era when manufacturing tolerances were much greater than to-day, and when it was impossible to measure complex spatial geometries. Since the worm-wheel geometry could not be measured, the final wheel geometry could only be wildly guessed. The same design philosophy is still applied today, so that it is difficult to apply theoretical mesh analysis with confidence. EHL (Elasto-hydrodynamic Lubrication) analysis and theoretical optimization of tooth forms of worm gearing are also of doubtful value if the worm wheel geometry is uncertain.

Worm gearing is therefore limited in two respects:

1) The theoretical analysis of worm gearing is based on the assumption that the wheel will be cut or "bedded-in" to a known geometry. This has been shown to be very unreliable assumption (see below).

2) The wheel geometry is not optimised for lubrication and load conditions, but generated fairly randomly by the bedding-in process, so that the best performance is unlikely to be achieved.

The situation in relation to the theoretical modelling of 'real' worm gearing is thus far from satisfactory.

1.3.2 Potential for improvement of worm gearing

The potential for improvement of worm gearing has increased significantly in recent years. Computers have given the power to apply modern 3D meshing theory to worm gearing of any type of geometry, and (potentially) optimise their design. CAM (computer-aided manufacturing) and CAI (computer-aided inspection) have given the possibility of optimising their production and inspecting them very accurately.

In the last decade, the potential accuracy achievable with modern CNC worm grinders has also improved dramatically. The development of full CNC grinding wheel dressers gives complete freedom in the choice of grinding wheel profiles, so that any desired worm(hob) profile can now be form ground with extreme accuracy. CNC hobbing machines are also capable of hobbing high accuracy worm wheels, which potentially allows worm wheels with optimised profiles to be generated. Coordinate measuring machines are available and software for measuring complex worm and wheel geometries can be developed.

The combination of modern gear tooth meshing analysis, and the experience gained from application of elastic mesh analysis and elastohydrodynamic lubrication theory to spur and helical gears, means that is now potentially possible to obtain better fundamental understanding of worm gearing and optimize worm gear profiles. The combination of accurate, flexible coordinate measuring machinery, modern manufacturing capability, and better lubricants opens up the possibility for physically achieving the optimised wormgear profiles with desired accuracy. The future redevelopment of worm gears could thus provide more accurate transmission, higher load capacity and higher efficiency, at effectively lower cost, without sacrificing the intrinsic advantages of low noise and high stage ratio.

1.3.3 Recent worm gear research activity in the UK

The BGA Research Framework published in 1990 (Transmission Engineering Research for the UK) has encouraged an interest in worm gear research at UK universities, and four major projects of research have been carried out.

1) "Improving the Stress Analysis of Worm Gearing" involved development of an elastic mesh model for worm gears to provide the same accuracy of analysis as is possible for spur and helical gears (University of Newcastle upon Tyne, EPSRC Grant NO: GR/G57932);

2) "Worm Gear Design: Elastohydrodynamics" involved a study of the lubrication of worm gears using EHL theory to develop a procedure for calculating oil film thickness and friction coefficients (University of Wales, Cardiff, EPSRC Grant NO: GR/H41546);

3) "The Analysis and Measurement of Transmission Error in Precision Worm Gears" involves the development of a more fundamental understanding of the causes and means of reducing transmission error (University of Huddersfield, DTI Reference No: REH/LMS/600.1);

4) "Improving Worm Gear Performance" involves an attempt to develop new worm gear geometries with higher load capacity, using cheaper materials of construction (University of Newcastle upon Tyne, DTI Reference No: MP20/113).

The work described in this thesis was carried out in connection with projects (1), (2) and (4).

1.3.4 The objectives of this research

The kinematic techniques for the analysis of models of worm gearing and for worm gear metrology both needed to be developed to support the activities mentioned above. Complete flexibility in worm gear modelling was necessary in the initial stages to allow for future requirements. Only with precise absolute measurement capability is it possible to utilise fully the accuracy potential of modern gear grinding and hobbing machines, since without accurate measurement, accurate manufacture and optimization are impossible [Hofmann,1993].

Therefore the objectives of this PhD project were:

1) To develop a general 3D non-elastic analysis technique for worm gearing;

2) To develop a method of worm gear metrology for use with CNC measuring machines;

3) To apply the theory and software developed to enable a particular coordinate gear measuring machine to measure ZI (involute form) worm gears and to assist in better understanding, practical investigation and further research into topics (1), (2) and (4) above.

CHAPTER 2

CONJUGATE ANALYSIS - HOBBING AND BEDDING-IN PROCESSES

2.1 Introduction to the B-matrix Method

The conjugate analysis described herein can be used to model the worm wheel flank generation processes during hobbing or bedding-in. Although the mechanism of the bedding-in process is still not quite clear to the author, measurement of bedded-in worm wheels shows that the worm part of the wheel flank is very nearly conjugate to the worm, as might be expected if the worm is regarded as equivalent to a rigid abrasive tool during bedding-in.

Many types of worm tooth geometry can be found in application and work is still going on to find new ones. The use of flexible mathematical models capable of analysing all types of worm gearing is thus very beneficial. The B-matrix method is a new approach to generalised 3D kinematic models of gearing which can potentially be used for writing kinematic models for any type of worm gear [Zhang and Hu, 1989]. The models are represented in the form of matrices. The major advantage of the B-matrix models over others is that the geometric parameters and the motional parameters appear in separate matrices. Therefore, the equations can be applied to different geometries without further derivations. Furthermore, dealing with matrix models in a computer is very convenient, because the individual matrices are of simple form and easily derived, so that the computer can be left to carry out the multiple matrix operations without the need for complex algebraic analysis.

The B-matrix method is developed from Litvin's proposal [Litvin, 1968] that the necessary condition for the existence of an envelope of a family of surfaces can be represented in

the form

$$N \cdot V^{(12)} = 0$$

where N is the normal to the generating surface, $V^{(12)}$ is the relative (sliding) velocity of surface 1 with respect to surface 2. For a given generating surface, N is a known geometric vector but $V^{(12)}$ contains both geometric and motional parameters. In most formulations of the meshing equations, $V^{(12)}$ needs to be derived for each individual generating surface even if the motional parameters and the coordinate systems for two different gear pairs are the same. This is inconvenient, tedious and prone to error.

The B-matrix method describes the necessary condition for the existence of an envelope of a family of surfaces in the form

$$N \cdot B \cdot r = 0$$

Where N denotes the normal vector to the generating surface as before and r denotes the position vector of the generating point on that surface. Both N and r depend only on the geometric parameters of the generating surface. B is a matrix containing only the motional parameters and, is therefore, independent of the geometry of the generating surface. Once a coordinate system is set up, change of generating tooth geometry does not cause any change in the B matrix.

2.2 Coordinate Systems

To derive the kinematic models for both conjugate and non-conjugate analysis, the set of Cartesian coordinate systems shown in Figure 2.1 is set up. The system $S_m(O_m, x_m, y_m, z_m)$ represents a coordinate system with origin O_m and coordinates (x_m, y_m, z_m), with m = 0,1, 2, f, p. $S_f(O_f, x_f, y_f, z_f)$ is the fixed system, rigidly connected to the frame with x_f along common line of centres, y_f parallel to the wheel axis, and z_f in the wheel middle plane. Moving system $S_1(O_1, x_1, y_1, z_1)$ is rigidly connected to the worm/hob and rotates

with it around axis z_1. It coincides with the stationary system S_0 (O_0, x_0, y_0, z_0) when the worm/hob rotation angle $\Phi_1=0$. Moving system $S_2(O_2, x_2, y_2, z_2)$ is fixed in the worm wheel, with z_2 along the worm wheel axis. It coincides with the stationary system S_p (O_p, x_p, y_p, z_p) when the worm wheel rotation angle $\Phi_2=0$.

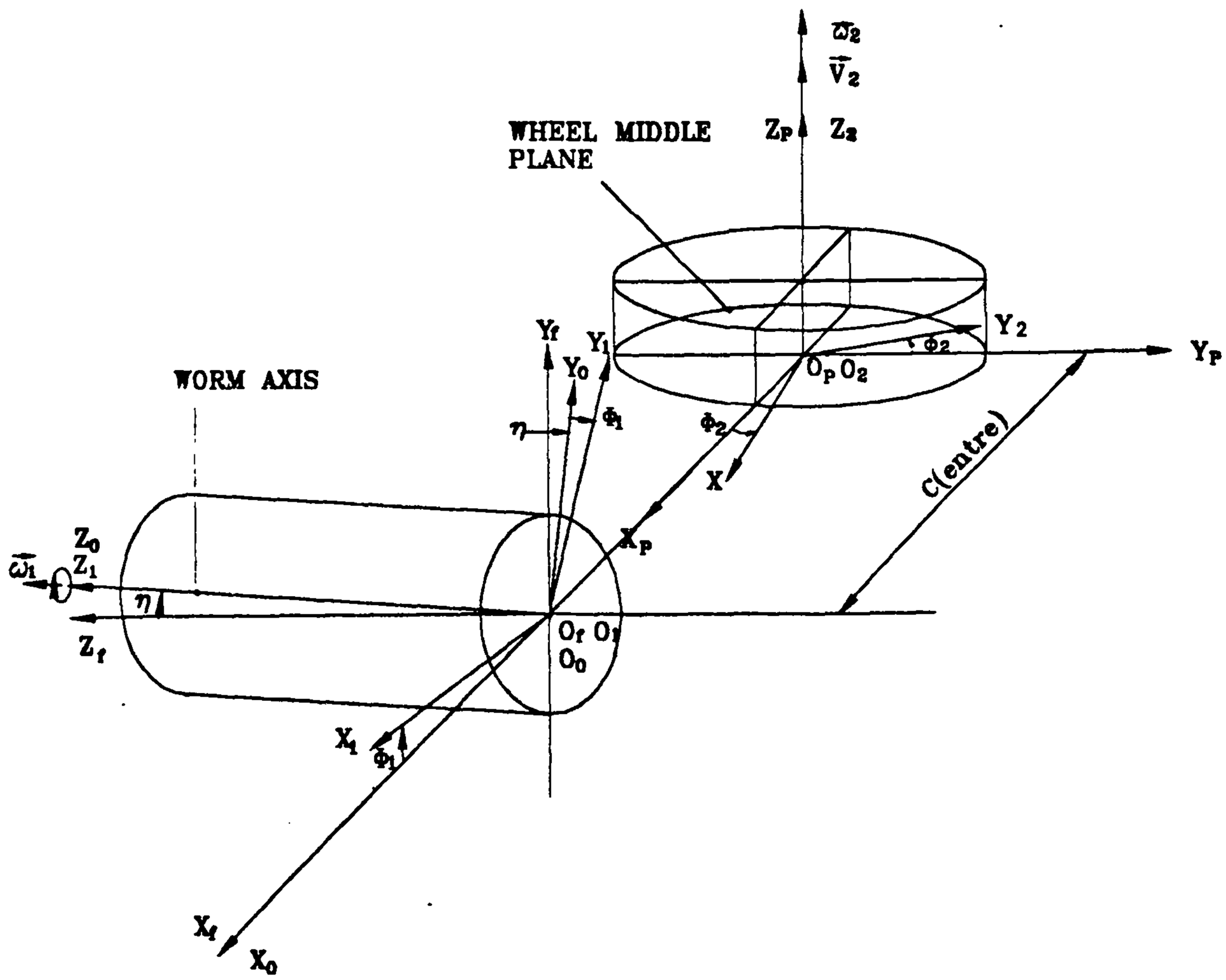


Figure 2.1 Coordinate Systems

2.3 The B-Matrix Kinematic Models

The main kinematic problems of the theory of gearing are:

(i) Determining generated surfaces as the envelope of a family of generating surfaces.

(ii) Determining lines of contact on conjugate surfaces.

(iii) Defining the limits for no undercutting of the generated surface during generation.

(iv) Calculation of the relative normal curvatures and the sliding and rolling velocities at contact points of the mating surfaces.

(v) Determining the surfaces of tools used to manufacture gears.

(vi) Simulation of mismatched (non-conjugate) gear drives, including prediction of the path of contact, entrainment velocity, contact gap contour and transmission error.

(vii) Prediction of the effects of manufacturing errors on no-load performance.

The first four problems are solved in this chapter and the last three are solved in chapters 4 and 5.

2.3.1 The Worm Surface and its Normal

In common practice, only the worm wheel hob has conjugate action with the wheel during the hobbing process. However, the hob is virtually a worm. Therefore, the term " worm" is used in this chapter to represent either a worm or a worm wheel hob, whichever has conjugate action with the worm wheel.

Since it is necessary to transform position vectors repeatedly between the different sets

of coordinate systems shown in Figure 2.1, homogeneous coordinates are used throughout this thesis. The homogeneous coordinates of a position vector in 3D space are thus determined by 4 numbers (x, y, z, 1) and a normal /velocity vector by (n_x, n_y, n_z, 0) / (v_x, v_y, v_z, 0), so that the corresponding coordinate translational and rotation matrices are then 4×4 square matrices. Matrix representation of the coordinate transformation only needs multiplication of matrices if position vectors are determined by homogeneous coordinates.

The position vector of a general point on worm surface $\Sigma^{(1)}$ is thus given in parametric form by:

$$r_1^{(1)}(u,\theta) = \begin{pmatrix} x_1^{(1)}(u,\theta) \\ y_1^{(1)}(u,\theta) \\ z_1^{(1)}(u,\theta) \\ 1 \end{pmatrix} \quad \ldots\ldots \quad (2.1)$$

$$r_1^{(1)}(u,\theta) \in C^2$$

$$\frac{\partial r_1^{(1)}}{\partial u} \times \frac{\partial r_1^{(1)}}{\partial \theta} \neq 0 \qquad (u, \theta) \in G$$

The superscript (1) denotes the worm surface $\Sigma^{(1)}$ and the subscript 1 denotes that the vector is represented in terms of components of coordinate system $S_1(O_1, x_1, y_1, z_1)$. u and θ are curvilinear parameters. Symbol C^2 denotes that $x_1^{(1)}(u,\theta)$, $y_1^{(1)}(u,\theta)$, $z_1^{(1)}(u,\theta)$ have continuous derivatives to the second order. $(u,\theta) \in G$ denotes that the curvilinear parameters belong to the area G which is an open region (excluding part of its boundary).

The normal vector to the worm surface can be obtained from:

$$N_1^{(1)} = \frac{\partial r_1^{(1)}}{\partial u} \times \frac{\partial r_1^{(1)}}{\partial \theta} = \begin{pmatrix} N_{1x}^{(1)}(u,\theta) \\ N_{1y}^{(1)}(u,\theta) \\ N_{1z}^{(1)}(u,\theta) \\ 0 \end{pmatrix} \quad \ldots\ldots \quad (2.2)$$

where the element 1 in the fourth row of the position vector $r_1^{(1)}$ and the element 0 in the fourth row of the normal vector $N_1^{(1)}$ are used to create the homogeneous vectors which are conformable for matrix multiplication.

2.3.2 Translational and Rotational Matrices

The position coordinate transformation matrix from system S_m to system S_n can be represented in the form

$$T_{nm} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & b_{14} \\ a_{21} & a_{22} & a_{23} & b_{24} \\ a_{31} & a_{32} & a_{33} & b_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Then the cooresponding rotation matrix for other vectors is

$$W_{nm} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The coordinate transformation and rotation matrices referring to the coordinate systems in Figure 2.1 are as follows.

$$T_{01}=\begin{bmatrix} cos\phi_1 & -sin\phi_1 & 0 & 0 \\ sin\phi_1 & cos\phi_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_{f0}=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & cos\eta & sin\eta & 0 \\ 0 & -sin\eta & cos\eta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_{pf}=\begin{bmatrix} 1 & 0 & 0 & C \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_{2p}=\begin{bmatrix} cos\phi_2 & sin\phi_2 & 0 & 0 \\ -sin\phi_2 & cos\phi_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_{21}=T_{2p}\cdot T_{pf}\cdot T_{f0}\cdot T_{01} \quad \ldots\ldots \quad (2.3)$$

$$T_{10}=T_{01}^{-1} \quad T_{f0}=T_{0f}^{-1} \quad T_{fp}=T_{pf}^{-1} \quad T_{p2}=T_{2p}^{-1} \quad T_{12}=T_{21}^{-1} \quad \ldots\ldots(2.4)$$

$$W_{21}=W_{12}^{T} \quad \ldots\ldots(2.5)$$

The superscipt -1 denotes the matrix inverse operation and the superscript T denotes the matrix transpose operation. T_{21} is the coordinate transformation matrix from coordinate system S_1 To S_2. W_{12} is the coordinate rotation matrix from S_2 to S_1. From (2.3) and (2.4) we obtain, after carrying out the matrix multiplications:

$$T_{21} = \begin{bmatrix} cos\phi_1 \cdot cos\phi_2 + sin\phi_1 \cdot sin\phi_2 \cdot sin\eta & cos\phi_2 \cdot sin\phi_1 + sin\phi_1 \cdot cos\phi_1 \cdot sin\eta & -sin\phi_2 \cdot cos\eta & C \cdot cos\phi_2 \\ -sin\phi_2 \cdot cos\phi_1 + cos\phi_1 \cdot sin\phi_1 \cdot sin\eta & sin\phi_1 \cdot sin\phi_2 + cos\phi_2 \cdot cos\phi_1 \cdot cos\eta & cos\phi_2 \cdot cos\eta & -C \cdot sin\phi_2 \\ sin\phi_1 \cdot cos\eta & cos\phi_1 \cdot cos\eta & sin\eta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \ldots \quad (2.6)$$

$$W_{12} = \begin{bmatrix} cos\phi_1 \cdot cos\phi_2 + sin\phi_1 sin\phi_2 \cdot sin\eta & -sin\phi_2 \cdot cos\phi_1 + cos\phi_1 \cdot sin\phi_1 \cdot sin\eta & sin\phi_1 \cdot cos\eta & 0 \\ -cos\phi_2 \cdot sin\phi_1 + sin\phi_1 \cdot cos\phi_1 \cdot sin\eta & -sin\phi_1 \cdot sin\phi_2 + cos\phi_1 \cdot cos\phi_2 \cdot cos\eta & cos\phi_1 \cdot cos\eta & 0 \\ -sin\phi_2 \cdot cos\eta & cos\phi_2 \cdot cos\eta & sin\eta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \ldots \quad (2.7)$$

2.3.3 The B Matrix and its Partial Derivative with Respect to Φ_1

For gearing which meshes at constant ratio i= Φ_2/Φ_1, the B-matrix is defined as

$$B(\phi_1, \phi_2) = B(\phi_1, i) = W_{12} \cdot \frac{\partial T_{21}}{\partial \phi_1} \quad \ldots \quad (2.8)$$

where T_{21} is the coordinate transformation matrix from coordinate system S_1 To S_2 and W_{12} is the rotation matrix from S_2 to S_1.

The B matrix is the basic matrix of the new kinematic method used in this thesis, and it appears throughout the calculations of conjugate analysis. This matrix and its derivative with respect to Φ_1 ($B_{\Phi 1}$) are given by equations (2.9) and (2.10) respectively.

$$\boldsymbol{B} = \begin{bmatrix} 0 & i\cos\eta - 1 & -i\sin\eta\cos\phi_1 & -iC\sin\phi_1\cos\eta \\ 1 - i\cos\eta & 0 & i\sin\eta\sin\phi_1 & -iC\cos\phi_1\cos\eta \\ i\sin\eta\cos\phi_1 & -i\sin\eta\sin\phi_1 & 0 & iC\sin\eta \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \ldots\ldots \quad (2.9)$$

$$\boldsymbol{B}_{\phi 1} = \begin{bmatrix} 0 & 0 & i\sin\eta\sin\phi_1 & -iC\cos\phi_1\cos\eta \\ 0 & 0 & i\sin\eta\cos\phi_1 & iC\sin\phi_1\cos\eta \\ -i\sin\eta\sin\phi_1 & -i\sin\eta\cos\phi_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \ldots\ldots \quad (2.10)$$

The matrices B and $B_{\Phi 1}$ have the following properties:

(1) Both contain only the motional variable Φ_1, as mentioned before. Once the coordinate systems (the hobbing settings and the worm gear axis positions) are decided, both matrices remain unchanged regardless of which type of worm gearing is to be analysed.

(2) All the elements in the fourth rows of both matrices are zero.

(3) The top left 3×3 sub matrices in both B and $B_{\Phi 1}$ are skew-symmetric.

These properties provide good guides for checking the B matrix for any particular set of coordinate systems.

It is worth mentioning that the B-matrix method is still valid with variable transmission ratio [Zhang and Hu, 1989], but only constant ratio needs to be considered for the conjugate analysis described in this thesis.

2.3.4 Equation of Meshing

The equation of meshing is used to determine contact points on the worm surface during the conjugate action. Only those points on the worm surfaces which satisfy this equation will take part in generating the worm wheel surface. As will be proved in section 2.4.2, the equation of meshing is

$$M=N_1^{(1)}\cdot B\cdot r_1^{(1)}=0 \qquad \ldots\ldots \quad (2.11)$$

Since the normal vector $N_1^{(1)}$ and the position vector $r_1^{(1)}$ are functions of the surface parameter u and θ, while B depends on i and Φ_1, it follows that the equation of meshing gives $M = M(u, \theta, i, \Phi_1) = 0$.

2.3.5 Generated Worm Wheel Surface and its Normal

Once the position vector of a contact point on a worm surface has be found by solving the equation of meshing (2.11), the corresponding point on the wheel flank can be obtained by coordinate transformation from the coordinate system $S_1(O_1, x_1, y_1, z_1)$, rigidly fixed to the worm surface, into the coordinate system $S_2(O_2, x_2, y_2, z_2)$, rigidly fixed to the wheel surface. This is because the position vector $r_1^{(2)}$ and normal vector $N_1^{(2)}$ of a contact point on the generated surface $\Sigma^{(2)}$ must be the same as those of the corresponding point on the generating surface $\Sigma^{(1)}$.

The position vector $r_2^{(2)}$ and normal vector $N_2^{(2)}$ of the worm wheel surface can thus be simply expressed by equations (2.12).

$$\begin{cases} M=N_1^{(1)}\cdot B\cdot r_1^{(1)}=0 \\ r_2^{(2)}=T_{21}\cdot r_1^{(1)} \\ N_2^{(2)}=W_{21}\cdot N_1^{(1)} \end{cases} \qquad \ldots\ldots \quad (2.12)$$

2.3.6 Relative (sliding), Rolling and Entrainment Velocities

The relative (sliding) and rolling velocities are defined in this thesis as $V_s^{(12)} = V^{(1)} - V^{(2)}$ and $V_r^{(12)} = V^{(1)} + V^{(2)}$ respectively. They are of important concern in studies of lubrication, wear and the heat generated in worm gear drives [Errichello, 1991]. The relative (sliding) and rolling velocity can both be calculated from the same equation in the B matrix method, but with the transmission ratio i assigned different signs. As proved below in sections 2.4.2 and 2.4.3, with a positive sign assigned to i as in equation (2.13), the equation will determine the relative (sliding) velocity, while with a negative sign assigned to i as in equation (2.14), the equation will determine the rolling velocity, so that:

$$V_{s1}^{(12)} = V_1^{(1)} - V_1^{(2)} = B(i) \cdot r_1^{(1)} \cdot \omega_1 \qquad \ldots\ldots \quad (2.13)$$

$$V_{r1}^{(12)} = V_1^{(1)} + V_1^{(2)} = B(-i) \cdot r_1^{(1)} \cdot \omega_1 \qquad \ldots\ldots \quad (2.14)$$

The entrainment velocity $V_e^{(12)}$ is required for EHL analyses, and is the mean velocity of the two mating surfaces relative to the contact point. Thus

$$V_e^{(12)} = 0.5 \cdot (V^{(1)} + V^{(2)}) - V_c = 0.5 \cdot V_r^{(12)} - V_c$$

where V_c is the velocity of the contact point in space, in other words the absolute velocity of the contact point along the path of contact. V_c can be easily obtained numerically with sufficient accuracy from this definition. It needs to be noticed, however, that, for EHL analysis, the entrainment velocity must be calculated over the entire wheel surface instead of just at the contact point. Details about the entrainment velocity are not required for the work described in this thesis.

2.3.7 Envelope of Contact Lines on a Worm Flank - (conjugate region on a worm flank)

Practice has revealed that heavy wear can often be observed on both the worm and the hob flanks. Hobs need to be reground after a certain period of use, so it is helpful to predict where the wear will appear, and where the edge of the working area will be

located on the worm/hob flank. This requires the calculation of the envelope of all the contact lines on a worm/hob surface. As proved in section 2.4.4, the envelope of all the contact lines which satisfy equation of meshing (2.11) is given by equation (2.15).

$$M_{\phi_1}=N_1\cdot B_{\phi_1}\cdot r_1^{(1)}=0 \qquad \ldots\ldots\ldots \quad (2.15)$$

Simultaneous solution of equations (2.11) and (2.15) determines a curved line as a locus of the points on the generating worm surface $\Sigma^{(1)}$. This curved line is the envelope of all the contact lines on the worm flank, which is illustrated in Figure 2.2. This envelope can be interpreted kinematically as potentially the edge of the working area of a generating surface. However, it should be noticed that the actual boundary of the working area on a worm flank may also be associated with the mating wheel tip diameter and rim profile. Further consideration of this is given in chapter 4.

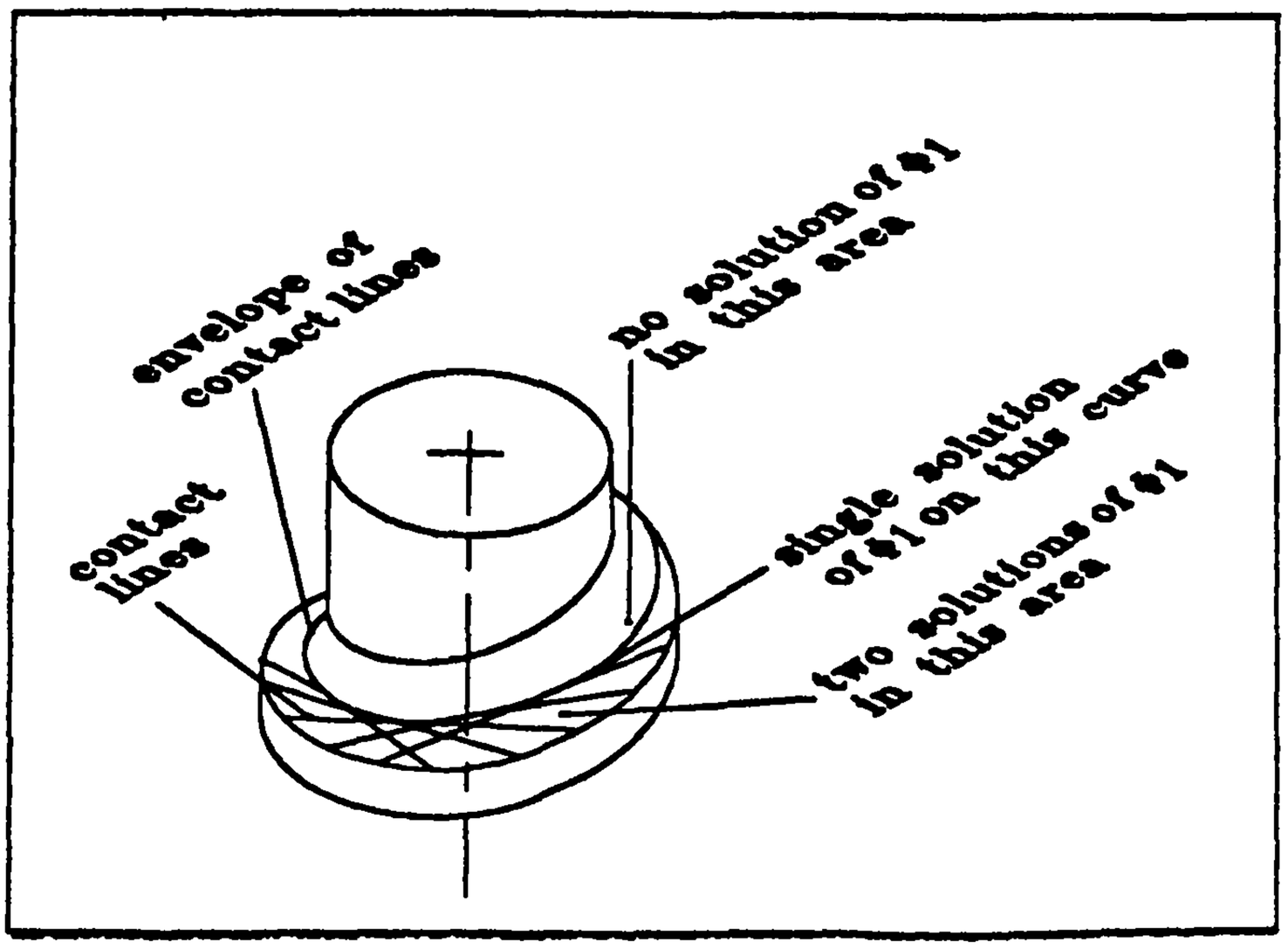


Figure 2.2 Envelope of Contact Lines on a Worm Flank

2.3.8 Singularities of a Worm Wheel Surface - (avoidance of undercut)

There are two main reasons why investigation of the singular points on the generated surface $\Sigma^{(2)}$ is needed:

- It helps in the avoidance of undercutting;

- It simplifies numerical solution of the worm gearing analysis equations, since it helps decide the range of curvilinear parameters (u, θ) that give feasible solutions. This is fully described in section 2.5.

Undercutting of worm wheel surfaces is associated with the appearance of singular points on them. When the singular points occur on the generated (wheel) surface $\Sigma^{(2)}$, the locus of these points may be the edge of a "cut away" part of the flank - that is, the line which connects two separate parts of the surface. The appearance of such a line results from the fact that part of generated wheel surface originally generated is later removed by the generating (worm) surface $\Sigma^{(1)}$. This can be demonstrated by the software, as illustrated in Figure 2.3. This shows that contact lines first move from the wheel tip towards the root, then turn backwards towards the tip once they have reached a certain limit. This limit is the line along which the singularity occurs.

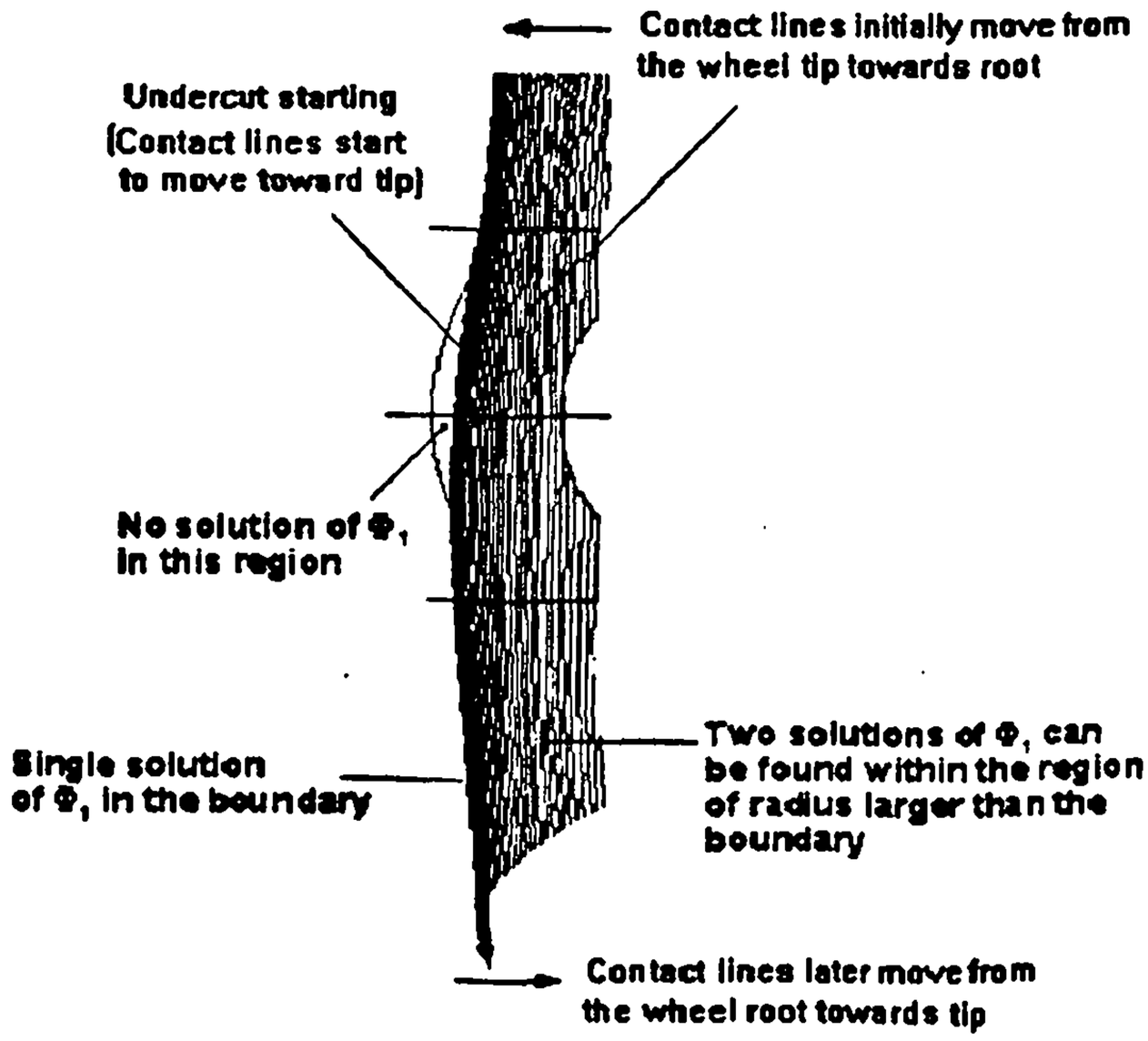


Figure 2.3 Undercut of Worm Wheel

The singularity occurs, according to differential geometry, when the generated wheel surface normal N_2=0. For practical purpose, it is better to control the range of curvilinear parameters (u,θ) (in other words, the size of the generating worm surfaces) to avoid undercut. Equation (2.16) can be used to determine the limit line on the generating worm surface $\Sigma^{(1)}$ as a locus of points which generate singular points on the generated surface $\Sigma^{(2)}$. This kinematic interpretation of equation (2.16) allows it to be used to avoid the problem of undercutting of the generated wheel.

Searching numerically for the values of the parameters (u,θ) of the generated wheel surfaces which gives N_2=0 requires very complicated iterative calculations [Litvin, 1992] to solve.

$$\begin{vmatrix} \frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial \theta} & Br_{x1} \\ \frac{\partial y_1}{\partial u} & \frac{\partial y_1}{\partial \theta} & Br_{x1} \\ \frac{\partial M}{\partial u} & \frac{\partial M}{\partial \theta} & \frac{\partial M}{\partial \Phi_1} \end{vmatrix} = 0 \qquad \ldots\ldots \quad (2.16)$$

where $\mathbf{B} \cdot \mathbf{r}_1^{(1)} = (Br_{x1}, Br_{y1}, Br_{z1}, 0)$.

Br_{x1}, Br_{y1} and the determinant can be obtained by library or user-defined functions.

2.3.9 Grid of Points on A Wheel Flank

In some applications it becomes necessary to determine the coordinates of a selected grid of points on the wheel surfaces. For an example, to measure wheel flanks, the coordinates of particular points on a wheel flank need to be given. In this case, the selected points on a wheel flank are the intersections of the specified measuring plane/cylinder with the generated wheel surface.

The wheel surface is defined by

$$\begin{cases} M=N_1^{(1)}\cdot B\cdot r_1^{(1)}=0 \\ r_2^{(2)}=T_{21}\cdot r_1^{(1)} \\ N_2^{(2)}=W_{21}\cdot N_1^{(1)} \end{cases} \quad \ldots\ldots \quad (2.12)$$

Geometric conditions for obtaining a uniform grid of points may be expressed by:

$$\begin{cases} z_2^{(2)} = H \\ (x_2^{(2)})^2+(y_2^{(2)})^2 = R_2 \end{cases} \quad \ldots\ldots \quad (2.17)$$

The satisfaction of equations (2.12) and (2.17) determines a point on the worm wheel surface at radius R_2 in the section distant H from the wheel mid plane. If R_2 is a constant and H varies, the solution of equations (2.12) and (2.17) generates the wheel lead at radius R_2. If H is a constant and R_2 varies, the solution of equations (2.12) and (2.17) generate the wheel profile in the section at H. If H and R_2 are both given a set of uniformly - spaced values, a regular 'rectangular' grid of points on the flank is generated.

2.3.10 Principal Relative Normal Curvatures - Contact Geometry

The relative normal curvature of two mating surfaces at the contact point is the difference between their normal curvatures taken in a common section normal to both. If the normal curvatures of the worm and wheel are $k_n^{(1)}$ and $k_n^{(2)}$ respectively, their relative normal curvature k_R is defined as [Litvin, 1992]

$$k_R = k_n^{(2)} - k_n^{(1)}$$

A relative normal curvature of two mating surfaces in a specific normal section can be regarded as a measure of conformity between the two surfaces in that direction. This is illustrated in Figure 2.4. There is an infinite number of relative normal curvatures at a point because of the infinity of possible normal sections. The principal relative curvatures are the two extreme (maximum and minimum) values of the relative normal curvatures.

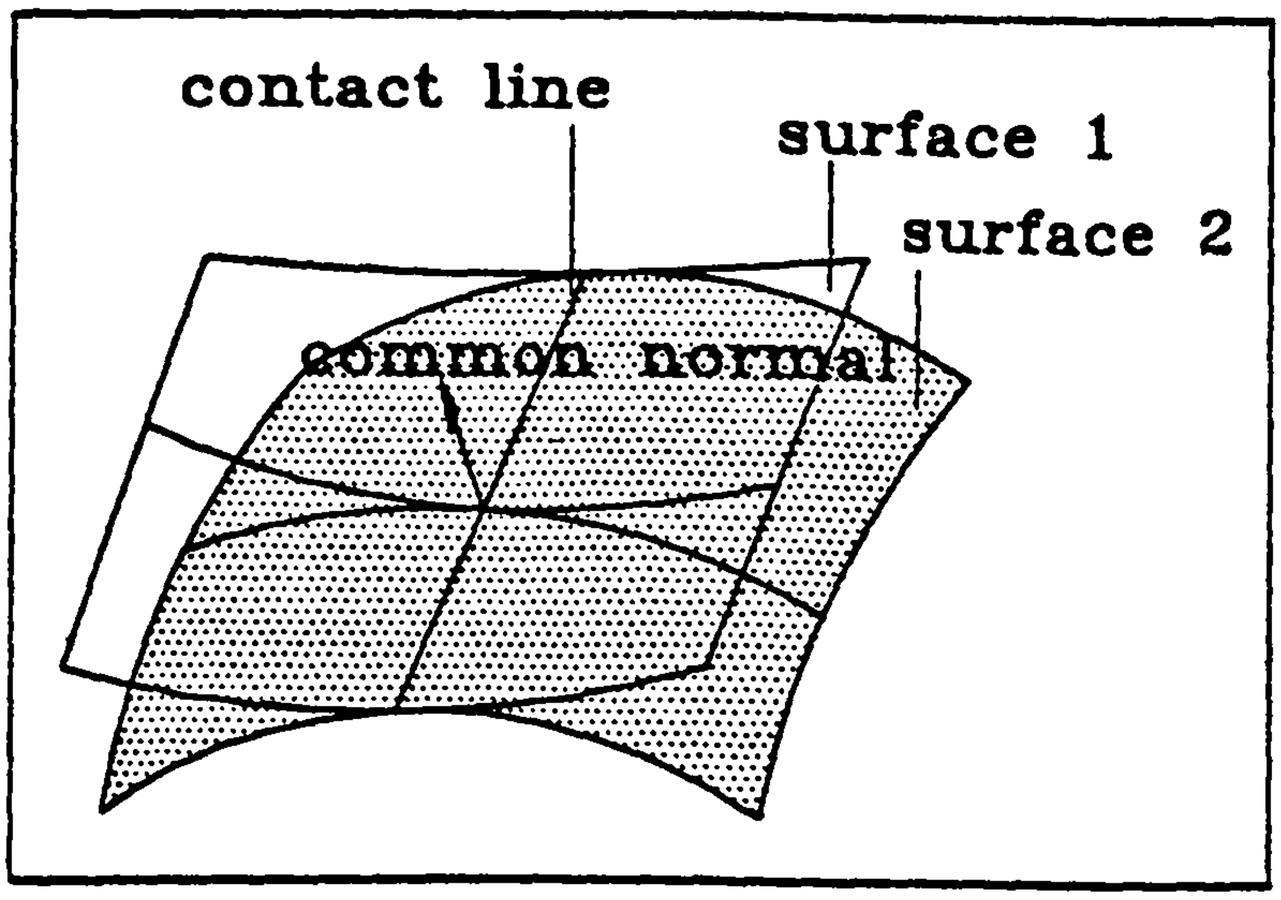


Figure 2.4 Relative Normal Curvature

The calculation of principal relative normal curvatures is recognized as one of the most difficult problems in the theory of worm gearing. However, these normal curvatures are critical for the calculation of both the oil film thickness [Wu,1986] [Dowson, 1968] [Errichello, 1991] and the Hertzian contact stresses [Pennell, 1990] [Evans and Snidle, 1992]. It is well known that, in general, the smaller the relative normal curvature is, the better the contact and lubrication conditions are. Therefore, making an effort to find relative normal curvatures is well worthwhile.

An approximation can be used to find the maximum relative normal curvatures in the plane perpendicular to the contact line, which can be obtained from the non conjugate gap contours around each contact point. However, many kinematic analyses do not go as far as to produce the gap contour. In this case an exact analytical model may be the easier means to determine the relative normal curvatures. In fact, with the power of modern personal computers, it is no longer a problem to determine relative normal curvatures accurately.

The principal relative curvatures are in two common normal planes which are perpendicular to each other [Litvin,1989]. We limit our discussion to the line contacts

between the conjugate worm and wheel surfaces. The minimum principal normal curvature in this case is obviously zero, and occurs in the plane specified by the common normal and the tangent to the instantaneous contact line. Therefore the maximum relative normal curvature must be in the plane specified by the common normal and the normal of the instantaneous contact line. We are interested in this curvature.

Principal relative curvatures are determined by the geometry of the generating (worm) surface and the relative motion. At each point of the worm surface, there are two principal curvatures, and two corresponding principal directions. Let us denote the worm surface principal directions by unit vectors $g_{minf}^{(1)}$ and $g_{maxf}^{(1)}$, and the two principal curvatures by k_{min} and k_{max}. Unit vectors $g_{minf}^{(1)}$ and $g_{maxf}^{(1)}$ are represented in terms of components of coordinate system S_f (O_f, x_f, y_f, z_f). The maximum relative normal curvature between the worm and wheel surfaces at the contact point is then given by [Ye, 1990]:

$$k_n^{(1,2)} = \frac{(m_1^2 + m_2^2)}{\psi} \quad \ldots\ldots \quad (2.18)$$

where

$$m_1 = k_{1min} \cdot (g_{minf}^{(1)} \cdot V_{sf}^{(12)}) + g_{maxf}^{(1)} \cdot \omega_f^{(12)}$$

$$m_2 = k_{max} \cdot (g_{maxf}^{(1)} \cdot V_{sf}^{(12)}) - g_{minf}^{(1)} \cdot \omega_f^{(12)}$$

$$\Psi = m_1 \cdot (g_{minf}^{(1)} \cdot V_{sf}^{(12)}) + m_2 \cdot (g_{maxf}^{(1)} \cdot V_{sf}^{(12)}) + M_{\Phi_1}$$

and

$$V_{sf}^{(12)} = T_{f1} \cdot B(l) \cdot r_1^{(1)} \cdot \omega_1$$

$$M_{\Phi_1} = N_f^{(1)} \cdot B_{\Phi_1} \cdot r_f^{(1)} = N_1^{(1)} \cdot B_{\Phi_1} \cdot r_1^{(1)}$$

$$\omega_f^{(12)} = \omega_f^{(1)} - \omega_f^{(2)}$$

It must be noticed that all the vectors in equations (2.18) are represented in terms of components of coordinate system S_f (O_f, x_f, y_f, z_f). However, scalar values, like $M_{\Phi 1}$, may clearly be calculated in any convenient coordinate system as shown.

2.3.11 The Orientation of the Contact Line

The orientation of the contact line is another parameter required for the lubrication analyses. It is well accepted that a larger angle between the contact line and the sliding velocity normally results in better lubrication conditions. The values of m_1 and m_2 obtained in equations (2.18) are all we need to decide the orientation of the contact line, since the angle between the tangent of the contact line and the principal direction $g_{minf}^{(1)}$ of the worm surface is given by [Ye et al, 1990]:

$$\Omega = \tan^{-1}(-m_1/m_2) \quad \ldots\ldots \quad (2.19)$$

It follows that the angle between the normal to the contact line and $g_{minf}^{(1)}$ is given by:

$$\Theta = \tan^{-1}(m_2/m_1) \quad \ldots\ldots \quad (2.20)$$

Because $g_{minf}^{(1)}$ is known for a given worm, the orientation of the contact line can thus be determined.

2.4 Derivations of the B-Matrix Equations

In the previous section, the B-matrix equations used in the analysis of worm gearing were simply quoted without proof. In this section, these equations are derived from the basic meshing theory, which is taken as the starting point.

2.4.1 Definition of the B-matrix

The B-matrix is the basic matrix for the conjugate analyses and used throughout the calculations of B-matrix method, its definition is given by form (2.8):

$$B(\phi_1, i) = W_{12} \cdot \frac{\partial T_{21}}{\partial \phi_1} \quad \ldots\ldots \quad (2.8)$$

Where T_{21} is the coordinate transformation matrix from coordinate system S_1 To S_2. W_{12} is the coordinate rotation matrix from S_2 to S_1. They are given by equations (2.6) and (2.7).

It should be noticed that the B-matrix depends only on motional parameters, as discussed before in section 2.1.

2.4.2 Derivation of the Relative (Sliding) Velocity and the Equation of Meshing

The relative velocity $V_s^{(12)}$ can be regarded as the velocity of a contact point on surface $\sum^{(1)}$ observed by an observer standing on surface $\sum^{(2)}$. The position vector of the contact point on $\sum^{(1)}$ is represented by coordinates of $S_2(O_2, x_2, y_2, z_2)$ as $r_2^{(1)}$. Then the relative velocity represented by coordinates of $S_2(O_2, x_2, y_2, z_2)$ can be determined by formula (2.21).

$$V_{s2}^{(12)} = \frac{dr_2^{(1)}}{dt} = \frac{d(T_{21} \cdot r_1^{(1)})}{dt}$$

$$= \frac{dT_{21}}{dt} \cdot r_1^{(1)} + T_{21} \cdot \frac{dr_1^{(1)}}{dt} \qquad (\frac{dr_1^{(1)}}{dt} = 0)$$

$$= (\frac{\partial T_{21}}{\partial \phi_1} \cdot \frac{d\phi_1}{dt}) \cdot r_1^{(1)} = \frac{\partial T_{21}}{\partial \phi_1} \cdot r_1^{(1)} \cdot \omega_1 \quad \ldots \quad (2.21)$$

Let us represent the relative velocity $V_2^{(12)}$ by coordinates of $S_1(O_1, x_1, y_1, z_1)$ as formula (2.13)

$$V_{s1}^{(12)} = W_{12} \cdot V_{s2}^{(12)} = W_{12} \cdot \frac{\partial T_{21}}{\partial \phi_1} \cdot r_1^{(1)} \cdot \omega_1$$

$$= B \cdot r_1^{(1)} \cdot \omega_1 \quad \ldots \quad (2.13)$$

where ω_1 is the speed of rotation of surface $\sum^{(1)}$ about its fixed axis.

$V_{s1}^{(12)}$ is used in this thesis to represent relative (sliding) velocity in order to distinguish it from rolling velocity $V_{r1}^{(12)}$.

Since no separation or penetration of the two mating surfaces can occur, the relative velocity must be in the common tangent plane. This is mathematically expressed in form (2.22).

$$f = N_1^{(1)} \cdot V_{s1}^{(12)} = 0 \quad \ldots \quad (2.22)$$

$$\because \quad V_{s1}^{(12)} = B \cdot r_1^{(1)} \cdot \omega_1$$

$$\therefore \quad f = N_1^{(1)} \cdot B \cdot r_1^{(1)} \cdot \omega_1 = M \cdot \omega_1 = 0 \quad (where\ M = N_1^{(1)} \cdot B \cdot r_1^{(1)}\ and\ \omega_1 \neq 0)$$

Therefore, the equation of meshing can be expressed as:

$$M = N_1^{(1)} \cdot B \cdot r_1^{(1)} = 0 \quad \cdots \quad (2.11)$$

2.4.3 Derivation of the Rolling Velocity

The rolling velocity is defined in the thesis as $V_r^{(12)} = V^{(1)} + V^{(2)}$.

$$V_r^{(12)} = V^{(1)} + V^{(2)} = V^{(1)} - (-V^{(2)})$$

$$\because \quad -V^{(2)}(\omega_2) = V^{(2)}(-\omega_2) \quad \text{(A worm wheel rotates about its fixed axis)}$$

$$\therefore \quad V_r^{(12)} = V^{(1)} - V^{(2)}(-\omega_2)$$

$$\because \quad -\omega_2 = -i \cdot \omega_1 \ (-\omega_2 \text{ can be interpreted as the direction of wheel rotation is reversed})$$

Substitute i with -i in equation (2.13), we have the rolling velocity represented by coordinates of S_1 as formula (2.14)

$$\therefore \quad V_{r1}^{(12)} = B(-i) \cdot r_1^{(1)} \cdot \omega_1 \qquad \cdots (2.14)$$

2.4.4 Derivation of The Envelope of Contact Lines on A Worm Surface

The generated surface is the envelope of a family of surfaces - the moving generating surface, in other words the generated surface is formed by contact lines. The contact lines are a family of curves on the worm/hob surface. The family is specified by the equation of meshing (2.11) and Φ_1 is the parameter of which any permitted value yields a member-curve of the family. If a curve exists that touches every member of the family, then this curve is known as the envelope of contact lines. According to the envelope theory described by Heading(1979), the envelope of contact lines on the worm flank which satisfy equation (2.11), can be obtained by following procedure.

$$\frac{\partial M}{\partial \phi_1} = \frac{\partial (N_1^{(1)} \cdot B \cdot r_1^{(1)})}{\partial \phi_1} = 0$$

Since $N_1^{(1)}$ and $r_1^{(1)}$ are independent of ϕ_1, this yields

$$\frac{\partial M}{\partial \phi_1} = N_1^{(1)} \cdot \frac{\partial B}{\partial \phi_1} \cdot r_1^{(1)} = N_1^{(1)} \cdot B_{\phi_1} \cdot r_1^{(1)} = 0 \quad \ldots \quad (2.15)$$

2.4.5 Derivation of The Singularities of the Worm Wheel Surface

According to [Litvin, 1989], a singularity will occur on the generating surface wherever

$$\begin{vmatrix} \frac{\partial x_1^{(1)}}{\partial u} & \frac{\partial x_1^{(1)}}{\partial \theta} & V_{x1}^{(12)} \\ \frac{\partial y_1^{(1)}}{\partial u} & \frac{\partial y_1^{(1)}}{\partial \theta} & V_{y1}^{(12)} \\ \frac{\partial f}{\partial u} & \frac{\partial f}{\partial \theta} & \frac{\partial f}{\partial \Phi_1} \cdot \frac{d\phi_1}{dt} \end{vmatrix} = 0$$

where $f = N_1^{(1)} \cdot V_{s1}^{(12)}$

and

$$\because \; V_{s1}^{(12)} = B \cdot r_1^{(1)} \cdot \omega_1$$

$$= (Br_{x1} \cdot \omega_1, Br_{y1} \cdot \omega_1, Br_{z1}, \cdot \omega_1, 0)$$

$$= (V_{x1}^{(12)}, V_{y1}^{(12)}, V_{z1}^{(12)}, 0) \qquad \cdots\cdots \qquad (2.23)$$

$f = M \cdot \omega_1$ (refer to section 2.4.2)

Thus

$$\begin{vmatrix} \frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial \theta} & Br_{x1} \cdot \omega_1 \\ \frac{\partial y_1}{\partial u} & \frac{\partial y_1}{\partial \theta} & Br_{y1} \cdot \omega_1 \\ \frac{\partial M}{\partial u} \cdot \omega_1 & \frac{\partial M}{\partial \theta} \cdot \omega_1 & \frac{\partial M}{\partial \Phi_1} \cdot \omega_1 \cdot \omega_1 \end{vmatrix} = \begin{vmatrix} \frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial \theta} & Br_{x1} \\ \frac{\partial y_1}{\partial u} & \frac{\partial y_1}{\partial \theta} & Br_{y1} \\ \frac{\partial M}{\partial u} & \frac{\partial M}{\partial \theta} & \frac{\partial M}{\partial \Phi_1} \end{vmatrix} \cdot \omega_1 \cdot \omega_1 = 0$$

because $\omega_1 \neq 0$, *this gives*

$$\begin{vmatrix} \frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial \theta} & Br_{x1} \\ \frac{\partial y_1}{\partial u} & \frac{\partial y_1}{\partial \theta} & Br_{y1} \\ \frac{\partial M}{\partial u} & \frac{\partial M}{\partial \theta} & \frac{\partial M}{\partial \Phi_1} \end{vmatrix} = 0 \qquad \cdots\cdots \qquad (2.16)$$

2.4.6 Derivation of Principal Relative Normal Curvatures and the Orientation of the Contact Line

The equations used in this part are developed from equations given by [Ye et al, 1990]. The development involves simply introducing the B-matrix method to determine the relative velocity and the envelope of contact lines on the worm flank into the equations.

2.5 Solution technique and considerations

Most of the B-matrix kinematic equations, as with other 3D kinematic models of worm gearing, are nonlinear. Except in special cases, nonlinear equations cannot be solved analytically, so that numerical solutions are required.

The numerical methods employed in this thesis are the Golden Section (0.618 interval) and the Newton-Raphson methods. Details of these methods are not discussed here since they can be found in many text books. The discussion given by Heading (1979) on stationary values of functions was also found to be helpful in determining the boundaries of a root. The major problems involved in using these methods are that there is no guarantee of convergence to the correct solution for "practical" worm gearing, although, in most cases, these problems can be overcome by combining practical knowledge of worm gears with the numerical solution techniques. Therefore, attention to the physical significance of the parameters is crucial to success in finding the correct roots of the mathematical equations of worm gearing.

The following discussion is not confined to a particular type of worm gearing, although it arose from the study of convolute (straight line generator) helicoid worm gears. Some points may be obvious but can be easily forgotten when attention is wholly focussed on the complicated computation procedure. Important points are:

1) Break the 3D models into 2D projections for finding solutions when necessary. This leads to better convergence and to an easier way of verifying the solution, for example, by determining the singular point in a particular section. If the golden section

method is used, first determine the boundary, then "capture" the root confidently. If the Newton-Raphson method is used, it is essential to check the solutions against physical sense, particularly since circular functions of Φ_1 and θ are used, so that "solutions" in the wrong range can easily be found that are beyond the limits of "real" gears.

2) The mapping (transformation) between position vectors of the worm flank and curvilinear parameters u and θ is a single value mapping. However, since the radial position of a point on a worm/hob surface depends only on the parameter u and the axial position dominantly on the parameter θ, the boundaries of the roots for parameters u and θ are easily found.

3) Viewed from the end of the worm/hob shaft, the angle $(\Phi_1 + \theta)$ determines the angle of generator rotation. The generator must, in this view, lie within the wheel facewidth. This determines the boundaries of the roots of parameter θ when Φ_1 is known.

4) Each individual value of Φ_1 determines a unique contact line, although the contact line may sometimes split into two segments (for ZC and globoidal worm gears, for example). The position of the contact lines on the wheel tooth depends on the value of Φ_1 (they move from wheel tip to the wheel root when the worm drives the wheel).

5) If the equations are being solved for Φ_1, three possible cases need to be considered in terms of physical sense: no solution, single solution or two solutions. This is illustrated in Figures 2.2 and 2.3 (pp. 23, 24). In addition, it is important to remember that circular functions of Φ_1 are again involved. Therefore, to input Φ_1 and solve for other parameters is usually an easier way. If solving for Φ_1 is unavoidable, determining the singular points on the wheel surface and the envelope of contact lines on the worm/hob surface in advance is a good practice from the point of view of computation, since this effectively eliminates the chance of multiple solutions. However, subsequently verifying the solution against physical sense is also necessary.

6) Summarising the above discussion, finding the surface of action is a good way to start the computation. This helps in defining the boundaries of the variables. The surface of action is the locus of contact lines and normally it has a four-sided boundary.

Three of the four sides are determined by the wheel tip and the wheel facewidth ends. Another side is determined by the worm/hob tip or the maximum worm/hob radius without undercutting, whichever is least. The accurate method of determining the surface of action is to employ equations (2.16) and (2.17). However, an easier alternative method may also be used. This uses only equation (2.12). If the latter is chosen, the following procedure can be applied:

step 1- chose an adequate range of angle of worm rotation Φ_1;

step 2- fix rotation angle Φ_1;

step 3- choose a value of curvilinear parameter θ;

step 4- determine the curvilinear parameter u from the equation of meshing;

step 5- record the position vector and the parameters (u,θ) if the contact point is on/near the boundary of the surface of action. The singular points on the wheel conjugate surface are the points in each transverse plane where the minimum radius occurs.

step 6- increase Φ_1 in small increments until all four sides of the surface of action have been determined.

step 7- take the corresponding values of the parameters (u,θ) on the boundary of the surface of action at each value of Φ_1 as the boundaries of the parameters.

2.6 Validation of the B-matrix models and calculation procedures

In addition to validating each individual equation against Litvin's method, validation against the involute section calculation has also been carried out with good agreement. A ZI worm/hob generates a "thin spur gear" in a special section of the worm wheel tooth when the two shafts are at 90^0[Hu and Pennell, 1994]. This provides a method to check the 3D kinematic models for one special condition. We can, in this way, check the maximum radius of a worm/hob without undercutting, the relative curvature and orientation of the contact line, because the results obtained from the B-Matrix models include all this information. The results in the wheel involute section calculated by the B-matrix models agree very well with those given by Litvin's equations and the

involute/basic rack theory, as shown in Table 2.1.

	B-matrix	Litvin method	Spur gear
Undercut radius	124.325 mm		124.325 mm
Relative curvature	-0.017691	-0.017690	-0.0176904
Orientation of contact line (to transverse plane)	89.99999^0		90^0

Table 2.1 Validation of Conjugate Calculation

The worm gear data used for above calculation is given in Appendix 2. The method for determining data of the equivalent spur gear from the worm gear data is given in Appendix 4.

It can be concluded that the B matrix method has reached the stage of solving all the major problems presently involved in conjugate analysis of worm gearing. The main advantages of B-Matrix method over other methods is that it fully uses the power of modern computers and simplifies the models for general 3D gearing problems, giving less chance of error.

The equation of meshing using B-matrix method was initially developed by the author's colleagues at the Huazhong University of Science and Technology, China. The further development of the method described above was carried out by the author at the University of Newcastle upon Tyne, UK.

CHAPTER 3

THE ANALYTICAL MEASUREMENT OF WORM WHEELS OF CYLINDRICAL WORMGEARS

Measurement of cylindrical worms presents few problems, since such a worm is essentially a thread of a particular thread form. Most modern CNC gear measuring and industrial CMMs can measure ZI and/or ZA worms without difficulty. The development of a technique for measurement of globoidal worms was achieved by the author in 1988 [Hu, 1989]. However, because of their much more complex geometry, successful absolute measurement of cylindrical worm wheels had not been reported when the author started his research.

This chapter describes a method of measurement of the worm wheels of cylindrical worm drives. The method presented here is an absolute measurement in which the profile and lead errors are measured relative to the ideal (theoretical) tooth form. The calculated tooth geometry is used to control a coordinate measuring machine so that the difference between the theoretical and actual tooth form can be obtained directly.

The measurement principle can be applied to any cylindrical worm gear. The mathematical model given here can be applied to worm gears of any convolute helicoid form, i.e. involute helicoid (ZI), straight sided axial profile (ZA) and straight sided normal profile (ZN).

3.1 Worm wheel testing in industry and the need for analytical measurement

Even on cylindrical worm gears, the geometry of the worm wheel is complicated. This has led to the widespread practice of only inspecting the worm wheel indirectly by

means of a composite meshing test with a master worm - the blue marking test. The normal procedure for carrying out a blue marking test is:

(1) Set the hobbing machine with calculated data, and hob a wheel preliminarily;

(2) Set the wheel in mesh with a standard worm in a meshing frame, with the worm coated with a thin layer of soft blue dye, to get a contact pattern on the wheel tooth flanks. The soft blue dye is commonly painted onto the worm flank with a brush;

(3) Check the contact pattern and backlash. If the test results are not satisfactory then adjust the machine settings and hob the wheel again. The direction and amount of the adjustment are frequently estimated by the hobbing machine operator, based on his experience.

(4) Repeat above procedures (2) and (3) until the required contact pattern and backlash are achieved.

To optimize the worm wheel geometry effectively, it is important that adjustments to the hobbing settings are relative to the initial settings, since both these and the hob accuracy may remain uncertain.

Typical contact patterns are given in Figure 3.1 [ANSI/AGMA 6022-C93, 1993]. Usually the marking is just checked visually, but a permanent record can be taken by transferring the marking onto a strip of 'Sellotape' to which the blue dye adheres.

The advantages of the blue marking test are that it is easy, rapid and requires no precise data for hobbing settings to start with. However, the contact pattern is a functional inspection result, which is not good for identifying individual errors, and subsequently, distinguishing individual error sources. It is thus quite difficult to interpret the results quantitatively in terms of changes required to tooling, gear mounting or machine settings.

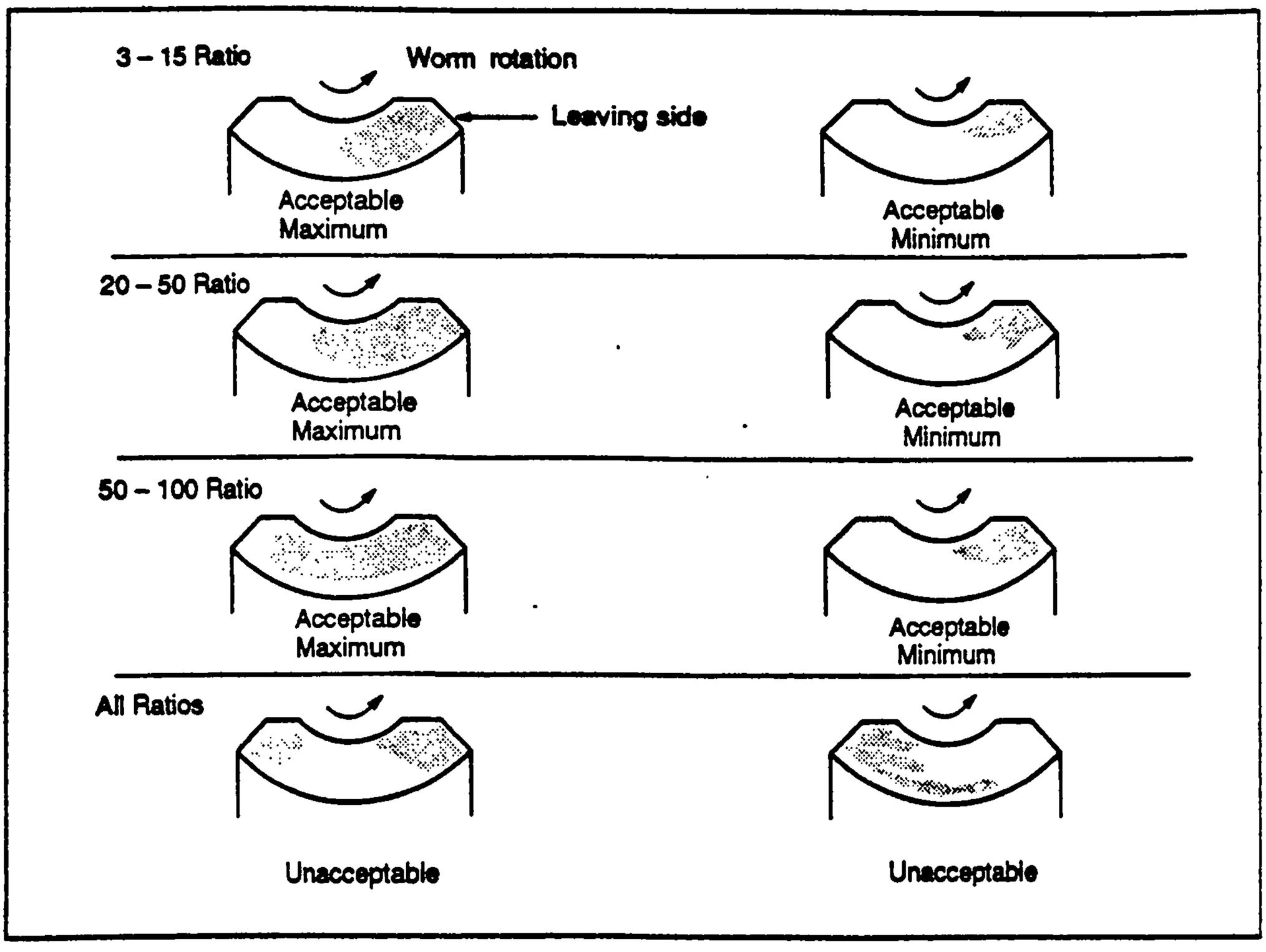


Figure 3.1 Typical Contact Patterns

Some modern CNC gear measuring machines can measure worm wheel profile and lead errors individually with respect to the tooth form of a master worm wheel, but these are only relative measurements, in which the relation of the actual tooth geometry to that ideally intended is never known. The results from these comparative measurements are, thus, of limited value in practice, since effective quality control calls for an inspection procedure of verifiable accuracy. They merely serve to demonstrate the repeatability of the production process.

Analytically measuring the individual errors becomes essential when high precision worm wheels must be made or 'real' worm gear performance predicted. Analytical measurements give the following advantages:

- Accuracy is verifiable, since traceable calibration is possible,

- Errors due to the master worm wheel are excluded,

- It is easier to determine the source of individual errors,

- Entry/exit gaps, profile barrelling, etc., can all be quantified straight away, and much more accurately than with the blue marking test.

- The results obtained are of direct significance for worm gear performance,

- Measured data for 'real' wheel flank forms, before and after running, can be directly transferred for contact stress, tooth deflection, lubrication or transmission error analysis,

- Optimum geometry can be more easily achieved since the confusion caused by unknown manufacturing errors has vanished,

- Investigation of worm gear drive problems is greatly facilitated.

3.2 Measurement principle

Difficulties involved in analytically measuring individual errors are mainly caused by the complexity of worm wheel flank geometry. It was impossible to carry out a fully analytical measurement for a worm wheel tooth flank until the advent of CNC measuring machines. The measuring probe of a coordinate measuring machine can be commanded to trace any prescribed spatial curve or scan any prescribed spatial surface, and it was this capability which opened up the possibility of analytically measuring worm wheel flanks.

Once the theoretical geometry of a worm wheel tooth flank has been determined, it is possible to measure its profile errors in any transverse plane and its lead errors on any cylinder with a suitably-programmed coordinate measuring machine. However, to simplify the measurement for practical use as a routine quality control technique, lead errors are measured on the reference cylinder and profile errors at any selected transverse section

as in Figures 3.3 and 3.4. This inspection procedure is very similar to that normally used for involute spur and helical gears, in which profile and lead errors are measured automatically on both flanks of three or four teeth equally-spaced round the gear. The associated measurement of pitch errors and tooth thickness presents no new problems on a worm wheel. In addition, the combination of lead errors on the reference cylinder and profile errors on several transverse sections can be used to give topological measurement results for the whole worm wheel tooth flank.

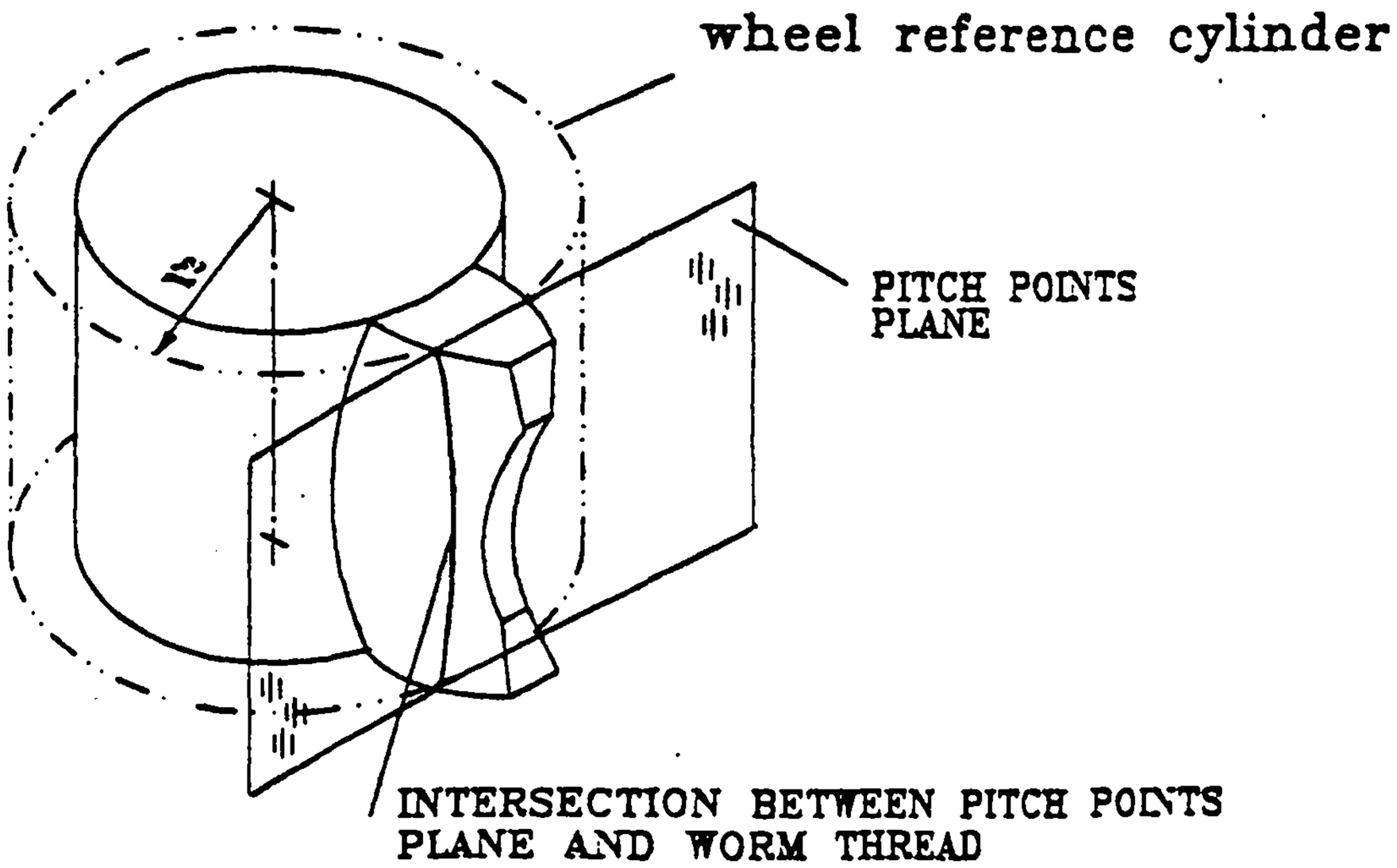


Figure 3.2 Pitch Points Plane

The wheel measurement procedure presented here defines the errors relative to the mating worm. In other words, the deviations to be measured are from a wheel tooth flank which has perfect conjugate action with the worm. To increase the overall measurement speed, the proposed method of measurement is based on a kinematic model of a conjugate worm and wheel with axes at 90°, since this is the usual meshing arrangement [Hu and Pennell, 1994].

The theoretical worm wheel is thus here generated by the mating worm. The theoretical worm wheel tooth profile in each transverse section is generated by the corresponding offset worm axial section just as a gear tooth profile is generated by a rack [Buckingham, 1949]. To explain the lead measurement principle, the concept of a "pitch points plane" is introduced.

The pitch points plane is the plane tangent to the worm wheel reference cylinder and parallel to the worm axis (Figure 3.2). The intersection between the worm helicoid and the pitch points plane is a 2D curve. Rolling the 2D curve around the worm wheel reference cylinder will give the lead trace on the theoretical worm wheel reference cylinder. This is because the rotary motion of the theoretical worm wheel can be nominally regarded as a pure roll of the reference cylinder on the pitch points plane.

The effect of the finite radius of a spherical probe must be also considered. The `theoretical' normal to the wheel tooth surface is used to position the probe centre so that a probe would contact a perfect wheel tooth surface at the 'correct' point. On tooth flanks with deviations from the theoretical form, this choice of contact point results in slightly different deviations being measured with probes of differing size, which contact the flank at slightly different positions. In practice, however, these differences are very small and, just as in the measurement of involute spur and helical gears, they can be ignored, since the errors introduced are negligible compared with the measured deviations from the theoretical form which cause them.

3.3 Kinematic equations

Kinematic equations for the measurement have been derived which allow analytical solutions. This facilitates practical implementation of the technique (see section 3.5). For ease of representation, the coordinate system S_2 (O_2, x_2, y_2, z_2) has been redrawn on a coordinate gear measuring machine as in Figure 3.6 (b) to which the following equations (3.1) to (3.6) refer:

3.3.1 Worm wheel profiles in transverse sections

The profile to be measured is in a user selected section at a distance D from the worm wheel mid plane shown in Figure 3.3. The profile and corresponding normal to the wheel tooth flank can be determined from equations (3.2) and (3.3). See [Hu, 1992] for details:

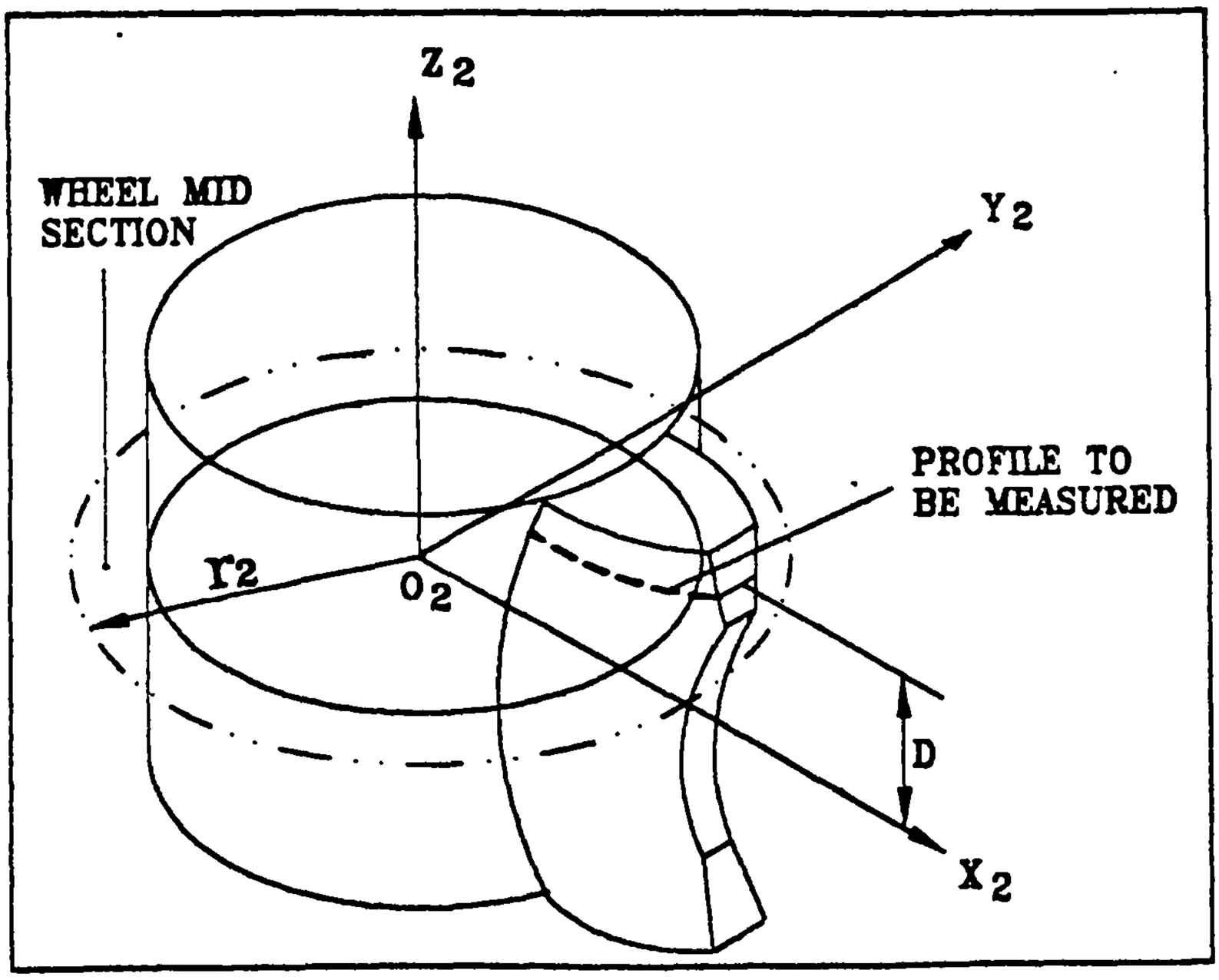


Figure 3.3 Profile to be measured

$$r_2=\frac{m_a \cdot z_2}{2} \quad r_1=C-r_2$$

$$x_1^{(1)}=\sqrt{r^2-D^2} \quad r \in (r_{f1},\ r_{a1})$$

$$y_1^{(1)}=-D$$

$$\theta=cos^{-1}\left(\frac{y_1^{(1)}}{r}\right)-sin^{-1}\left(\frac{r_b}{r}\right) \quad u=\frac{\sqrt{r^2-R_b^2}}{cos\lambda}$$

$$K=\frac{p_z}{2\cdot\pi} \quad z_1^{(1)}=u\cdot sin\lambda-K\cdot\theta$$

$$N_1^{(1)}=\begin{pmatrix} N_{1x}^{(1)} \\ N_{1y}^{(1)} \\ N_{1z}^{(1)} \\ 0 \end{pmatrix}=\begin{pmatrix} -K\cdot cos\lambda\cdot cos\theta+r_b\cdot sin\lambda\cdot cos\theta+u\cdot cos\lambda\cdot sin\lambda\cdot sin\theta \\ r_b\cdot sin\lambda\cdot sin\theta-u\cdot cos\lambda\cdot sin\lambda\cdot cos\theta-K\cdot cos\lambda\cdot sin\theta \\ u\cdot cos^2\lambda \\ 0 \end{pmatrix} \quad \ldots (3.1)$$

$$L_1=z_1^{(1)}-\frac{N_{1z}^{(1)}}{N_{1x}^{(1)}}\cdot(x_1^{(1)}-r_1) \quad \Phi_2=\frac{L_1}{r_2}$$

$$r_2^{(2)}=\begin{pmatrix} x_2^{(2)} \\ y_2^{(2)} \\ z_2^{(2)} \\ 1 \end{pmatrix}=\begin{pmatrix} -x_1^{(1)}\cdot cos\Phi_2-z_1^{(1)}\cdot sin\Phi_2+C\cdot cos\Phi_2+L_1\cdot cos\Phi_2 \\ x_1^{(1)}\cdot sin\Phi_2-z_1^{(1)}\cdot cos\Phi_2-C\cdot sin\Phi_2+L_1\cdot cos\Phi_2 \\ -y_1^{(1)} \\ 1 \end{pmatrix} \quad \ldots (3.2)$$

$$N_2^{(2)}=\begin{pmatrix} N_{2x}^{(2)} \\ N_{2y}^{(2)} \\ N_{2z}^{(2)} \\ 0 \end{pmatrix}=\begin{pmatrix} -N_{1x}^{(1)}\cdot cos\Phi_2-N_{1z}^{(1)}\cdot sin\Phi_2 \\ N_{1x}^{(1)}\cdot sin\Phi_2-N_{1z}^{(1)}\cdot cos\Phi_2 \\ -N_{1y}^{(1)} \\ 0 \end{pmatrix} \quad \ldots (3.3)$$

3.3.2 The worm wheel lead trace on the reference cylinder

The worm wheel lead trace to be measured is on its reference cylinder. The measuring range is from +V to -V, as shown in Figure 3.4. V is chosen as a half of the effective face width defined in BS 721(or the actual half face width of the worm wheel, whichever is least). The lead on the reference cylinder and corresponding normal to the wheel flank can be calculated from equations (3.5) and (3.6):

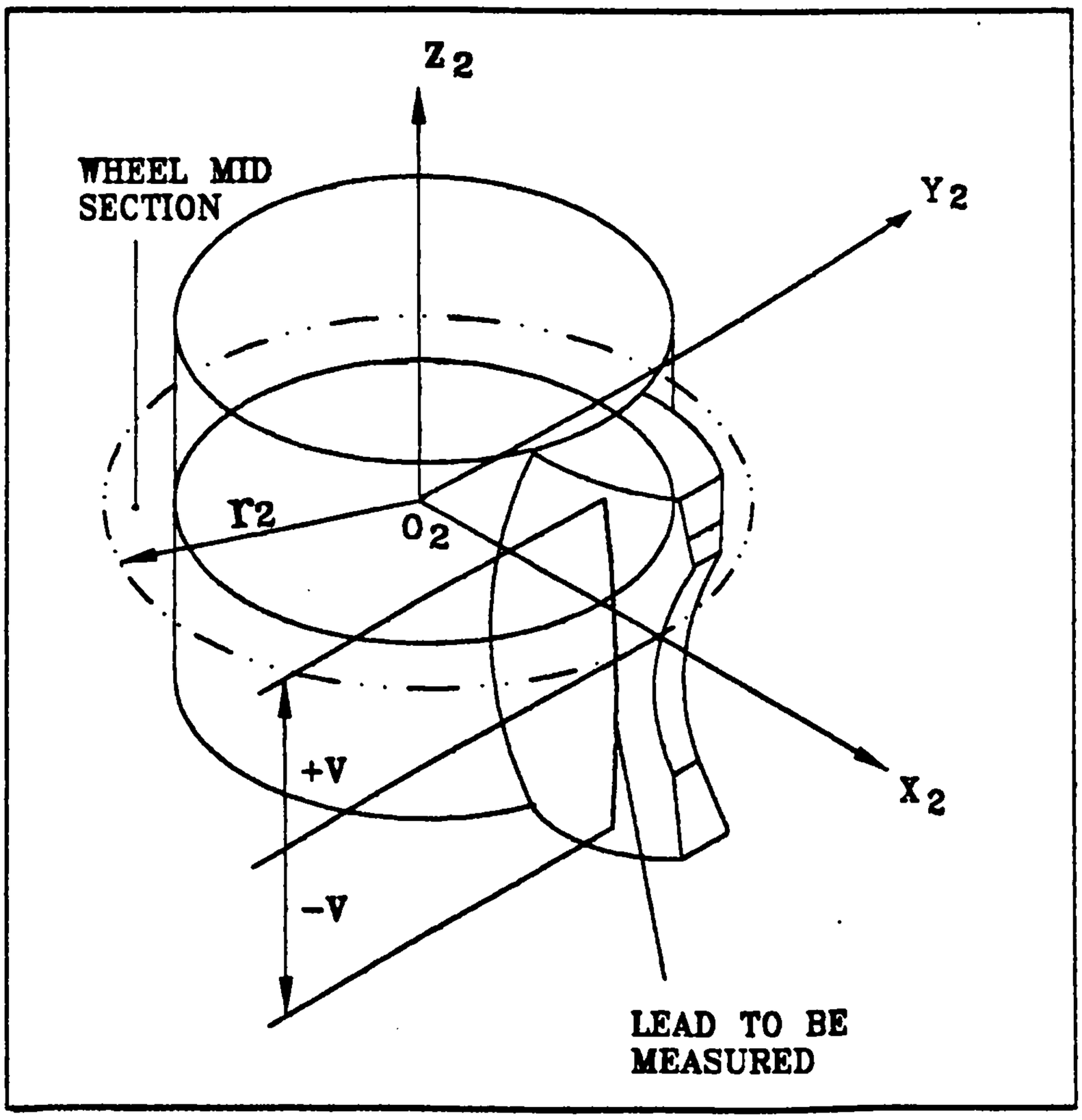


Figure 3.4 Lead to be measured

$$r_2=\frac{m_s \cdot z_2}{2} \qquad r_1=C-r_2$$

$$x_1^{(1)}=C-r_2 \qquad y_1^{(1)}=-H \qquad H \in [V,-V]$$

$$r=\sqrt{(x_1^{(1)})^2+(y_1^{(1)})^2}$$

$$\theta=sin^{-1}(\frac{H}{r})-sin^{-1}(\frac{r_b}{r}) \qquad u=\frac{\sqrt{r^2-R_b^2}}{cos\lambda}$$

$$K=\frac{p_z}{2\cdot\pi} \qquad z_1^{(1)}=u\cdot sin\lambda-K\cdot\theta$$

$$N_1^{(1)}=\begin{pmatrix} N_{1x}^{(1)} \\ N_{1y}^{(1)} \\ N_{1z}^{(1)} \\ 0 \end{pmatrix}=\begin{pmatrix} -K\cdot cos\lambda\cdot cos\theta+r_b\cdot sin\lambda\cdot cos\theta+u\cdot cos\lambda\cdot sin\lambda\cdot sin\theta \\ r_b\cdot sin\lambda\cdot sin\theta-u\cdot cos\lambda\cdot sin\lambda\cdot cos\theta-K\cdot cos\lambda\cdot sin\theta \\ u\cdot cos^2\lambda \\ 0 \end{pmatrix} \quad \ldots (3.4)$$

$$L_1=z_1^{(1)}-\frac{N_{1z}^{(1)}}{N_{1x}^{(1)}}\cdot(x_1^{(1)}-r_1) \qquad \phi_2=\frac{L_1}{r_2}$$

$$r_2^{(2)}=\begin{pmatrix} x_2^{(2)} \\ y_2^{(2)} \\ z_2^{(2)} \\ 1 \end{pmatrix}=\begin{pmatrix} r_2\cdot cos\phi_2 \\ -r_2\cdot sin\phi_2 \\ -y_1^{(1)} \\ 1 \end{pmatrix} \quad \ldots (3.5)$$

$$N_2^{(2)}=\begin{pmatrix} N_{2x}^{(2)} \\ N_{2y}^{(2)} \\ N_{2z}^{(2)} \\ 0 \end{pmatrix}=\begin{pmatrix} -N_{1x}^{(1)}\cdot cos\phi_2-N_{1z}^{(1)}\cdot sin\phi_2 \\ N_{1x}^{(1)}\cdot sin\phi_2-N_{1z}^{(1)}\cdot cos\phi_2 \\ -N_{1y}^{(1)} \\ 0 \end{pmatrix} \quad \ldots (3.6)$$

3.3.3 Spatial trajectory of the measuring probe

Equations (3.1) to (3.6) describe the curves on the wheel tooth surface that are to be measured. These curves are target contact traces of the measuring probe. Most coordinate measuring machines use a spherical probe and control the position of the probe centre. Due to its finite radius ρ the centre of a spherical probe is not at its target contact point on the wheel surface. To ensure that contact with the worm wheel flanks does occur at the target contact point, the centre of the spherical probe should be positioned at a distance ρ from the target contact point in direction of the normal to the surface. This is illustrated by Figure 3.5.

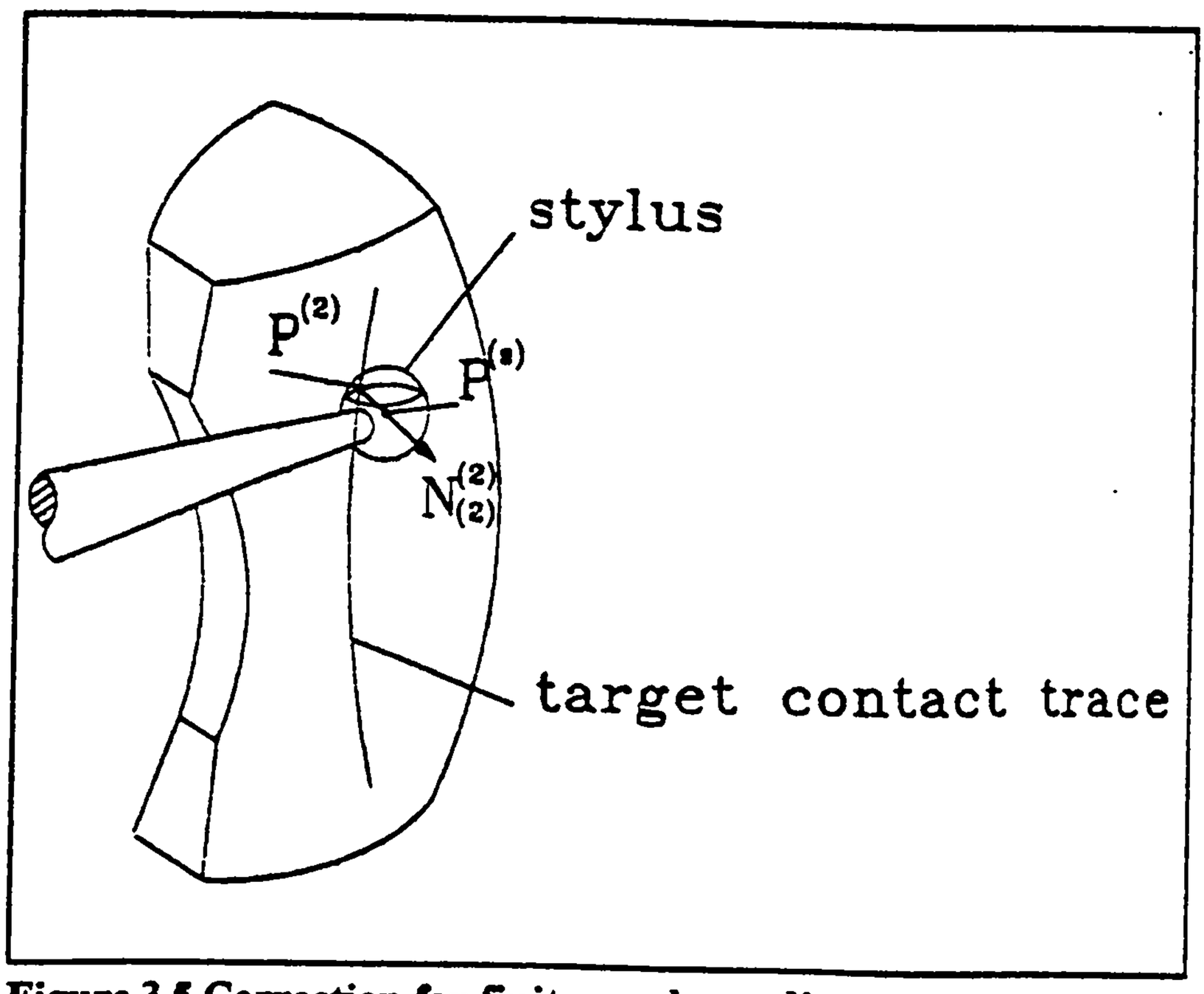


Figure 3.5 Correction for finite probe radius

When contact occurs at point $P^{(2)}$, as required, the probe centre will be at point $P^{(s)}$ (Figure 3.5), whose position vector is

$$r_2^{(s)} = r_2^{(2)} + \Delta r^{(s)} \quad \ldots \quad (3.7)$$

where

$$\Delta r^{(s)} = \begin{pmatrix} \pm\rho \cdot \dfrac{N_{2x}^{(2)}}{|N_2^{(2)}|} \\ \pm\rho \cdot \dfrac{N_{2y}^{(2)}}{|N_2^{(2)}|} \\ \pm\rho \cdot \dfrac{N_{2z}^{(2)}}{|N_2^{(2)}|} \\ 0 \end{pmatrix} \quad \ldots \quad (3.8)$$

where $N_2^{(2)}$ is directed towards into the wheel surface from space and $|N_2^{(2)}|$ is its magnitude. Since the arrangement of coordinates may vary from one CNC measuring machine to another, the signs in equation (3.8) must be chosen to suit the particular machine.

3.4 Practical Considerations

3.4.1 Choice of measurement with respect to the mating worm

It would be better if a worm wheel could be analytically measured with respect to both the mating worm and the hob. This would be easier in some respects since the general mathematical models presented in Chapter 2 could be used straight away. However, the following facts led to the decision to measure the worm wheel only relative to the mating worm at this time:

1) Precise machine settings are seldom available for the hobbing process, although the actual wheel flank profile generated is very sensitive to both the hob accuracy and the machine settings (refer to Chapter 5). Precise worm data is nevertheless always available. Getting reliable data for hobbing settings has also proved to be also very difficult in practice.

2) The CNC measuring machine used for this project is equipped with a slow computer (HP 9000/300). Therefore, the solution of the measurement equations needs to avoid iterative procedures. The solution is much slower if the hob parameters and hobbing settings are used to calculate a wheel tooth flank.

3) The effect of changing the hobbing settings and hob profile errors can be simulated by separate worm gearing analysis software to assist in rectifying a particular hobbing process. This can quickly indicate changes needed in the hobbing process, although this trial and error procedure is not the ideal way to cut accurate worm wheels.

4) Measurement relative to the mating worm is often more useful, since it can be taken before and after running, and the measurement results can still be used to investigate worm gear performance or transferred directly to analysis software. The curvature of the lead and profile traces is also a direct indication of the working gap profile (entry and exit gaps) between the worm and wheel.

3.4.2 Choice of the coordinate measuring machine

The arrangement of slide-ways may vary significantly from one coordinate measuring machine to another. The equations in this chapter are derived for a Gleason GMS 430 CNC gear measuring machine which has 3 linear axes and 1 rotary axis, as illustrated in Figure 3.6 (a) and (b). Described by the equations above in section 3.3, four separate motions of the machine are thus required to position the probe. However, it is possible to apply the equations to a different type of coordinate measuring machine without major changes.

3.4.3 Selection of Axial Datum

When measuring spur and helical gears, precise setting of the axial datum position is not critical, and because of the existence of the 'very thin spur gear' section, measurement of profile errors on convolute worm wheels is also relatively insensitive to the axial datum [Hu and Pennell, 1994], as shown in Figure 3.7 (a).

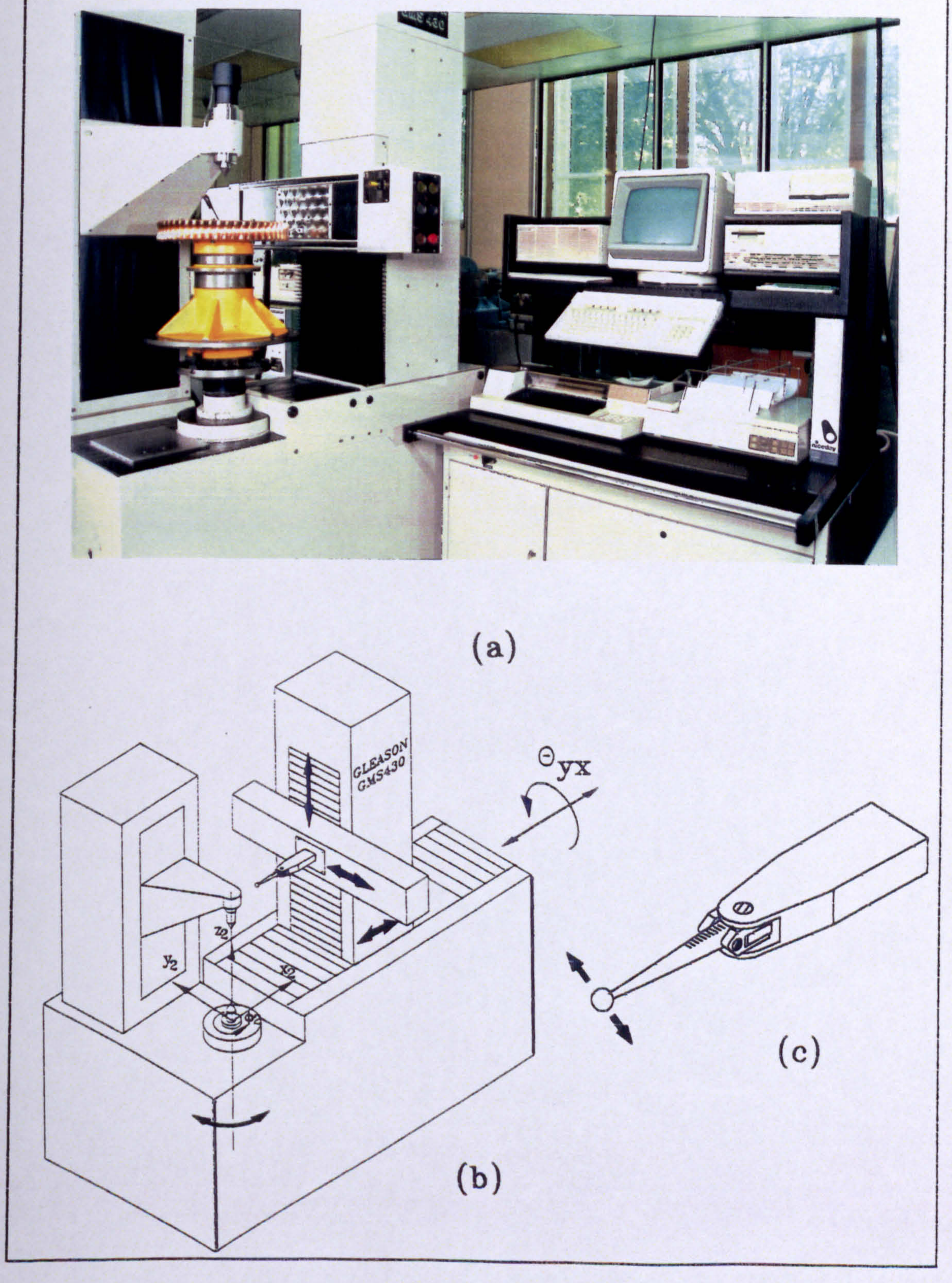


Figure 3.6 The Gleason GMS 430 Coordinate Gear Measuring Machine

When measuring worm wheel lead errors, however, the situation is quite different, since a datum error of 50 μm can easily give rise to, as a rough guide, measuring errors of 20 μm, even on single start worms [Hu and Pennell, 1994]. Accurate setting of the datum for lead measurement is thus essential.

If the axial datum face on the worm wheel is probed before measuring the lead errors, the mid-plane of the teeth can be located using the offset given on the wheel drawing. If the actual offset is incorrect, however, large lead errors will be measured on a wheel whose tooth form is otherwise `perfect'. Figure. 3.7(b), for example, shows lead traces on a wheel measured relative to three different datums separated by only 50 μm.

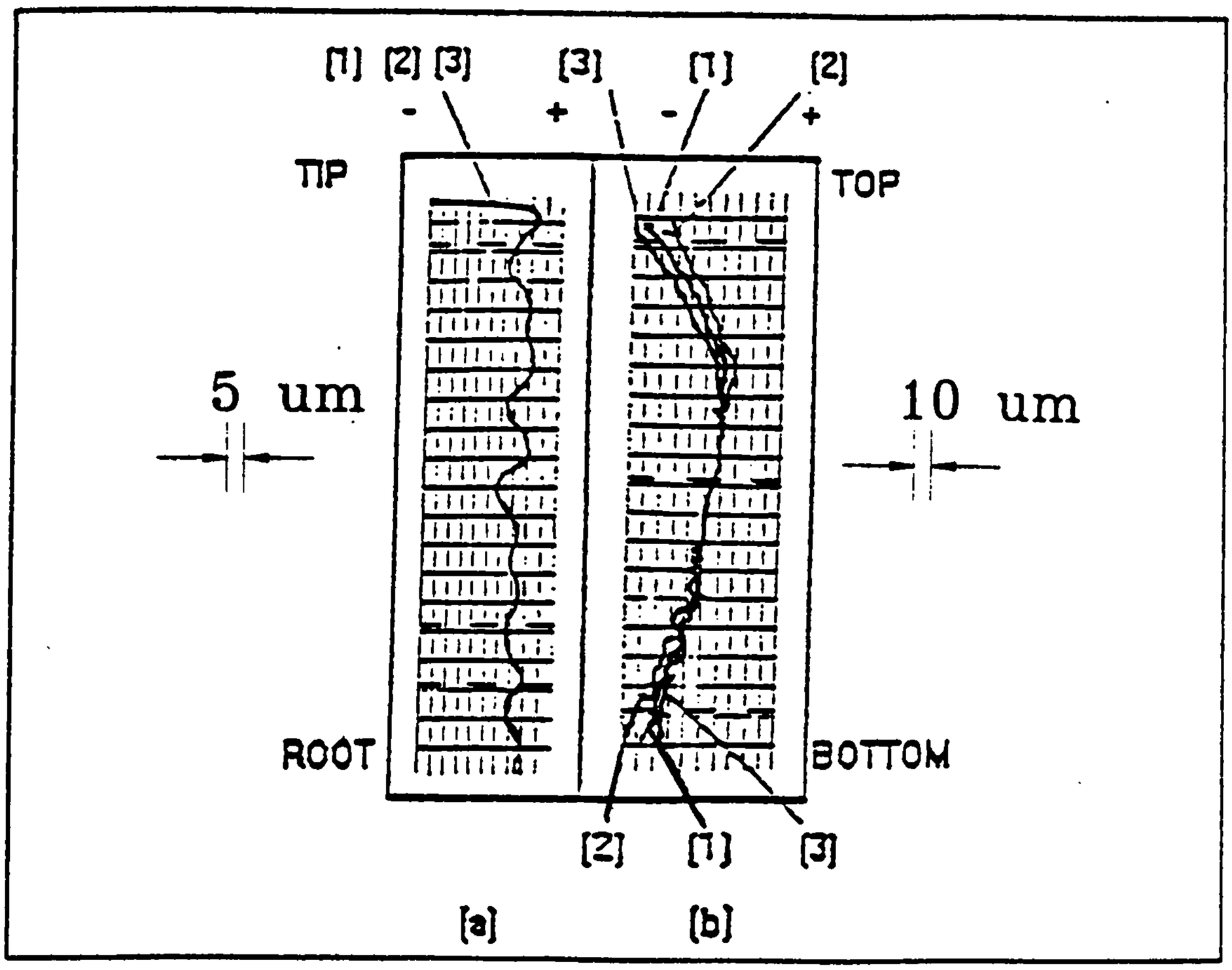


Figure 3.7 Profile and lead deviations measured relative to theoretical geometry at different axial datums. (a) Profile (b) Lead.

(1) Measured at nominal axial datum,

(2) Measured at 50 μm above the nominal datum,

(3) Measured at 50 μm below the nominal axial datum.

Rather than rejecting a wheel with an incorrect datum offset, it is possible to use the theory to determine the correct datum offset for the gear being measured. This can be done by adjusting the offset to minimize the lead errors. Use of this actual datum face offset in place of its nominal value greatly reduces (and may eliminate altogether) the amount of axial adjustment required during assembly. This is particularly helpful for hardened worm wheels since hardened tooth surfaces can not be bedded-in and the running contact will stay as its initial contact.

3.4.4 Range of Measurement

To avoid possible collisions with the tooth fillet surfaces, the range of measurement must be limited. If the lead measurement range is chosen as the effective face width b_{eff} of the worm wheel (or the actual face width b_2, whichever is least), contact between the probe and the root section of the worm wheel cannot occur, since b_{eff} is exactly the width of the intersection between the pitch points plane and the worm tip cylinder.

The profile measurement range is determined by the full worm tooth depth at the user-selected transverse section. The worm tip generates the nominal start of active profile of the wheel. The wheel tip should never reach the worm root due to the clearance. However, on the GMS 430, the measurement process will automatically stop at the wheel tip because the probe reading goes out of range.

The software also checks the radius of the probe against the clearance (defined by BS 721) and assures no collision will occur during measurement procedure.

3.4.5 Measuring Probe Limits

The probe used (Tesa GT-31, bi-directional, single axis probe) only measures in one direction, as shown in Figure 3.6(c), and so indicates deviations in the wheel transverse section. It must be corrected for each individual point on the wheel flank when normal deviations are required since the direction of the normal to the wheel flank varies throughout the wheel tooth flank. This is quite different from measuring an involute spur or helical gear for which the normal direction is constant and always in the base tangent

plane.

A further difficulty arises when measuring wheel lead errors on the GMS 430 because of the limited probe measuring range of ± 102 μm. This range is normally adequate for measuring spur and helical gears. However, it is too small for measuring lead deviations on worm wheels because of the relatively large entry/exit gaps. The problem has been overcome by rotating the wheel to give a probe reading of 'zero' at each contact point and taking the angle encoder reading on the rotary axis. The deviation is then calculated from the difference between the theoretical and actual encoder readings. Any actual probe reading deviation from the nominal 'zero' is corrected.

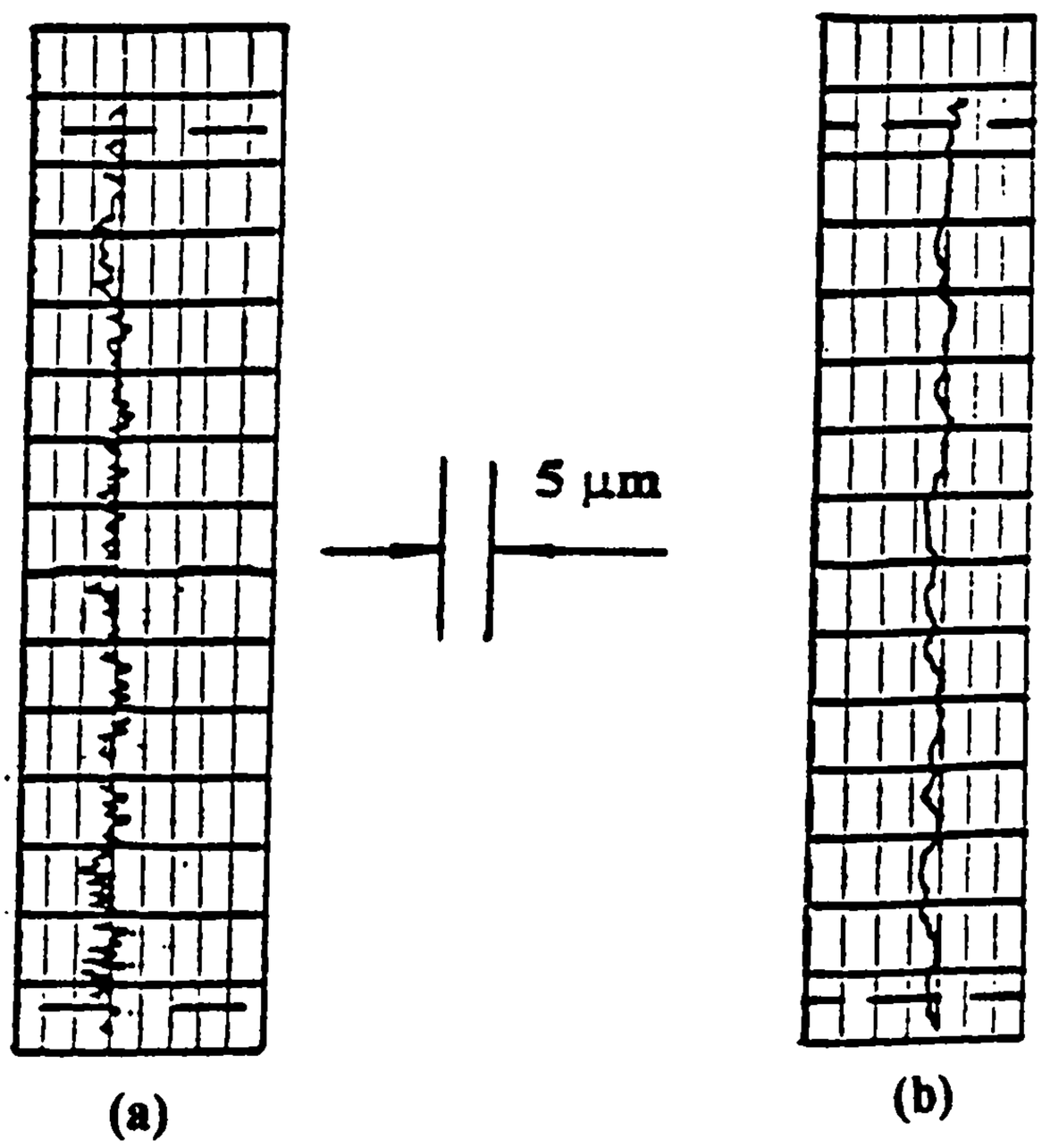


Figure 3.8 Filtering the ripples caused by the measuring system
(a) Before filtering. (b) After filtering

3.4.6 Ripples on measurement results of heavy wheels

High frequency ripples with a magnitude up to ± 2 μm were noted in measured traces on heavy wheels (see Figure 3.8 (a)). These ripples look like form errors or surface texture but, in fact, they are mainly caused by the measuring system. One of the differences between the measurement procedure for worm wheels and that for spur and helical gears, is that the rate of worm wheel rotation and axial motions of the probe must, in this case, be non-uniform and discontinuous. Errors caused possibly by inertia or friction of the moving parts may thus be quite large. After investigating this phenomenon, it was concluded that this was caused by torsional wind-up of the machine spindle due to static friction while 'nulling' the probe reading. A null-band range of ±2μm was applied to the probe, the residual probe reading being accounted for as part of the measurement process. These measuring errors are of significant concern when using the measured flank to analyse worm gear performance. A peak of form error could result, during software analysis, in an enormous and spurious theoretical contact stress. These high frequency ripples should not be counted as wheel flank errors in any way, since a bronze wheel flank surface will soon be polished during bedding-in. This problem has been solved by use of a digital filter. This filter will not affect the validity of the measured result.

Figure 3.8 shows the results before and after filtering.

3.4.7 Symmetry of the two flanks

It is beneficial to be aware that the geometry of the two flanks is symmetric about the wheel middle face plane. Once the geometry of one flank is known, that of the other can be obtained by simply "reversing" the wheel tooth in the software.

3.5 Validation of measurement results

Since analytical measurement of worm wheels has been achieved for the first time, no comparison with any previous publications or documents is possible. Very careful

validation of this measuring procedure has therefore been carried out over five years using four different methods.

The following sections will represent the validation methods and results. The idea of these methods is based on the fact that measurement results of a wheel flank should agree with each other no matter how the wheel is measured. However, it needs to be realized that the measurement results can agree with each other only to a certain extent because of the accuracy limits of the measuring machines. For the Gleason GMS 430 used, the measuring uncertainty U_{95} (determined by calibration tests) is known to be $\pm$ 1.8 μm over 100 mm face width for lead ($f_{H\beta}$) and $\pm$ 1.5 μm for profile ($f_{H\alpha}$) measurements on spur/helical master gears and artifacts.

3.5.1 Method 1: Change of measuring probe

As discussed above, the coordinate measuring machine positions the probe centre to ensure contact at the correct point on a wheel tooth flank. The wheel tooth flank and measurement result should thus be independent of the measuring probe used. If the calculated geometry is incorrect, the actual contact points on a wheel flank will be wrong and measurement results will vary with the probe radius and shape. Figure 3.9 shows results from a wheel tooth flank measured by using a disc probe and two spherical probes of different radius. The differences between these measurement traces are all within the machine's measuring uncertainty.

3.5.2 Method 2: Measurement of the involute section of a ZI worm wheel

There is an involute section on a ZI worm wheel flank. This section can easily be measured as a `very thin involute gear', even on conventional mechanical involute measuring machines (or on CNC machines without special `worm wheel' software) [Hu and Pennell, 1994]. Results for the profile in this special section, measured respectively by the worm wheel measuring software and the involute gear measuring software, are compared in Figure 3.10. Although the measuring principles and procedures of the two packages of software are completely different, the results are identical.

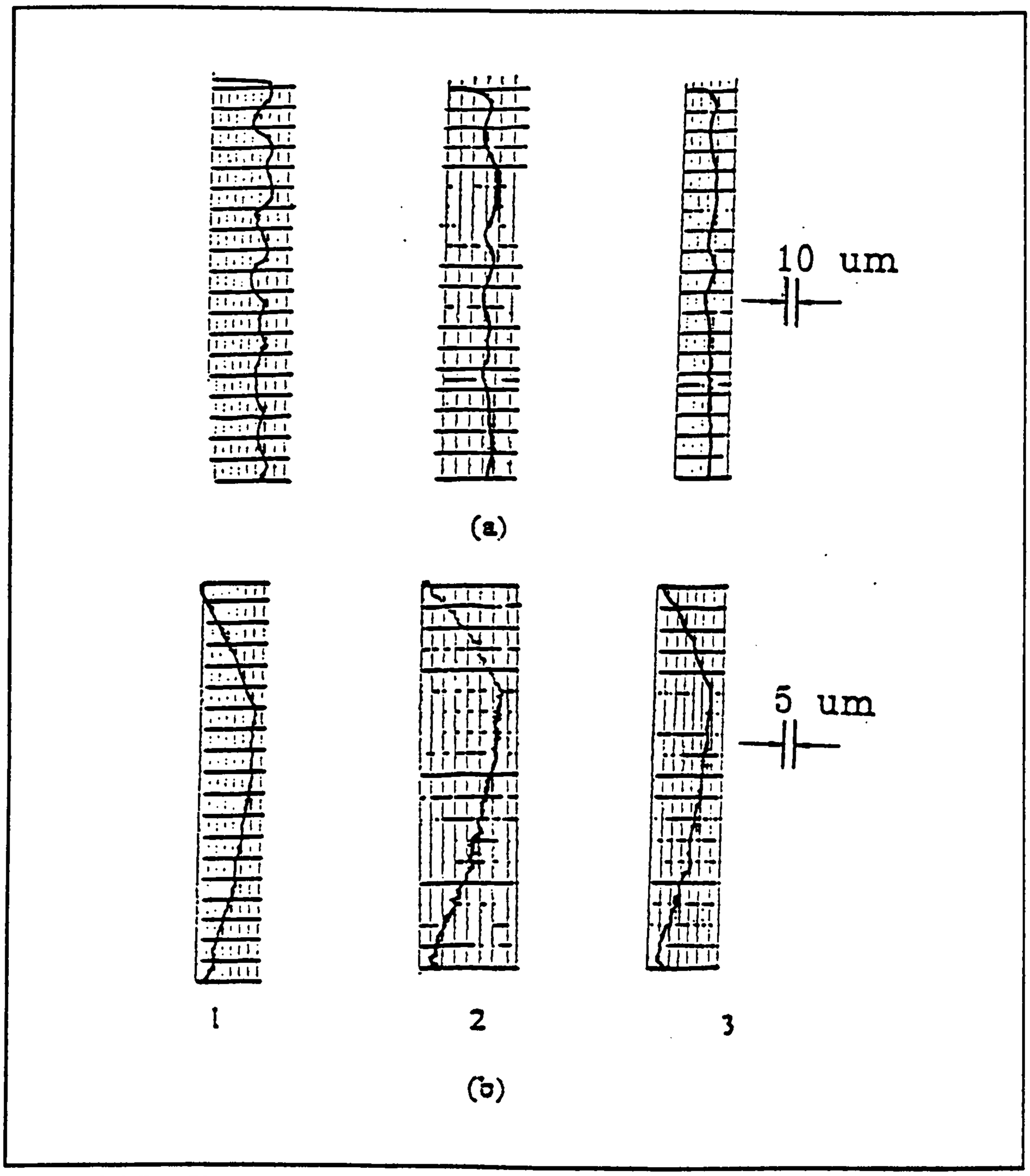


Figure 3.9 Profile and lead deviation measured with different probes
(a) Profile. (b) Lead.
(1) Measured with a 0.809 mm disc stylus,
(2) Measured with a 1.0 mm spherical stylus,
(3) Measured with a 2.0 mm spherical stylus,

Probe	0.809 mm disc	1.0 mm sphere	2.0 mm sphere
fHα	1.3 μm	-1.3 μm	-2.3 μm
fHβ	-27.7 μm	-27.8 μm	-28.6 μm

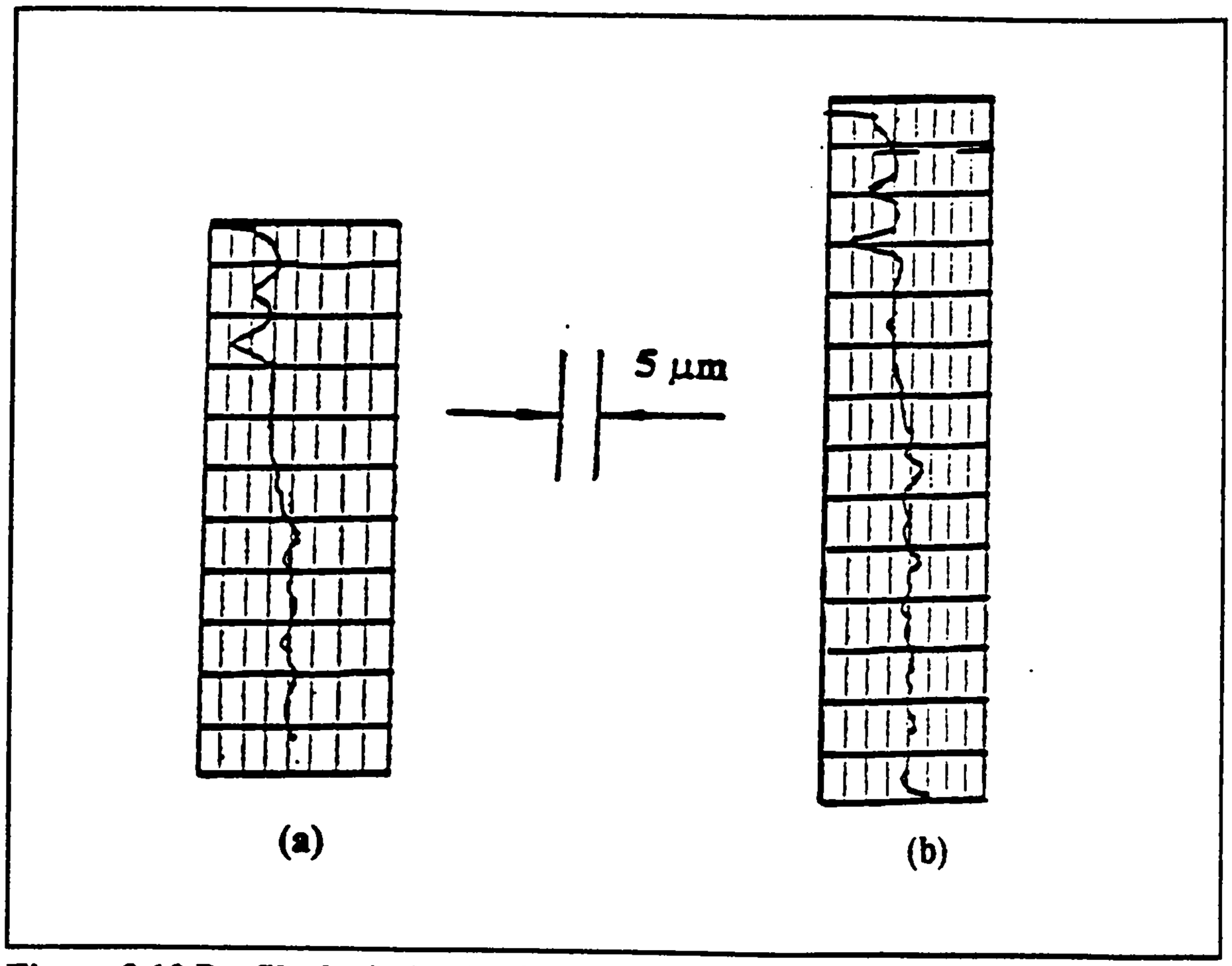


Figure 3.10 Profile deviation in the involute section
(a) Measured with the worm wheel measurement software
(b) Measured with the involute gear measurement software

Note: The involute gear measurement software plots the profile trace against length of roll. However, the worm wheel measurement software plots the profile trace against the radius. This causes the difference in the length of the two traces.

3.5.3 Method 3: Comparison of measurement results with blue marking patterns

Measured lead and profile traces should correspond to the blue marking patterns obtained from industrial blue marking tests. The comparison presented here was obtained by Stirzaker (1994) and the blue marking test was carried out by Highfield Gears Ltd. Huddersfield. Figure 3.11 shows that blue patterns appeared where the lead and profile traces showed plus metal (indicating the minimum contact gap).

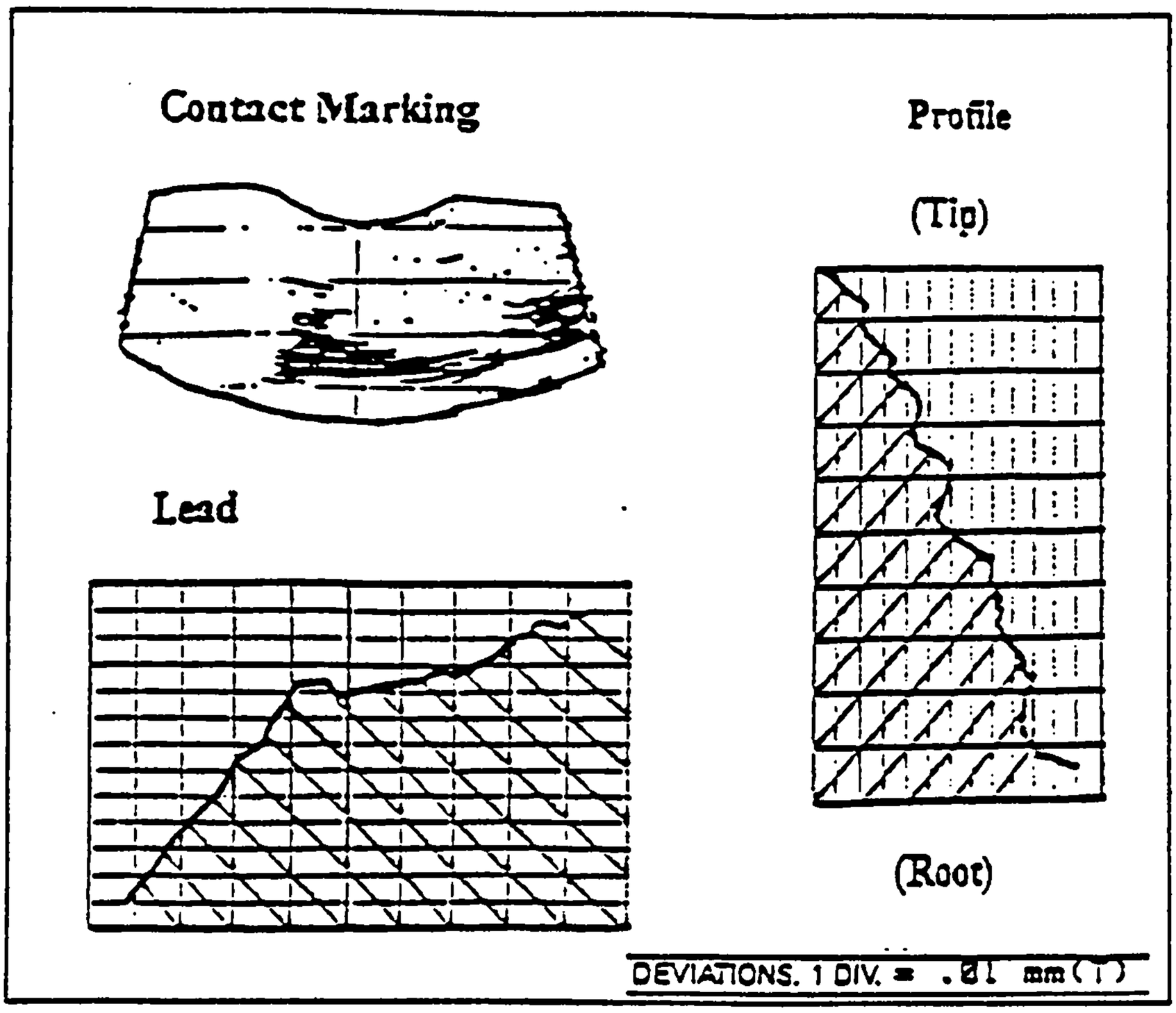
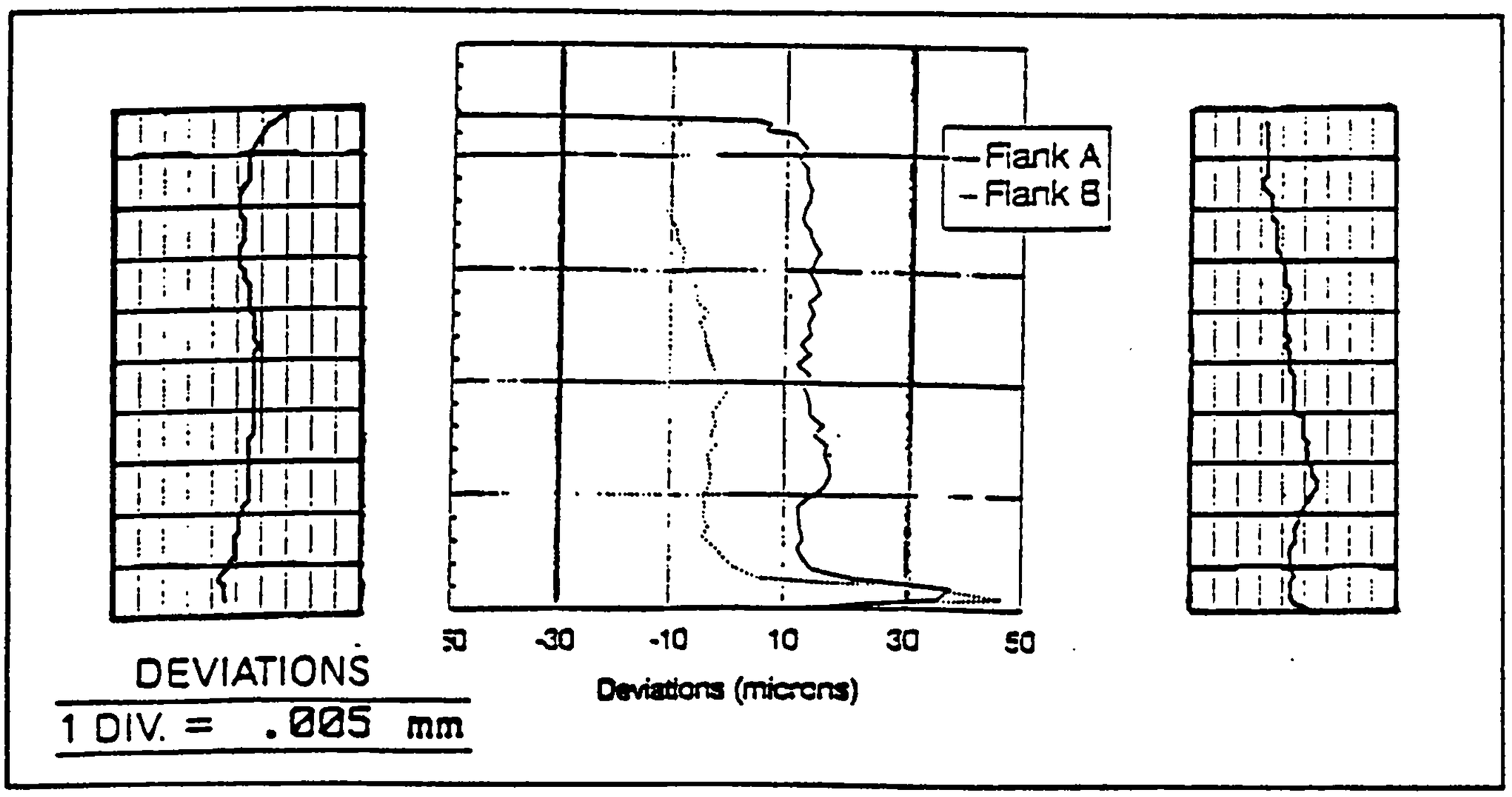


Figure 3.11 Comparison between measurement traces and blue pattern

3.5.4 Method 4: Comparison between wheel and fly tool measurement

Measurement traces on a worm wheel cut on an accurate hobbing machine should reflect the profile of its fly tool. The comparison shown in Figure 3.12 was carried out by Fish (1995) at the University of Huddersfield using a fly tool measured by Holroyd Machine Tool, Rotors & Precision Gears. The worm wheel was measured at Newcastle University by the author.



Wheel flank B **Fly tool** **Wheel flank A**

Figure 3.12 Comparison between wheel profile and fly tool profile

CHAPTER 4

NON-CONJUGATE ANALYSIS: WORM AND WHEEL MESH

The worm wheel hob is usually a few percent oversize relative to the worm. As a result of this oversize, some parameters of a worm wheel hob are different from those of the worm and the relative position settings for the hobbing process are also different from those of the worm gear unit. These differences are called mismatches in this thesis.

Two gears are said to have a conjugate action if they transform motion with a prescribed relationship between Φ_1 and Φ_2, so that $\Phi_2 = f(\Phi_1)$ [Litvin,1994]. (Usually I = constant and $\Phi_2 = \Phi_1 \cdot I$, but conjugate gearing with a variable transmission ratio can also be designed, for instance, unequal pitch worm gears [Zhang and Hu, 1989]). If there are no mismatches, a pair of worm gears will have conjugate action under no load. However, real worm gears normally have non-conjugate action under no load because of the mismatches. The conjugate analysis, described in Chapter 2, based on a given hob/worm tooth flank and relative motion, generates the corresponding wheel tooth flank and determines the geometric relationship between the generating and generated surfaces. In the non-conjugate analysis, described in this chapter, it is assumed that the profiles of both worm and wheel tooth flanks are known. The purpose of non-conjugate analysis is to calculate the transmitted motion and determine the geometric relationships between the driving and driven surfaces during this motion.

The fundamental idea of the non-conjugate analysis in this thesis is to bring the wheel tooth surface into contact with the worm flank at each successive angle of worm rotation. The transmission error and the gaps round the contact point can then be determined. Although this thesis is mainly concerned with cylindrical ZI worms, the methods developed herein can be applied for both cylindrical and globoidal worm gearing since the angle Φ_1 of worm rotation (instead of the notional worm thread axial translation used

by Colbourne (1989)) is used to determine the contact points. In addition, the effects of worm and wheel tooth form and manufacturing errors can be allowed for (as discussed in chapter 5).

The main objective of this chapter is thus to predict the worm gear performance under no-load conditions, and to facilitate the interpretation of the results of the worm gear tooth form errors measured as described in chapter 3.

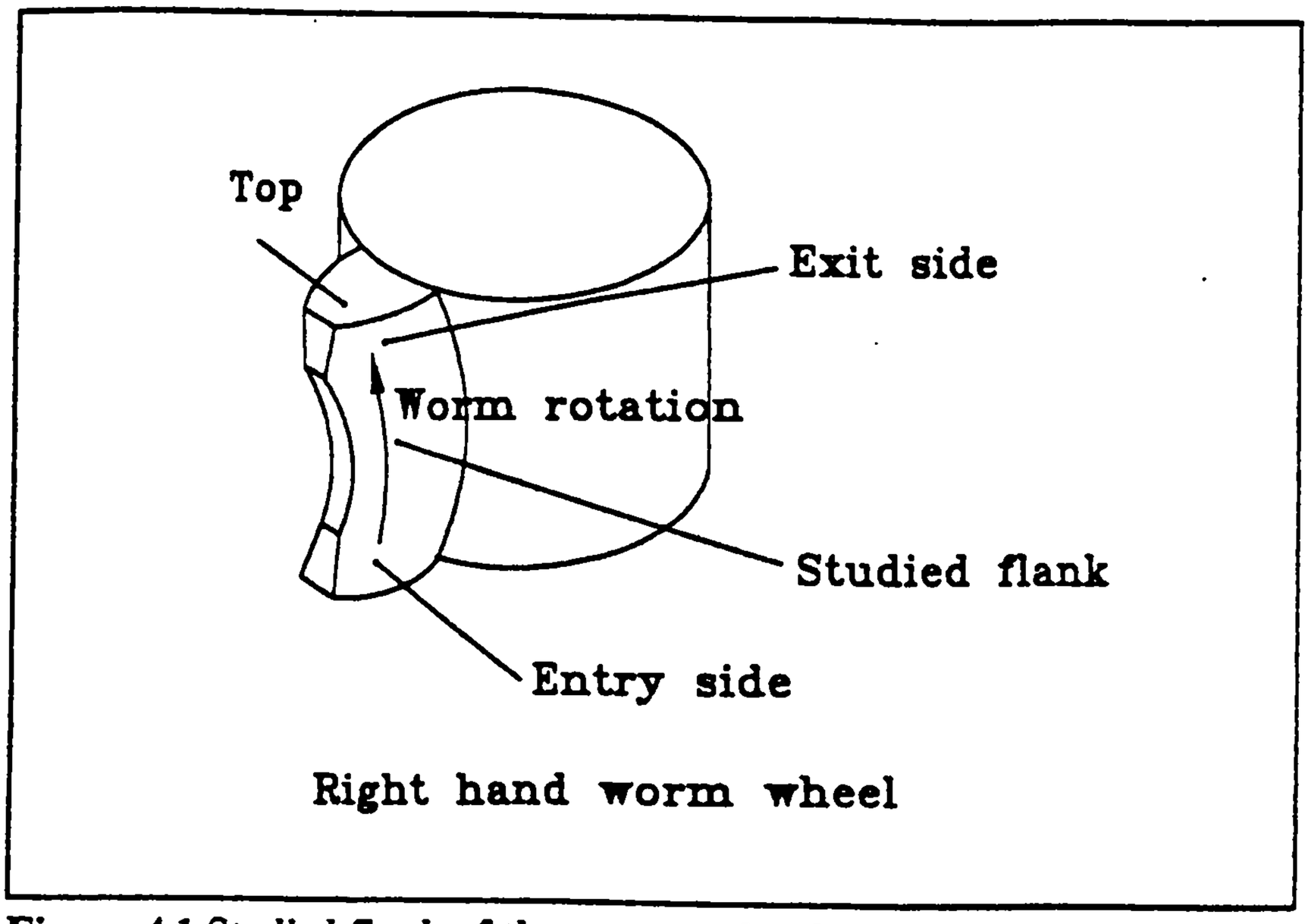


Figure 4.1 Studied flank of the worm wheel

The problems to be solved in this chapter are to:

(1) Generate a general model for defining the geometry of the worm/hob helicoids,

(2) Determine the instantaneous contact points and transmission error for a given worm rotation angle Φ_1,

(3) Determine the instantaneous contact gap contour at a given value of Φ_1 and predict the overall contact marking pattern,

(4) Calculate the entry and exit gaps,

(5) Predict the total contact area at any instant (for comparing the load capacity of different designs),

(6) Predict the wear marking on the worm flanks,

The coordinate systems used in this chapter are the same as those in chapter 2 (Figure 2.1). Right hand worm gears are studied with the worm driving the worm wheel. The worm wheel flank studied is shown in Figure 4.1. The analysis of left hand worms and of those rotating in the opposite sense requires only minimal changes in the presentation of final results.

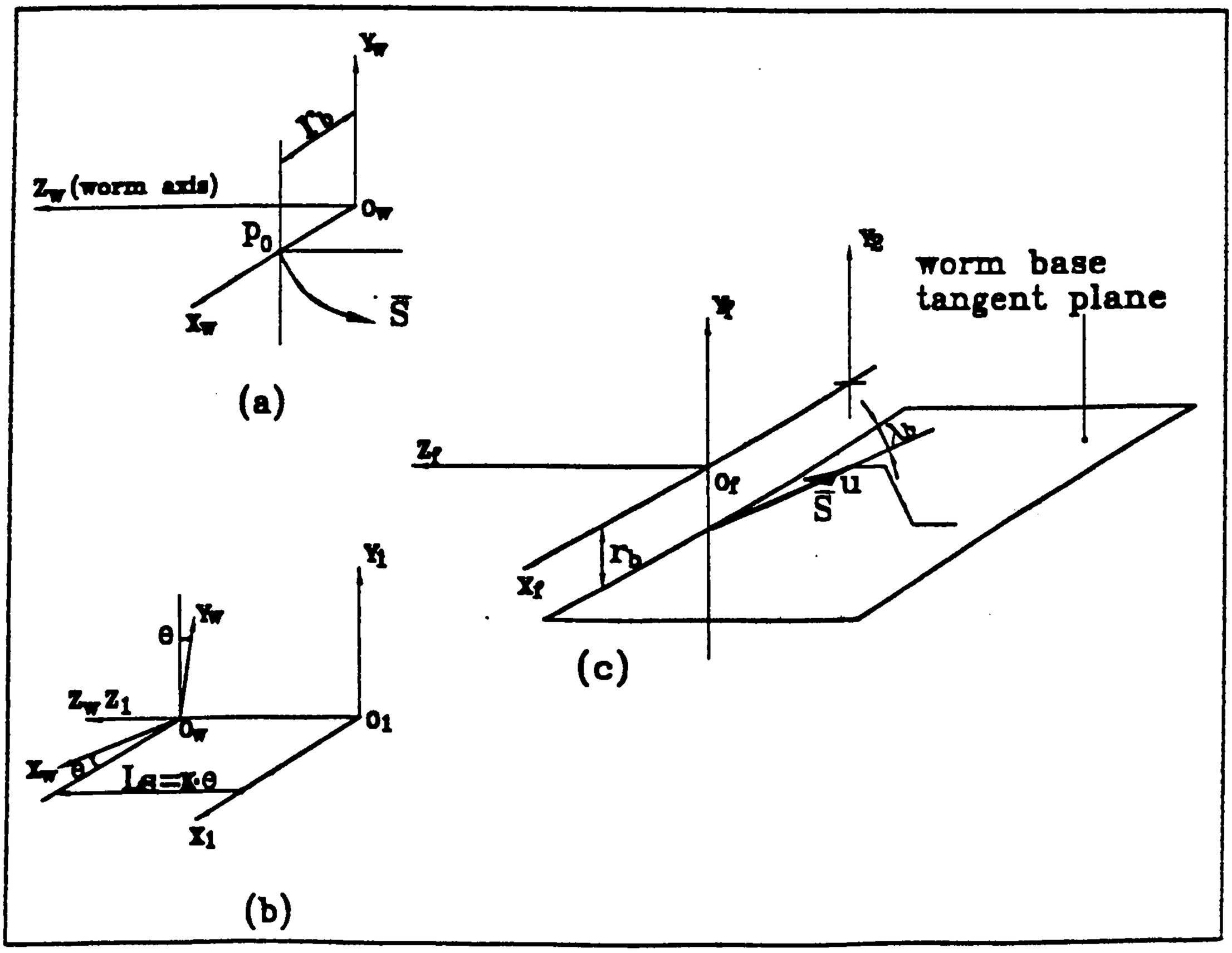


Figure 4.2 Coordinate systems for generation of worm helicoid

4.1 General Model of a Cylindrical Worm Thread

A worm flank is a helicoid surface generated by screw motion of a curve. Without loss of generality, a planar curve **s** can be used to generate a worm surface, as shown in Figure 4.2. The coordinate system $S_w(O_w, x_w, y_w, z_w)$ is rigidly connected to the planar curve **s**. The planar curve **s** can be represented in terms of components of coordinate system $S_w(O_w, x_w, y_w, z_w)$ in form (4.1)

$$\boldsymbol{s} = \left(r_b, y_w, z_w, 1 \right)^T \quad \ldots\ldots \quad (4.1)$$

Where r_b (= O_wP_0 in Figure 4.2 (a)) is a constant for cylindrical worms. Equation (4.1) gives the worm axial profile when $r_b = 0$.

The planar curve **s** has a screw motion, with lead p_z, along $axis_w z$. Its linear displacement L_s during the screw motion is proportional to its angular displacement θ, so that

$$L_s = K \cdot \theta, \quad \ldots\ldots \quad (4.2)$$

where $K = p_z / (2\pi)$.

The screw motion of the planar curve **s** forms a helicoid surface represented in terms of components of coordinate system $S_1(O_1, x_1, y_1, z_1)$ of Figure 4.2 (b). The coordinate systems $S_w(O_w, x_w, y_w, z_w)$ and $S_1(O_1, x_1, y, z_1)$ are coincident at the starts of the screw motion (when $\theta = 0$).

The coordinate transformation matrix from $S_w(O_w, x_w, y_w, z_w)$ to $S_1(O_1, x_1, y_1, z_1)$, as shown in Figure 4.2 (b), is given by

$$T_{1w} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & K\cdot\theta \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \ldots\ldots \quad (4.3)$$

The coordinates of points on the generated helicoid surface represented in system $S_1(O_1, x_1, y_1, z_1)$ can then be obtained from

$$r_1^{(1)} = T_{1w} \cdot s \qquad \ldots\ldots \quad (4.4)$$

from which it follows that

$$r_1^{(1)} = \begin{pmatrix} r_b \cdot cos\theta - y_w \cdot sin\theta \\ r_b \cdot sin\theta + y_w \cdot cos\theta \\ K \cdot \theta + z_w \\ 1 \end{pmatrix} \qquad \ldots\ldots \quad (4.5)$$

Equation (4.5) is valid for any helicoid. It represents a convolute helicoid (ZA, ZN and ZI) if the planar curve **s** is a straight line [Buckingham, 1949] or a concave helicoid if the planar curve **s** is a circular arc (a concave helicoid has a circular arc axial profile) [Winter and Wilkesmann, 1981]. The planar curve **s** is generally a function of θ on a globoidal worm.

When the planar curve **s** is a straight generating line, equations (4.1) and (4.5) reduce to

$$s = \left(r_b, \; -u \cdot cos\lambda, \; -u \cdot sin\lambda, \; 1 \right)^T \qquad \ldots\ldots \quad (4.6)$$

$$r_1^{(1)} = \begin{pmatrix} r_b \cdot cos\theta + u \cdot cos\lambda \cdot sin\theta \\ r_b \cdot sin\theta - u \cdot cos\lambda \cdot cos\theta \\ K \cdot \theta - u \cdot sin\lambda \\ 1 \end{pmatrix} \qquad \ldots\ldots \quad (4.7)$$

where λ is the angle of generation (the angle between the straight line generator and transverse planes perpendicular to the helicoid's axis). r_b is the radius of the base

cylinder of the helicoid and u and θ are curvilinear parameters. (Figure 4.2 (c)).

Equation (4.7) represents:

1) a ZA helicoid if $r_b = 0$;

2) a ZN helicoid if λ is replaced by -λ;

3) a ZI helicoid if $\lambda = \gamma_b = \tan^{-1}(p_z / (2 \pi r_b))$, where γ_b is the worm base lead angle.

The angle of generation of a ZI worm helicoid is thus equal to its base lead angle.

The normal vector to a convolute worm flank can be obtained from equations (2.2) and (4.7). This gives

$$N_1^{(1)} = \frac{\partial r_1^{(1)}}{\partial u} \times \frac{\partial r_1^{(1)}}{\partial \theta} = \begin{pmatrix} -K \cdot cos\lambda \cdot cos\theta + r_b \cdot sin\lambda \cdot cos\theta + u \cdot cos\lambda \cdot sin\lambda \cdot sin\theta \\ r_b \cdot sin\lambda \cdot sin\theta - u \cdot cos\lambda \cdot sin\lambda_b \cdot cos\theta - K \cdot cos\lambda \cdot sin\theta \\ u \cdot cos^2\lambda \\ 0 \end{pmatrix} \quad \ldots (4.8)$$

$N_1^{(1)}$ is directed outwards from the worm surface.

4.2 Hob design and Hobbing settings

A hob is usually designed with oversize to allow for resharpening and thus give longer life. Consequently, the hobbing centre distance usually differs from the centre distance of the worm gear unit, and the hob axis is set with a swing angle relative to the worm axis. More importantly, the hob oversize should be understood as a means of introducing modifications to improve the performance of the worm gear drive. The

oversize introduces deliberate mismatch to minimize the effects of deflection under load, small errors in the axial position of the worm wheel, and other mounting tolerances. It also establishes a gap profile between the worm and wheel teeth for easier entry of lubricant into the mesh [AGMA 6022 - C93] [Janninck, 1987].

Janninck (1987) introduced three methods of hob design. Those methods are the axial pitch design method, the normal pitch design method and the extra thread method. Colbourne(1989) proposed a new concept of hob design defined by him as the base pitch design method.

More details of these methods of hob design are given by references [Janninck, 1987] and [Colbourne, 1989]. The details of the normal pitch method for hob design are also given in Appendix 3, since this method has practically become an industry standard [Janninck, 1987].

4.3 Contact points and transmission error of single tooth pair

To carry out a non-conjugate analysis with the method proposed in this thesis, the coordinates of a grid of points on wheel tooth flank and the normal to the wheel flank at each of these points should have already been calculated. A regular grid of points on a wheel tooth flank may be obtained by:

i) Conjugate analysis of the generation process as described in chapter 2;

ii) Analytical measurement of the wheel flanks as described in chapter 3;

iii) Any other independent method (for example, the tooth flank of an involute helical gear can be analysed).

Before bedding-in, the no-load contact between worm and wheel flanks is nominally a point contact. In this thesis, the following procedure is used to find this instantaneous contact point:

1) fix the angle of worm rotation Φ_1, as expressed by equation (4.9).

2) place the wheel at an arbitrary angular position Φ_2' (eg, at its previous position), as expressed by equation (4.10).

3) rotate the worm wheel until its tooth flank is in contact with the worm tooth flank, as expressed by equations (4.11) to (4.14).

When the wheel tooth flank is generated, the coordinates of the grid points on it are represented in terms of components of system $S_2(O_2, x_2, y_2, z_2)$. For the sake of convenience, we now transform the position vectors of the worm and wheel tooth flanks into stationary coordinate system $S_p(O_p, x_p, y_p, z_p)$. The fixed angle of worm rotation is denoted by Φ_1, and the initial (arbitrary) angle of wheel rotation by Φ_2'. The worm and wheel tooth flanks are then represented in terms of components of coordinates $S_p(O_p, x_p, y_p, z_p)$ by :

$$r_p^{(1)} = \begin{pmatrix} x_p^{(1)}(u,\theta,\phi_1) \\ y_p^{(1)}(u,\theta,\phi_1) \\ z_p^{(1)}(u,\theta,\phi_1) \\ 1 \end{pmatrix} = T_{p1}(\phi_1) \cdot r_1^{(1)} \quad \ldots\ldots\ (4.9)$$

$$r_p^{(2)} = \begin{pmatrix} x_p^{(2)}(\phi_2') \\ y_p^{(2)}(\phi_2') \\ z_p^{(2)}(\phi_2') \\ 1 \end{pmatrix} = T_{p2}(\phi_2') \cdot r_2^{(2)} \quad \ldots\ldots\ (4.10)$$

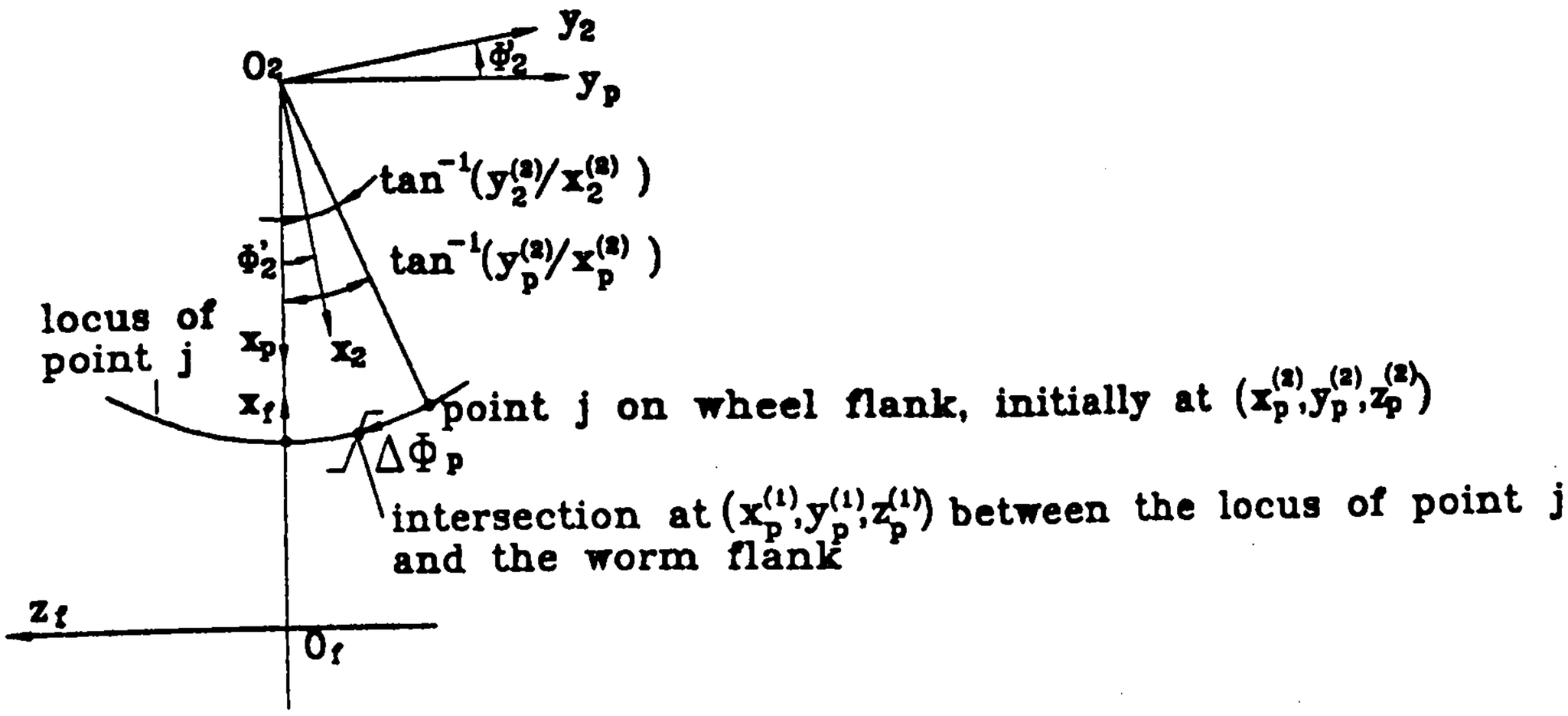
For a fixed angle Φ_1 of worm rotation, the worm surface is fully defined in coordinate

system $S_p(O_p, x_p, y_p, z_p)$ by the two curvilinear parameters (u, θ). In other words coordinates $(x_p^{(1)}, y_p^{(1)}, z_p^{(1)})$ are single valued functions of the curvilinear parameters u and θ, which are unknowns, since the objective at this moment is to find the contact point on the worm flank corresponding to the (known) grid point on the wheel flank. The actual angle of wheel rotation (denoted by Φ_2) corresponding to Φ_1 is the third unknown besides u and θ. The arbitrary angle Φ_2' of wheel rotation can be used instead of the actual value Φ_2 at the moment. The two unknowns u and θ can be found by solving the two equations (4.11) and (4.12), since the locus of each grid point on the wheel tooth flank is a circle of radius of R_2, in the wheel transverse section at $z_p^{(2)}$.

$$z_p^{(2)} = z_p^{(1)}(u,\theta) \qquad \ldots\ldots \qquad (4.11)$$

$$R_2 = \sqrt{(x_p^{(2)})^2 + (y_p^{(2)})^2} = \sqrt{(x_p^{(1)}(u,\theta))^2 + (y_p^{(1)}(u,\theta))^2} \quad \ldots\ldots\ (4.12)$$

Equations (4.11) and (4.12) can be interpreted as defining the point on the worm flank at which the circular path traced by the grid point j intersects the worm flank.



Intersection between the worm flank and the locus of a point on the wheel flank

Figure 4.3

We define the differential rotation required to bring a grid point on the wheel tooth flank onto the worm flank as the angular distance $\Delta\phi_p$ (Figure 4.3). Once the intersecting point $(x_p^{(1)}, y_p^{(1)}, z_p^{(1)})$ on the worm flank has been found, $\Delta\phi_p$ can be obtained from

$$\Delta\phi_p = \tan^{-1}(y_p^{(1)} / x_p^{(1)}) - \tan^{-1}(y_p^{(2)} / x_p^{(2)}) \qquad \ldots\ldots \qquad (4.13)$$

The differential rotation angle $\Delta\phi_p$ varies from one grid point to another. The minimum differential rotation angle among all the points is the wheel rotation needed to bring the wheel tooth flank into contact with the worm flank, so that the actual wheel angular position is:

$$\Phi_2 = \Phi_2' + \min \{\Delta\phi_p(j)\} \quad j = 1,2,3, \ldots N \quad \cdots \quad (4.14)$$

where j denotes the j'th grid point on the wheel flank and N denotes the total number of grid points on the wheel flank.

The next step is to rotate the worm to another angle Φ_1 and find the next Φ_2 by repeating the above procedures. The worm is rotated with a constant increment $\Delta\Phi_1$ until the wheel tooth goes out of mesh. For convenience of later calculations, it is best if $\Delta\Phi_1$ is chosen an integer factor of the angle equivalent to the worm axial pitch. This requires

$$\Delta\Phi_1 = 2 \cdot \pi/(Z_1 \cdot N_{pitch}) \qquad \cdots \quad (4.15)$$

where N_{pitch} is the number of steps of worm rotation per axial pitch.

In this way, a series of corresponding values of Φ_1 and Φ_2 can be calculated. The first Φ_1 and the first Φ_2 can be regarded as the datum values and denoted as $\Phi_1(0)$ and $\Phi_2(0)$ respectively. Each individual angle of worm rotation $\Phi_1(n)$ has a corresponding angle of wheel rotation $\Phi_2(n)$. The transmission error is defined as the deviation in the angular position of the wheel (for any given angular position of the worm), relative to the angular position that it would occupy if the worm and wheel were conjugate, geometrically perfect and undeformed. From this definition, the transmission error at the n'th step of worm rotation can be obtained from

$$f_i'[n] = \phi_2(n) - \phi_2(0) - [\phi_1(n) - \phi_1(0)] \cdot i \quad \ldots\ldots \quad (4.16)$$

where i is the transmission ratio, and $f_i'[n]$ is the single tooth transmission error at the

n'th step of worm rotation.

The method proposed above can obviously be applied for line contacts, but more than one contact point will be found at each instant.

Because of the finite spacing of the grid points on the worm wheel flank, equation (4.14) normally only gives an approximation to Φ_2, since the actual contact point will not, in general, occur precisely at a grid point (If it does, (4.14) gives Φ_2 exactly). Numerical tests have shown, however, that, for the fine grid spacing normally used, the errors involved in using (4.14) are negligible. However, if higher accuracy is required (eg. when a coarse grid is used) polynomial or cubic spline interpolation of the values of $\Delta\Phi_1(j)$ near $\min\{\Delta\Phi_1(j)\}$ is clearly possible to obtain a better estimate of Φ_2. Standard library procedures could be used for this purpose.

4.4 Axial and normal gap of a single tooth pair

The analysis presented below gives the "no-load gap" at each grid point on the wheel flank at an input worm angle $\Phi_1(n)$. The axial gap is defined along the direction parallel to the worm axis. The wheel flank is assumed to be in contact with the worm flank at one point with $\Phi_1(n)$ and $\Phi_2(n)$ already obtained. Obviously the gap at the contact point is zero.

For ease of presentation, let us transform the coordinates of the worm and wheel flanks into the stationary coordinate system $S_f(x_f, y_f, z_f)$. The calculation procedure is given by equations (4.17) to (4.21). For each grid point at $r_2^{(2)}$ on the wheel tooth flank, equations (4.19) and (4.20) determine the corresponding axially displaced point $r_1^{(1)}$ on the worm tooth flank. It must be noticed that this point on the worm tooth flank is nominally different from the one obtained by equations (4.11) and (4.12), although this difference may possibly be negligible in some cases.

$$r_f^{(1)} = \begin{pmatrix} x_f^{(1)}(u,\theta) \\ y_f^{(1)}(u,\theta) \\ z_f^{(1)}(u,\theta) \\ 1 \end{pmatrix} = T_{f1}(\phi_1) \cdot r_1^{(1)}(u,\theta) \quad \ldots\ldots\ldots \quad (4.17)$$

$$r_f^{(2)} = \begin{pmatrix} x_f^{(2)} \\ y_f^{(2)} \\ z_f^{(2)} \\ 1 \end{pmatrix} = T_{f2}(\phi_2) \cdot r_2^{(2)} \quad \ldots\ldots\ldots \quad (4.18)$$

$$x_f^{(1)}(u,\theta) = x_f^{(2)} \quad \ldots\ldots\ldots \quad (4.19)$$

$$y_f^{(1)}(u,\theta) = y_f^{(2)} \quad \ldots\ldots\ldots \quad (4.20)$$

$$\delta_a = z_f^{(1)}(u,\theta) - z_f^{(2)} \quad \ldots\ldots\ldots \quad (4.21)$$

where δ_a is the single tooth axial gap.

The normal gap δ_n is defined along the normal to the wheel surface and can be obtained by the following procedure:

1) Transform the coordinates of each grid point on the wheel surface and rotate the corresponding normal to the wheel flank into coordinate system $S_f (x_f, y_f, z_f)$;

2) Define a straight line along the normal to the wheel flank at a point (x_f^2, y_f^2, z_f^2);

3) Transform the coordinates of the worm flank into coordinate system $S_f (x_f, y_f, z_f)$;

4) Determine the intersection between the straight line and the worm flank. The intersecting point on the worm flank is presented in the form $(x_f^{(1)}, y_f^{(1)}, z_f^{(1)})$;

5) Determine the magnitude of normal gap which is the distance measured from the point $(x_f^{(1)}, y_f^{(1)}, z_f^{(1)})$ on the worm flank to the point $(x_f^{(2)}, y_f^{(2)}, z_f^{(2)})$ on the wheel flank;

Equations (4.22) to (4.24) determine the coordinates of the intersecting point between the normal to the wheel surface and to the worm flank. t is a parameter for computation, and will be the size of the normal gap if $N_f^{(2)}$ is a unit normal.

$$x_f^{(1)}(u,\theta)=x_f^{(2)} + t\cdot N_{xf}^{(2)} \quad \ldots\ldots \quad (4.22)$$

$$y_f^{(1)}(u,\theta)=y_f^{(2)} + t\cdot N_{yf}^{(2)} \quad \ldots\ldots \quad (4.23)$$

$$z_f^{(1)}(u,\theta)=z_f^{(2)} + t\cdot N_{zf}^{(2)} \quad \ldots\ldots \quad (4.24)$$

$$\delta_n=\sqrt{(x_f^{(2)}-x_f^{(1)})^2+(y_f^{(2)}-y_f^{(1)})^2+(z_f^{(2)}-z_f^{(1)})^2} \quad \ldots\ldots \quad (4.25)$$

where δ_n is the single tooth normal gap. Equations (4.22), (4.23) and (4.24) must be solved numerically for t, u, and θ so that the coordinates of the intersecting point can be found. The normal gap magnitude can then be obtained by equation (4.25).

In the neighbourhood of the contact point, the two mating surfaces are almost parallel, so that the normal to the wheel flank is similar to that to the worm flank ($N^{(1)} \approx -N^{(2)}$) in this region. As a result of this, having found the axial gap from equation (4.21) the normal gap also follows easily from

$$\delta_n \approx |\delta_a \cdot n_f^{(1)}|$$

where

$$\delta_a = [0, 0, \delta_a, 1]$$

and $n_f^{(1)}$ is the unit normal obtained from

$$n_f^{(1)} = N_f^{(1)} / | N_f^{(1)} |$$

(Note that, 'normal' in this case means 'normal to the worm flank'.)

If a ZI worm is used, the normal gap can be found more easily from

$$\delta_n \approx \delta_a \cdot \cos(\lambda_b)$$

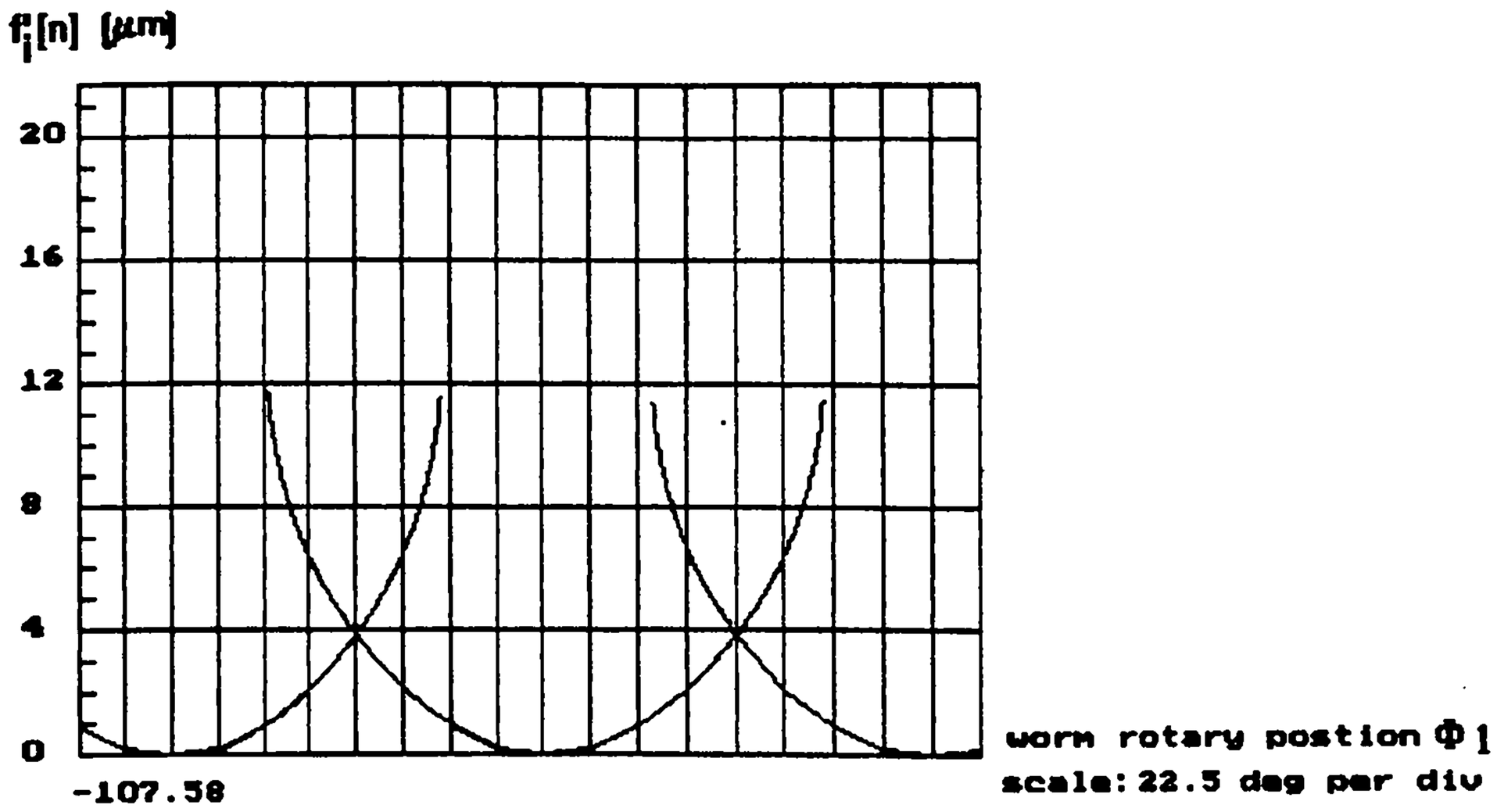
since the angle between the axis of a ZI worm and the normal to the worm flank is constant and equal to the base lead angle [Hu and Pennell, 1994].

If the worm base lead angle is, say, less than 25^0, the axial gap can thus be used as an approximation of the normal gap for certain applications (prediction of contact pattern etc), since $\cos(\lambda_b) \approx 1.0$.

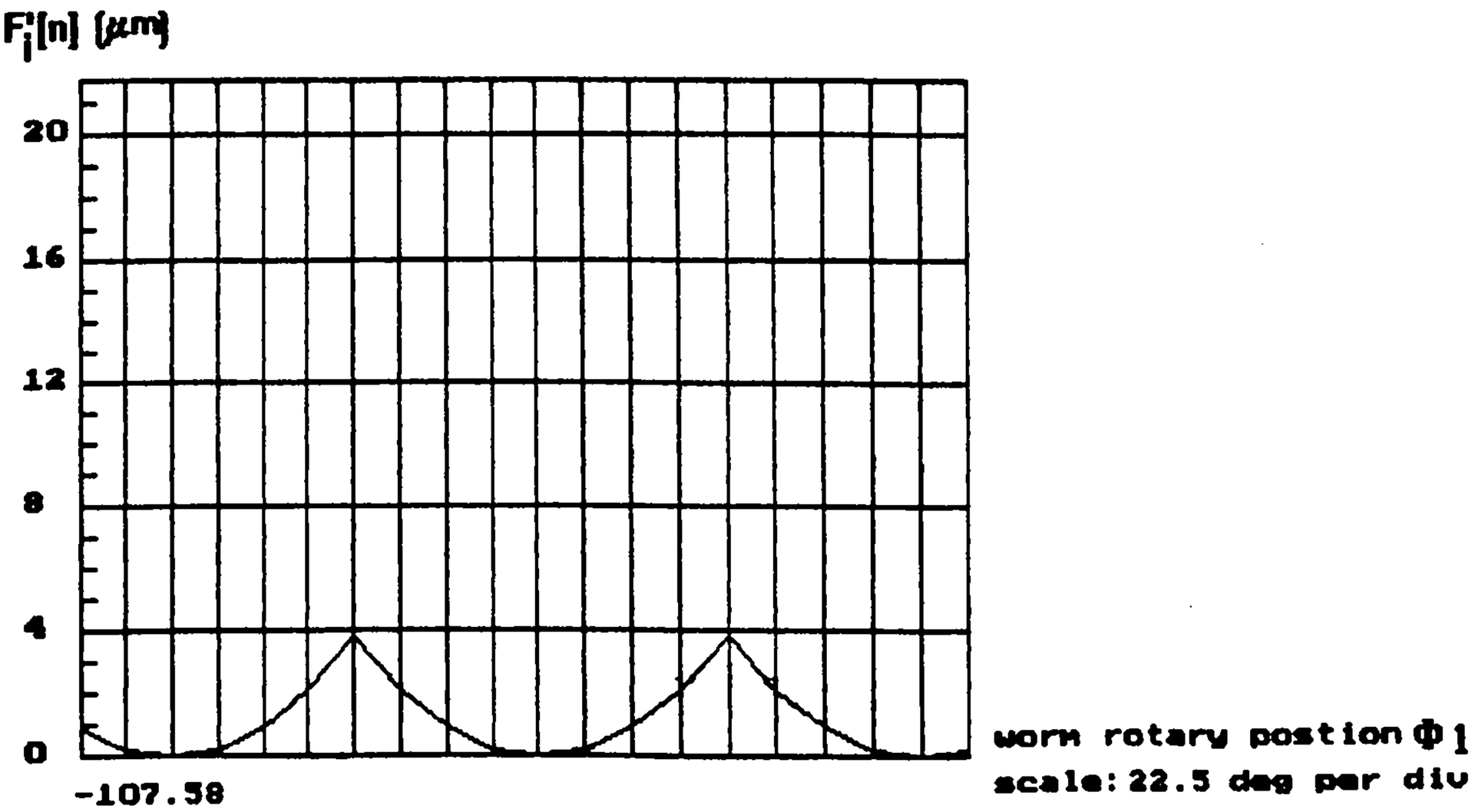
4.5 Gap contour and contact marking pattern

Gaps round a contact point can be displayed by a "no-load gap contour plot" (Figure 4.5 (a)). The no-load gap contour is useful for contact and lubrication analysis. However, the contact marking pattern must be clearly distinguished from the single tooth gap contour plot. A contact marking pattern results from a complete meshing cycle while a gap contour shows the gaps round the contact point at one instant of mesh.

So far we have only discussed the analysis for a single tooth contact. However, on average, more than one single tooth pair must be in mesh if a continuous motion is to be



(a) single tooth transmission error $f_i''[n]$



(b) overlapped transmission error $F_i''[n]$

plus values of $f_i''[n]$ and $F_i''[n]$ denote that the wheel rotated less than it should

Figure 4.4 Overlap of Transmission Errors

transmitted under load, so that there is potential for two or more simultaneously contacting tooth pairs. The transmission error, gap contour and contact marking pattern for each meshing cycle result from the overlapping action of all the engaged teeth, which can potentially engage simultaneously at any instant.

4.5.1 Determination of total number of mating tooth pairs

To determine which wheel teeth are potentially in mesh, we first need to estimate the ranges of worm rotation angle Φ_1 and wheel rotation angle Φ_2 for which contact is possible between one particular tooth pair. This is done during the single tooth pair meshing process, and yields

$$\Phi_1 \in [\Phi_{1min}, \Phi_{1max}]$$
$$\Phi_2 \in [\Phi_{2min}, \Phi_{2max}]$$

which signifies that the rotation angles Φ_1 and Φ_2 belong to closed intervals $[\Phi_{1min}, \Phi_{1max}]$ and $[\Phi_{2min}, \Phi_{2max}]$ respectively. Angles Φ_{1min} or Φ_{2min} and Φ_{1max} or Φ_{2max} correspond respectively to the start and the finish of the single tooth pair mesh. For the no-load analysis, required in this study, the intervals $[\Phi_{1min}, \Phi_{1max}]$ and $[\Phi_{2min}, \Phi_{2max}]$ can easily be determined precisely by the wheel flank generation process, since they normally correspond to the tip and root contact of the wheel tooth flank (see discussion in section 2.6 about boundaries). The interval $[\Phi_{1min} - \Delta\Phi_1, \Phi_{1max} + \Delta\Phi_1]$ is then used for the meshing analysis.

Consider now that the individual worm and wheel teeth are identified by k_1 and k_2, respectively, with $k_1 \in [1,2,\cdots, Z_1]$ and $k_2 \in [1,2,\cdots Z_2]$. Imagine the first single tooth pair to be the datum tooth pair. Without loss of generality, we can define worm tooth $k_1 = 1$ for this worm tooth and $k_2 = 1$ for this wheel tooth. Angles Φ_1 and Φ_2 are assumed to define the angular positions of the worm tooth $k_1 = 1$ and wheel tooth $k_2 = 1$ respectively. The corresponding angular position $\Phi_1(k_1)$ of any worm tooth k_1 is clearly relative to that of worm tooth 1. The corresponding angular position $\Phi_2(k_2)$ of any wheel tooth k_2 is likewise associated with that of wheel tooth 1.

The equivalent "rotation angles" for teeth k_2 on the worm wheel are then

$$\Phi_2(k_2) = \Phi_2 \pm (2\pi / Z_2)(k_2 - 1) \quad \cdots \quad (4.26)$$

From the analysis of a single tooth pair, we find that "proper" conjugate flank contact can only occur (on perfect rigid teeth) between angles Φ_{1min} and Φ_{1max}. However, on 'real' worm gears, (with errors, deflections, hob oversize, etc), contact may occur outside 'nominal' contact region defined by the intervals $[\Phi_{1min}, \Phi_{1max}]$ and $[\Phi_{2min}, \Phi_{2max}]$, so that margins $\Delta\Phi_1$ and $\Delta\Phi_2$ are introduced to allow for this. This gives $\Phi_1 \in [\Phi_{1min} - \Delta\Phi_1, \Phi_{1max} + \Delta\Phi_1]$ and $\Phi_2 \in [\Phi_{2min} - \Delta\Phi_2, \Phi_{2max} + \Delta\Phi_2]$. Although the situation may differ between the start and finish of the mesh cycle, to simplify the calculation, the margins actually used at start and finish of the cycle are all given a value equivalent to the worm rotation increment $\Delta\Phi_1$ ($\Delta\Phi_2 = \Delta\Phi_1 \cdot \Phi_2 / \Phi_1$) to cover the possible round-off errors of the numerical analysis.

The wheel teeth potentially in mesh are those for which

$$\Phi_{2min} - \Delta\Phi_2 \le \Phi_2 \pm (2\pi / Z_2)(k_2 - 1) \le \Phi_{2max} + \Delta\Phi_2 \quad \cdots \quad (4.27)$$

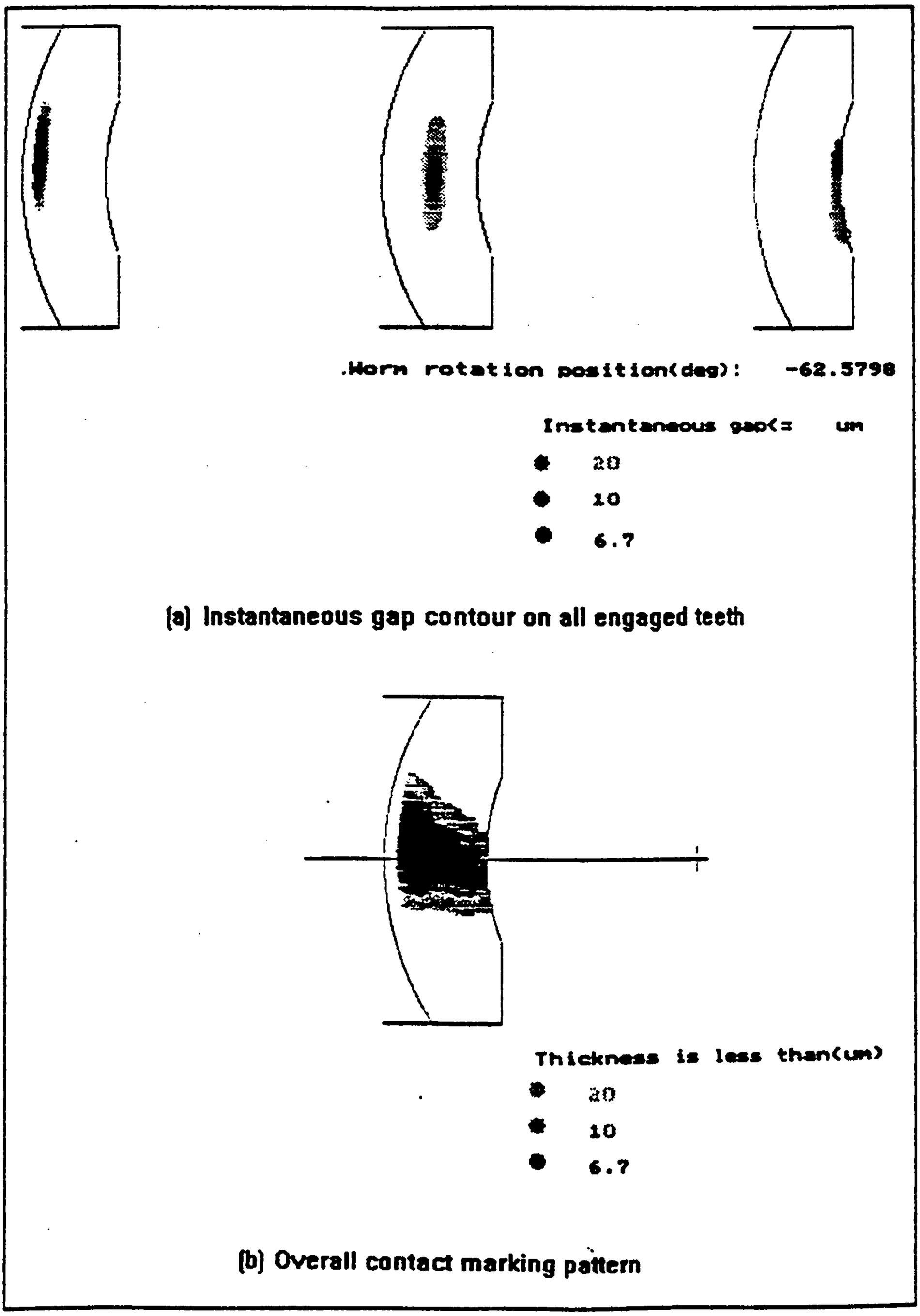
from which the values of k_2 can be found.

4.5.2 Determination of overlapped gap contours

To determine the overlapped gap contour we first need to know the transmission error.

An example of the single tooth pair transmission errors is plotted in Figure 4.4(a). It should be noticed that a positive value on the vertical scale represents a negative transmission error. The procedure required to obtain the overlapped no-load transmission errors shown in Figure 4.4 (b) is:

1) Superimpose the transmission error curves of all mating tooth pairs on the same diagram as in Figure 4.4(a) (the curves for successive tooth pairs are displaced laterally by the angle equivalent to one worm axial pitch);



(a) Instantaneous gap contour on all engaged teeth

(b) Overall contact marking pattern

Figure 4.5 Gap contour and contact pattern

2) Compare the value of f_i' of the curves at each instant (For each step of worm rotation) and take the lowest value as the overlapped transmission error F_i' at this step (since the lowest value corresponds to the maximum plus metal condition which defines the angular position of the driven worm wheel);

3) Plot F_i' to give the overlapped transmission error curve Figure 4.4 (b).

It is worth mentioning that the single tooth transmission error may vary from one tooth pair to another and that pitch errors also affect the overlapped transmission error. A more detailed explanation of this procedure is given by Munro (1990).

The horizontal axis of Figure 4.4 represents the angle of worm rotation Φ_1, whose full range for one single tooth contact is

$$\Phi_1 \in [\Phi_{1min} - \Delta\Phi_1,\ \Phi_{1max} + \Delta\Phi_1]$$

The horizontal scale is $\Delta\Phi_1$ (worm rotation increment) per division. Since $\Delta\Phi_1$ has been chosen as a submultiple of the angle equivalent to one worm axial pitch, as described by equation (4.15), the number of worm rotation steps over a worm axial pitch is:

$$N_{pitch} = 2 \cdot \pi / (Z_1 \cdot \Delta\Phi_1)$$

(In Figure 4.4, $N_{pitch} = 8$).

If the transmission error of the datum tooth pair at position n is represented by

$$f_i'[n] \quad n = 0, 1, 2, \cdots, M$$

where n denotes the n'th step of worm rotation and M is the total number of worm rotation steps, spaced at equal intervals $\Delta\Phi_1$ of worm rotation, required for the datum tooth pair to complete its single tooth mesh cycle.

The overlapped transmission error at the n'th step of worm rotation is denoted by $F_i'[n]$. At each contact point on the single tooth transmission error f_i' [n] curve, the axial gap for

(worm wheel tooth k_2) is zero (provided that tooth k_2 is actually in contact). At other points j, the gap on this tooth pair is $\delta_s[j]$ (from the single pair analysis).

If another worm wheel tooth is in contact, however, the gaps on worm wheel tooth k_2 will increase by the difference between the single tooth pair and overlapped transmission error curves, so that

$$\Delta_s[j] = \delta_s[j] + (f_i'[n] - F_i'[n])$$

where Δ_s is the overlapped gap and j denotes the j'th grid point on the wheel tooth k_2.

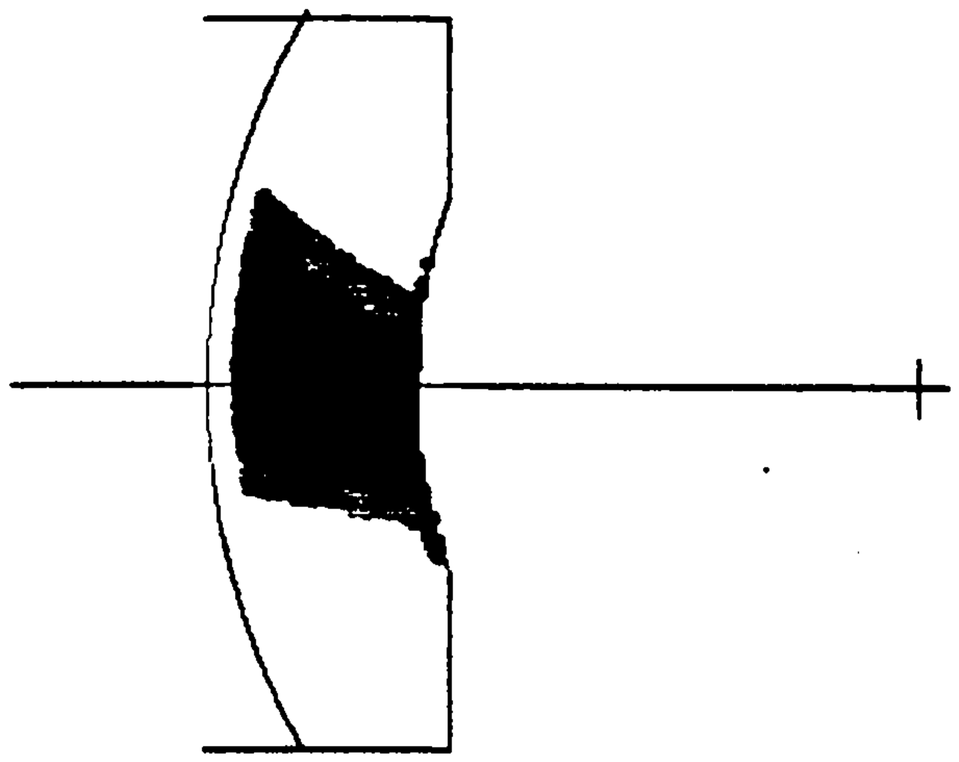
Clearly, if tooth k_2 is in contact ($f_i'[k_2, n] = F_i'[n]$), so that

$$\Delta_s[j] = \delta_s[j]$$

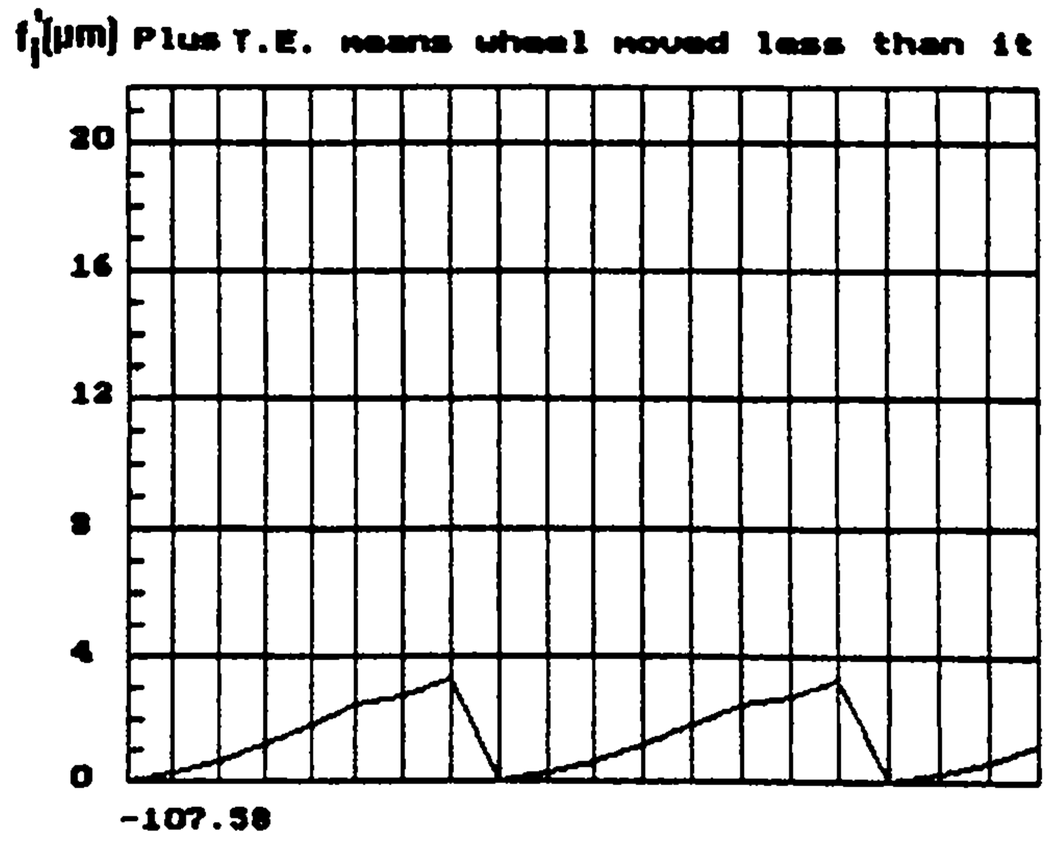
Once the overlapped gaps have been found for the worm rotation step n, plotting of the instantaneous gap contour is only a matter of graphics. An example of plotted instantaneous gap contours is given in Figure 4.5 (a) for the three simultaneously engaged teeth.

4.5.3 Determination of the contact marking pattern

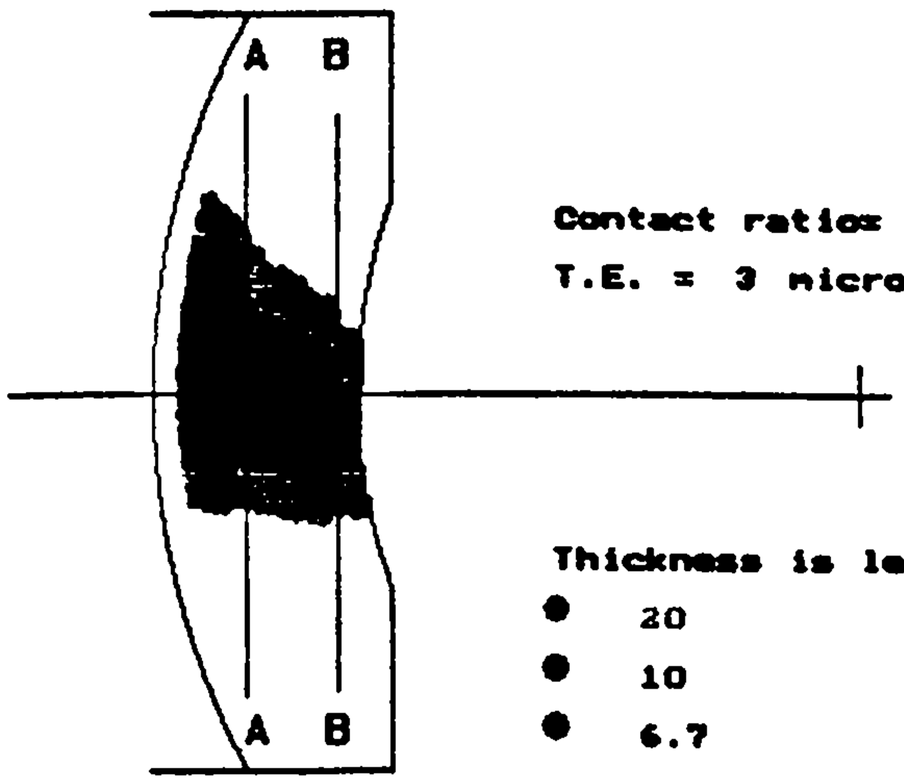
The normal procedure for the industrial contact marking test is described in chapter 3. The contact mark, which moves from the tip of a worm wheel tooth flank to the working depth as the gears rotate, sweeps out a blue area on the worm wheel tooth flank. Thus blue area is the contact pattern. It is not difficult to imagine that a blue mark appears at each meshing instant throughout the region at any point at which the gap size is less than or equal to the thickness of the blue dye painted on the worm flank. Therefore, the prediction of contact pattern is obtained by plotting each individual blue contact marking onto the wheel tooth flank, for every rotation increment n. The marked region actually comprises of all the grid points for which the minimum normal gap is less than the blue dye thickness.



(a) Single tooth contact pattern



(b) Transmission error



(c) Contact pattern

The tooth flank has plus metal at the root; The contact ratio is 2.4; The root dominates the contact; B-B corresponds to a tooth moving out of mesh; A-A corresponds to another tooth moving out of mesh. (Same mesh as Fig. 4.5).

Figure 4.6 Effect of profile error on contact pattern

An example of a contact marking pattern is given in Figure 4.5(b).

4.5.4 The effect of overlap on gap contour and contact pattern

There is a significant difference in the gap contours and contact marking pattern between a single tooth pair and an overlapped mesh with all the engaged tooth pairs. This can be caused by tooth flank form errors, tooth spacing errors, etc. all of which can be easily simulated. An example of the effect of tooth form errors on the contact pattern is given in Figure 4.6. Such variations in initial contact pattern are acceptable in most power drives using bronze worm wheels, since contact will become uniform after a few hours of operation, due to bedding-in wear. A significant variation in the contact marking pattern can also occur from tooth to tooth around the wheel as a result of tooth spacing errors, particularly with even ratios.

4.6 Prediction of entry and exit gaps

The entry (exit) gap discussed in this thesis is the no-load gap between the worm and wheel flanks on the entry (exit) side on a defined wheel cylinder. The initial entry / exit gaps are necessary to reduce manufacturing costs (see section 4.2) and allow initial setting of the centre of contact at the required position. Since the contact changes during bedding-in, the initial contact should be such as will achieve the desired contact under load after bedding-in (refer to section 4.7). Also the entry gap allows for easy entry of lubricant at the worm gear mesh, so that both the entry and exit gaps are important parameters to control. It is, thus, common practice to check the contact parking pattern particularly for the correct entry and exit gaps. However, the no-load contact pattern does not quantify the entry / exit gaps and therefore cannot indicate what the contact conditions will be under load or after bedding-in. This is not a major problem for standard production worm wheels since the required marking can be based on experience already gained. However, it could be of much greater concern for a new design or a special job.

The entry / exit gaps can either be calculated or measured. They are represented in this

thesis by the wheel lead traces whose measurement is discussed in chapter 3. The calculation of lead traces is discussed below.

The procedure for calculating the entry /exit gaps is as follows:

1) Generate a set of wheel flank coordinates with the actual hob and hob settings, as in chapter 2;

2) Generate a second set of wheel flank coordinates conjugate to the worm (using the worm geometry and actual meshing position);

3) Bring the two sets of worm wheel flanks into contact on a selected cylinder (the reference cylinder is the default), by relative rotation about their common axis;

4) Obtain the tangential entry/exit gaps between the two flanks on this cylinder.

The deviations at the two ends of a lead trace are the entry / exit gaps respectively. An example showing calculated and measured entry / exit gaps is given in Figure 4.7.

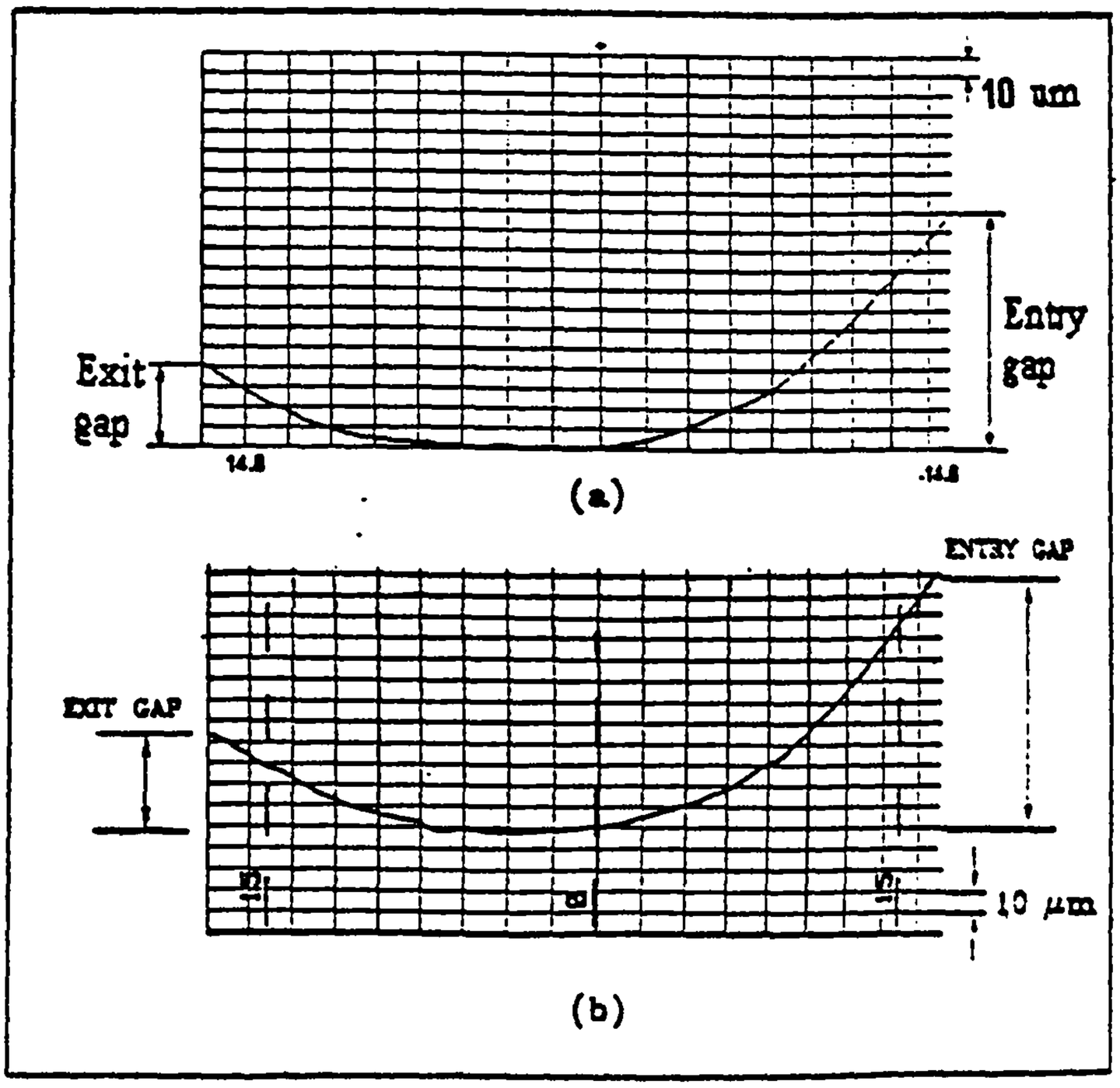


Figure 4.7 Entry / exit gaps (a)calculated (b) measured

The lead trace provides useful guidance for predicting the contact conditions after Obedding-in, since it is most likely that the contact region will extend over the region with the smallest gap. More important for this research project, however, is that, combined with the worm wheel and hob measurement, it assists investigations into the hobbing process. The theoretical lead trace is calculated from the hob geometry and hobbing settings. The difference between the measured lead trace and the predicted lead trace is then mainly caused by hobbing errors (see [Hu and Frazer, 1996] and chapter 7).

One purpose of this study was to allow simulation of a cutting process with measured hob geometry and to study the effect of hobbing settings on the wheel tooth flank (see chapter 5). However, a practical situation is generally much more complicated than this purely kinematic theory suggests and research in this field is still very new compared with that in involute gears, so that a full understanding of how this work can help in the investigation of worm gear hobbing can only be build up gradually.

4.7 Total contact area over all mating teeth

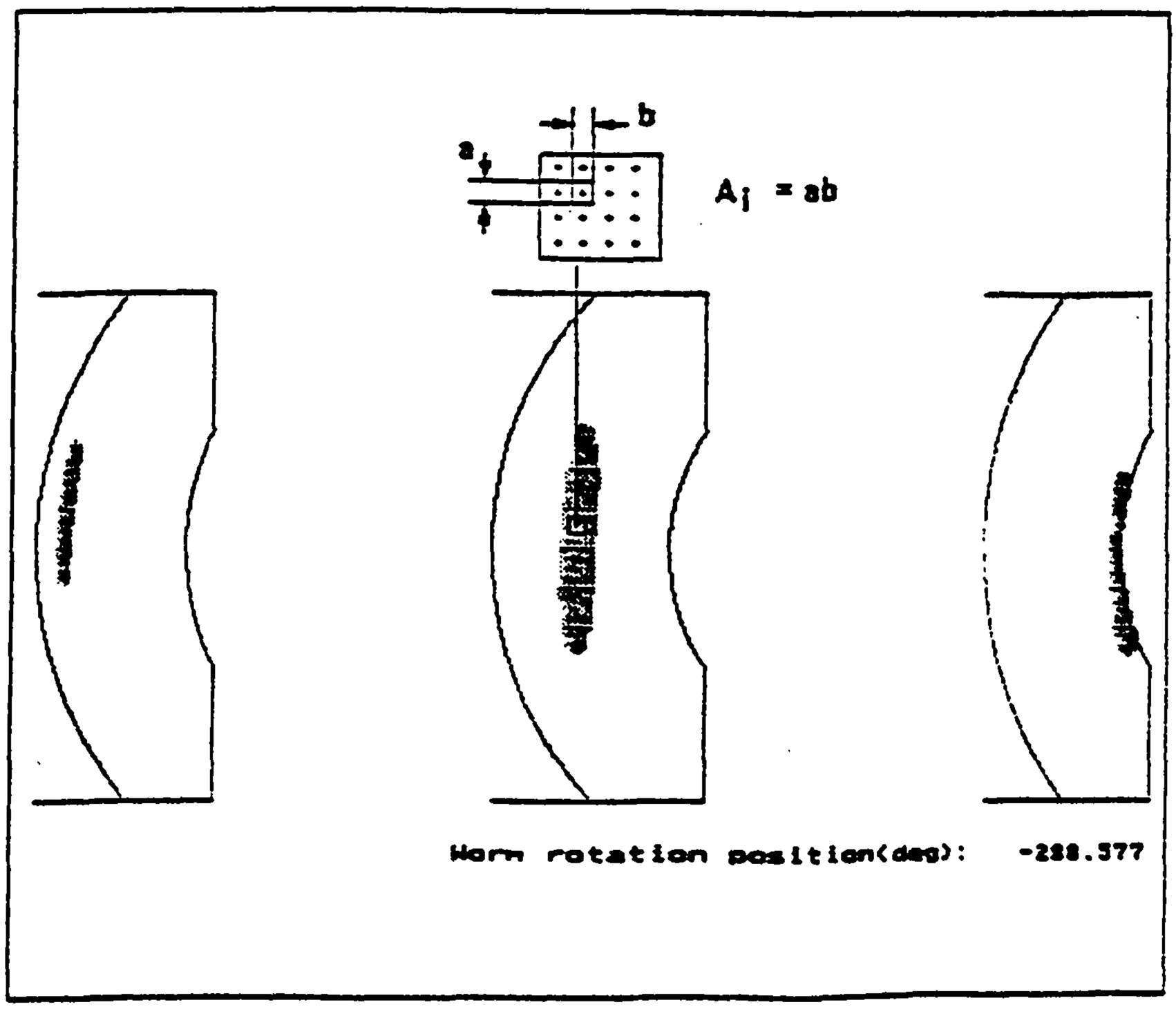


Figure 4.8 Contact Area: Total contact area=$\sum A_i$

Under load, the total transmitted load is shared by all mating tooth pairs, whose mating surfaces will deflect under load. The contact area under load is associated with the degree of conformity of the worm and wheel flanks.

To determine the contact area is beyond the scope of this thesis. The simple alternative approach described here can be used instead to approximately estimate the contact conditions, so that the relative load capacity of different worm design can be compared. In this approach, we assume a constant elastic contact deformation (or wear increment) and so sum the area of all the individual grid squares in which the normal gap magnitudes are less than this constant value.

Since the wheel flank is represented by a regular rectangular grid of points, the sum of all the areas round individual grid points that are "in contact" (ie. with gap less than the assumed deflection) gives an approximation to the total contact area. The total contact area is thus calculated using the following equation:

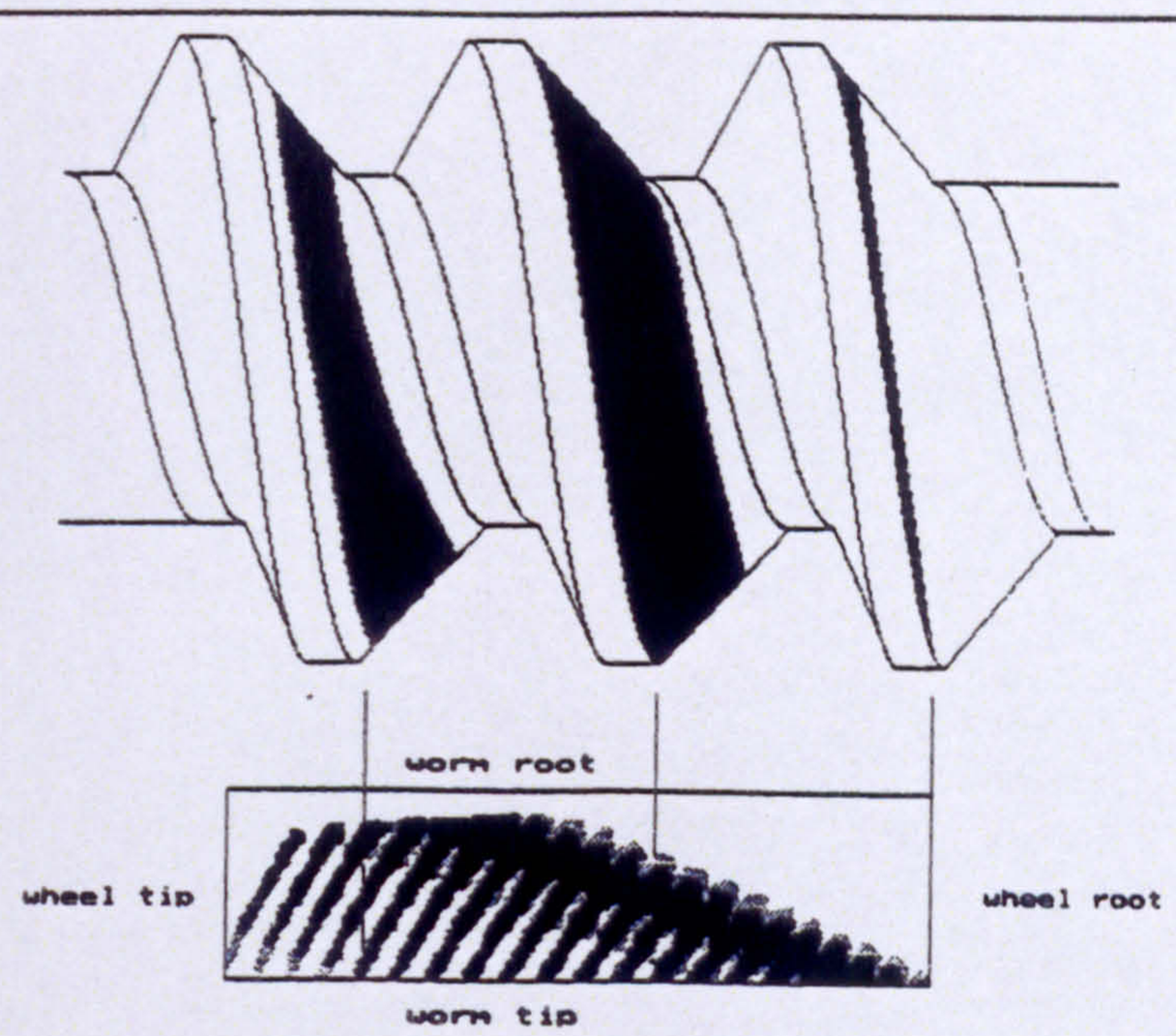
$$A_{Tot} = \sum A_j \qquad j = 0, 1, \ldots, N_c$$

where A_j is the area associated with the j'th grid point (Figure 4.8) and N_c is the total number of grid points that are "in contact". For a regular grid, A_j is constant and equal to the area of one grid square, determined using equations (2.12) and (2.17). For a given load, the total area of contact is clearly related to the contact stresses, since, when comparing two designs which are generally similar (same module, face width, tooth numbers, etc.), the elastic deflection of the teeth under load can also be assumed to be the same for each. The calculation described here is clearly only approximate, but when the transmitted load is given, the contact stress will usually be smaller with a larger total contact area.

4.8 Prediction of wear marking on worm flanks

The worm wear marking depends on the total wear and the combined contact deflection of both the worm and wheel tooth surfaces. The total wear is composed of wear on both

(a)



(b)

Figure 4.9 Worm wear marking (a) predicted (b) actual

worm thread is not symmetrical to the common centre line, its prediction is helpful for worm deflection calculations and determining of the effective length of the worm thread. More importantly, if the worm and wheel flanks have been measured, it also yields information about the contact deflection. With soft worm wheels, the shape of the wear marking on the worm is mainly affected by the wear of the wheel flanks. For hard-to-hard contacts, however, although the total wear would be significantly reduced, it would be more evenly distributed between worm and wheel, so that contact deflections would play a dominant role.

If the contact were a nominal point contact, the wear marking on a worm flank would be a curved line. Because of the contact deflection and wear, however, the marking becomes a band along the worm thread surface. The extent of this contact marking on the worm can be determined from the instantaneous gap magnitude in the same way as for the worm wheel.

An example of wear marking on a worm flank is given in Figure 4.9. The 'striped' appearance of the marking in the lower part of Figure 4.9 (a) is due to the relatively large worm rotation increment (22.5^0) used.

4.9 Validating the non-conjugate analysis against other results

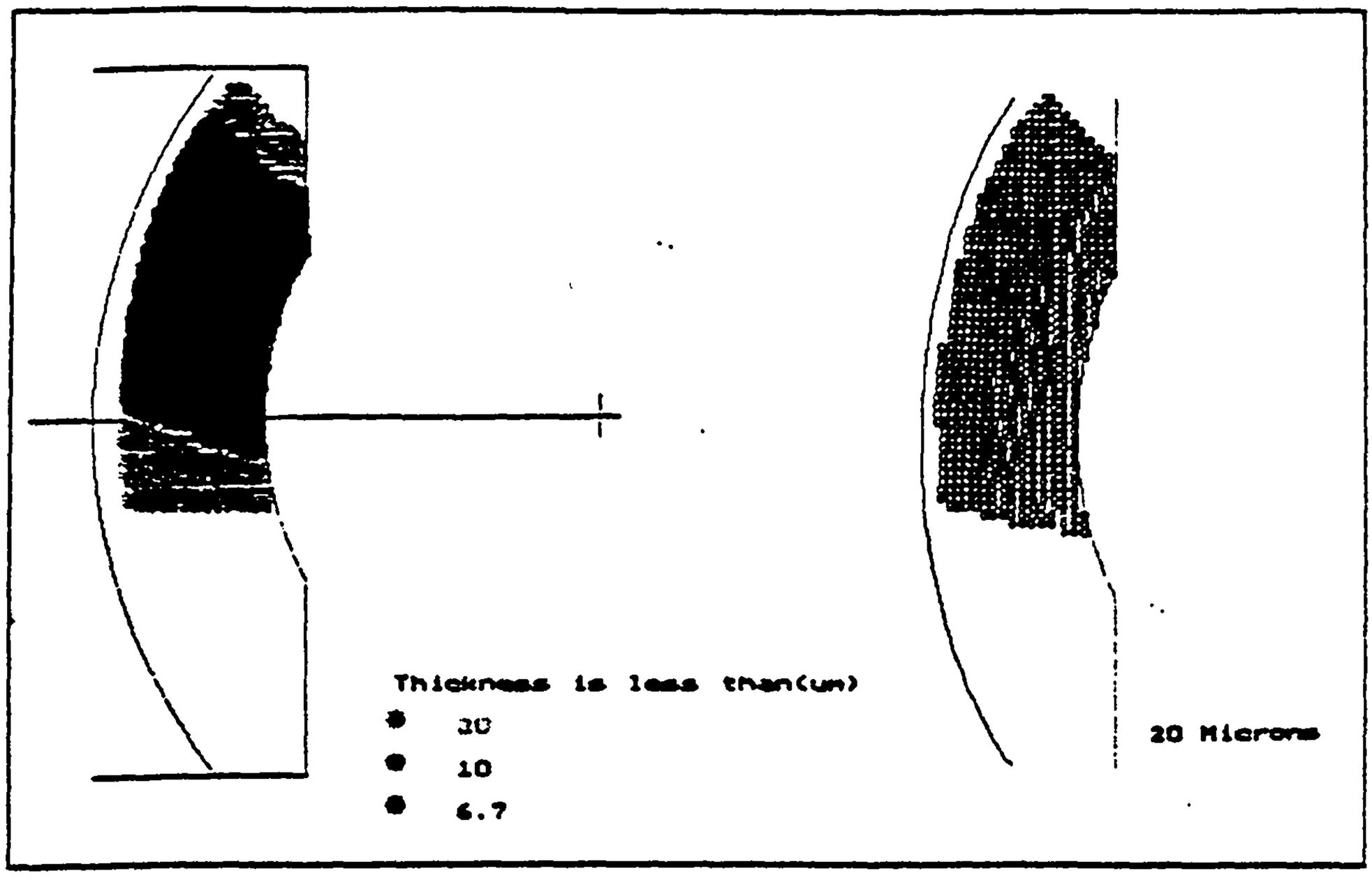
The mesh analysis software has been validated against other calculation methods and against test data from actual worms and wheels with excellent agreement.

4.9.1 Validation against other calculation methods

The contact pattern and transmission error has been validated against results calculated by Fish and Munro(1995). The worm gear data for one test case is given in section A 2.2 (Appendix 2) and the results are given in Figure 4.10.

Additional comparisons with results published by Colboune(1989) also show excellent agreement.

Additional comparisons with results published by Colboune(1989) also show excellent agreement.



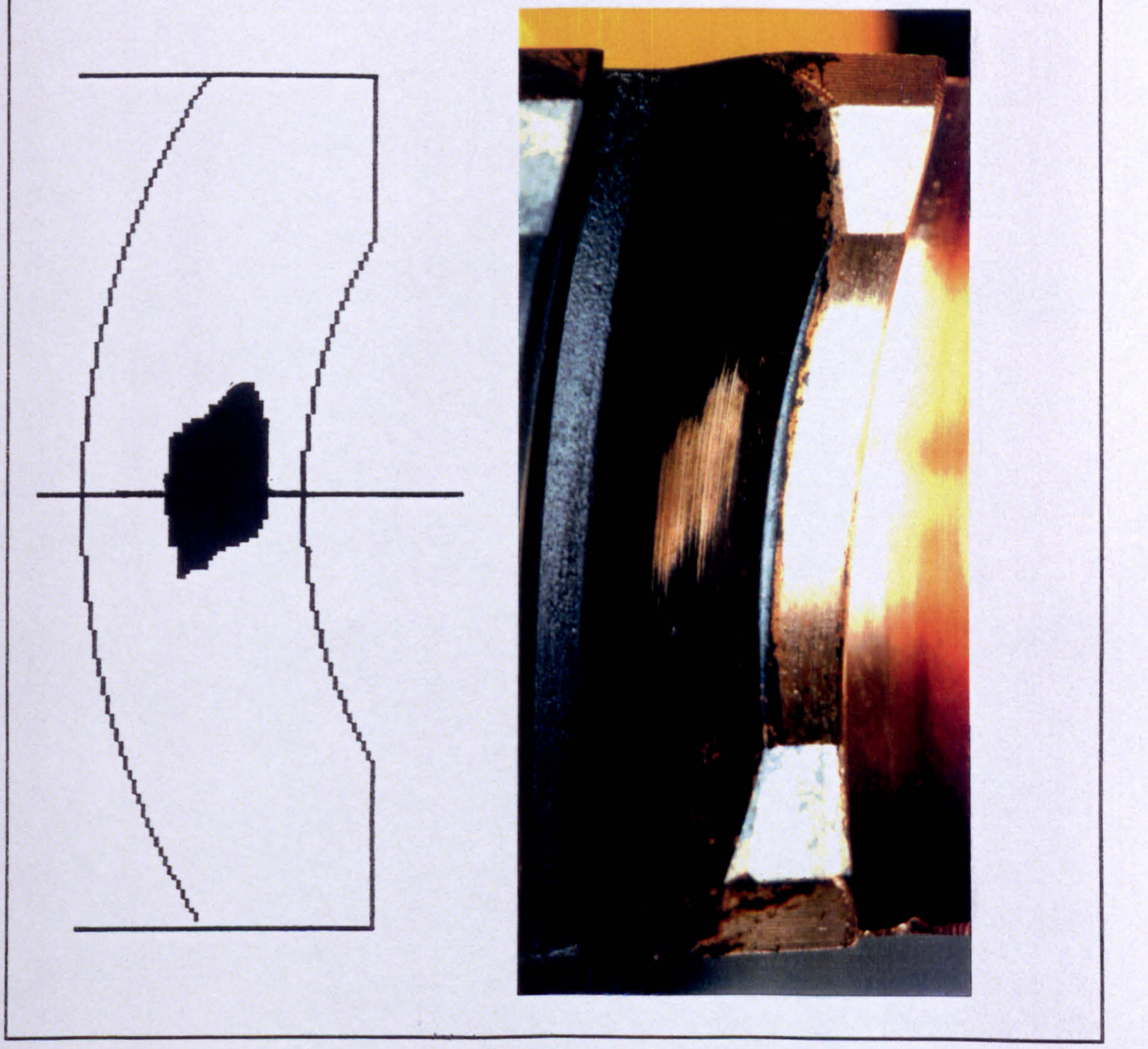
(a) (b)

Figure 4.10 Comparison of results obtained by the author (a) and Fish and Munro(b)

4.9.2 Validation against test results from actual worms and wheels

The software has been used to investigate production problems on many occasions. An example of one investigation is given here which was carried out on site for a lift manufacturer. The worm gear data is given in section A 2.3 (Appendix 2) and the predicted and measured contact marking pattern are compared in Figure 4.11 (see chapter 7 for further comment on this example).

Although the software has been used successfully for predicting contact marking patterns and worm gear performance in several applications, the experience so far gained from these applications show that one of the most important requirements is to obtain reliable input data for prediction. Obtaining reliable data is, in fact, the most difficult part in most applications. In many cases, this is because manufacturers simply do not know (or have not recorded) the precise hob geometry and hobbing settings. To obtain accurate



(a) (b)

(a) Predicted contact pattern

(b) Actual contact pattern

Figure 4.11 Comparison between predicted and actual contact pattern

predictions, the exact cutter profile and machine settings must clearly be known (see chapter 5), although it must be realised that elastic deflection of the cutter can be significant in some cases, leading to inaccuracies in this purely kinematic approach to the analysis of the generation process.

CHAPTER 5: IMPLEMENTATION AND APPLICATION I

THEORETICAL STUDY OF EFFECTS OF HOB FLANK ERRORS AND HOBBING SETTINGS ON THE WORM WHEEL FLANKS

5.1 Introduction to the non-elastic worm gearing analysis software

The non-elastic analysis software has been developed by employing two packages of commercial software - Microsoft Visual Basic 3.0 and Borland C/C++ 3.1.

Visual Basic was used by the author to create the user interface which runs in a standard Windows graphics environment. Using Visual Basic it is convenient to create pull-down menus and sub-menus, along with data input forms and dialogue boxes etc. and all the standard mouse-driven Windows opening, closing, re-sizing, and scrolling features. Visual Basic allows programmers to design the required interface, and when this is complete, it is used to produce a source-code file for the interface. The programmer then needs to modify this source code and insert additional code in order for items selected from the menus to actually run applications.

Borland C/C++ is a middle level computer language and has almost infinite flexibility. It can be used to create applications for both DOS and Windows. It is widely used for real-time applications which permits C/C++ programs to be adopted directly to measurement and manufacturing systems. Borland C/C++ was used by the author to write the programs for computation and analysis.

Visual Basic and Borland C/C++ can be easily linked together. One method is using Visual Basic's **Shell()** function to run executable programs made by Borland C/C++ (or other computer languages' compilers). Another method is using Borland C/C++ to

produce DLLs (dynamic-link library) for Windows. Visual Basic then can Declare references to the external procedures in DLL, including real-time DLL, and Call these procedures just as its own procedures or functions. All types of data can be passed as value or reference between Visual Basic and Borland C/C++ programs with no barriers.

The programs have been written with the following considerations:

1) Combination of GUI (Graphics User Interface) and powerful information flow analysis procedures modules. As discussed above, this results from using the advantages of Visual Basic and Borland C/C++.

2) Expandability and manageability - To meet the requirements of implementation of the general kinematic models, the programs are intellectually managed by modules and user defined function libraries, which allows further features to be added without in any way disturbing the integrity of the programs' ability to perform their existing functions.

3) Maintainability and understandability - Internal documentation has been written throughout the source code and meaningful labels are used for user defined functions and data / variables to guarantee that the programs are readable and maintainable.

4) Accuracy - Single precision floating-point data is used for the analysis procedures, so the accuracy is up to 7 digits, which meets the practical requirements for the analysis, bearing in mind that worm wheel dimensions are of order of 100 mm and manufacturing accuracy cannot presently achieve tolerances of ± 1 μm.

5) Reliability and error tolerance - The programs have been written with the ability to deal with input and run-time errors. Code for error testing, handling and messages has been written wherever necessary. More importantly, the programs carry out the double check of validating results against physical sense. If the results are doubtful, an alternative iterative procedure is employed to try again.

6) Flexibility and user friendliness - The programs have been written bearing users in mind. Default input and flexible and comprehensive options are provided throughout

with online text and graphics instructions. The output has been designed to be informative and easy to interpret.

The logical structure of the non-elastic analysis software is given in Figure 5.1. Details of the user interface are given in Figures 5.2 to 5.10. Table 5.1 shows the principal routes for operating the non-elastic analysis software.

The steps shown in Table 5.1 do not have to be followed successively, since the procedures are controlled by "events" instead of a sequential flow. For example, if a standard worm gear pair (BS 721) is analysed, then the user only needs to input the designation, centre distance and hob oversize: the software does all the rest. If a vital input error has been found by the software, the CALCULATE command will not be available until the error is corrected. The software reports the input errors found, which is very necessary considering that some dozens of input values may be needed to define and analyse a user-defined worm gear pair. Before doing any analysis, the software first checks whether any previously calculated results already exist in the hard disk. If result files for the worm gear to be analysed are found, the software will ask the user to decide between displaying the existing results or recalculating.

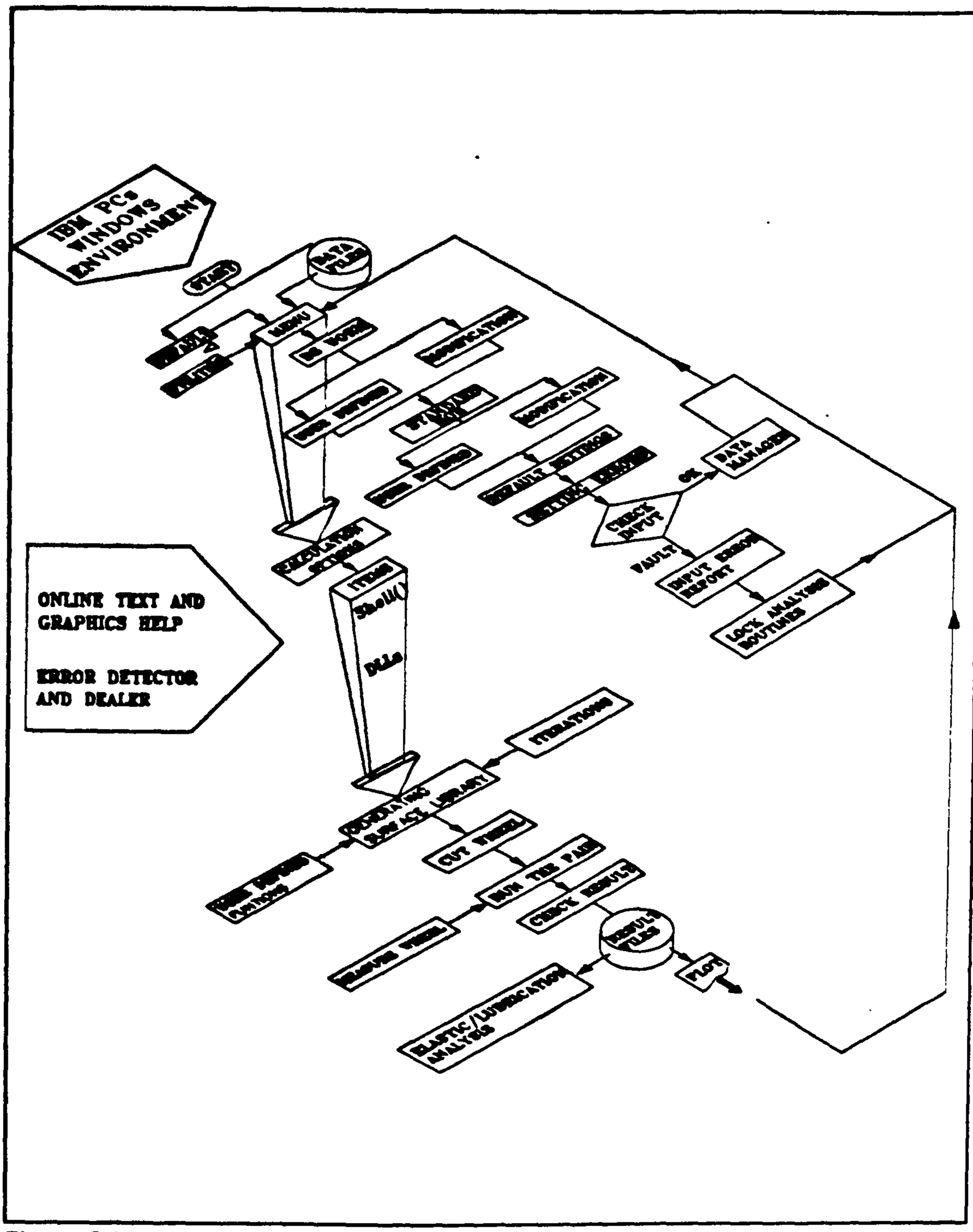
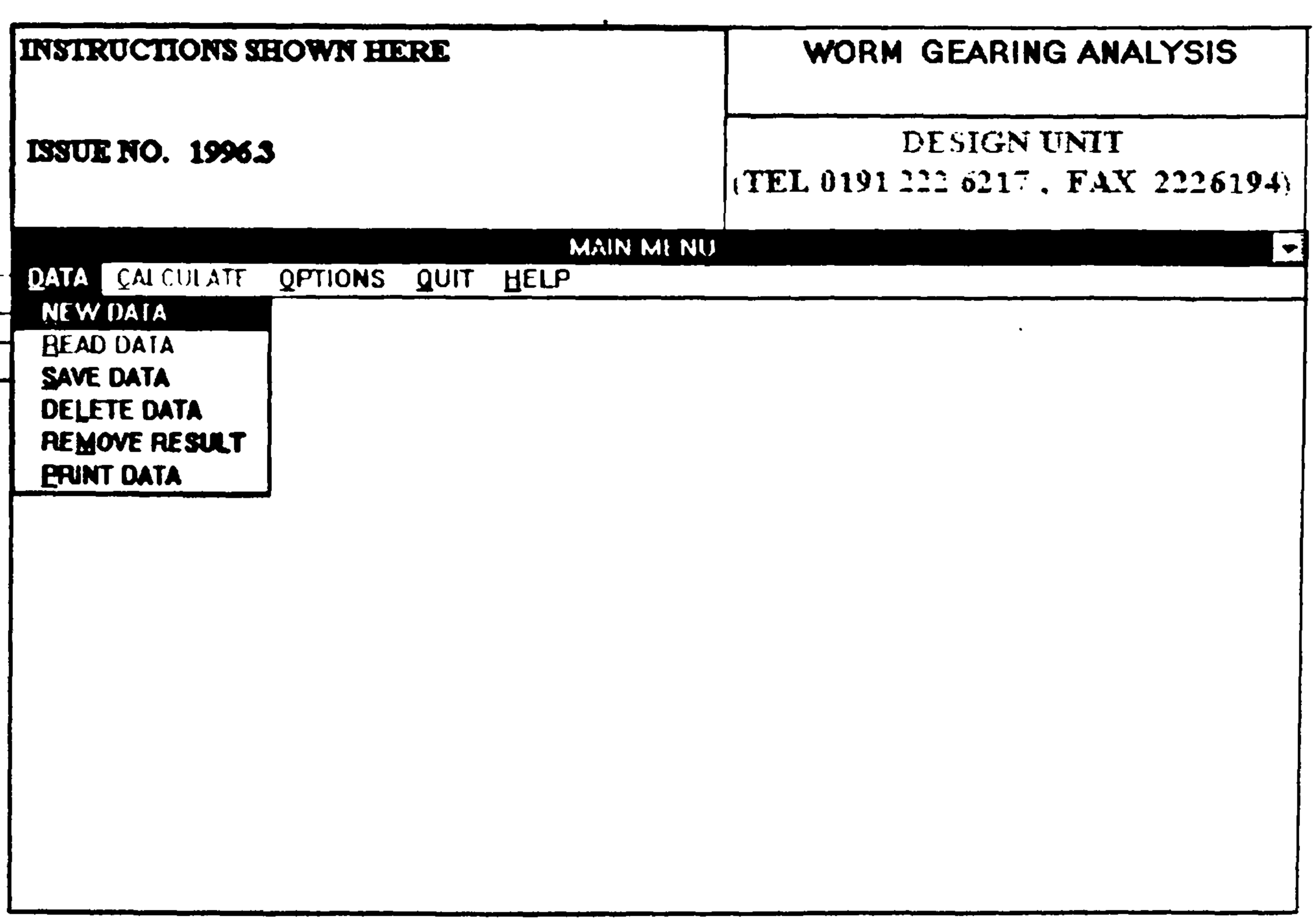
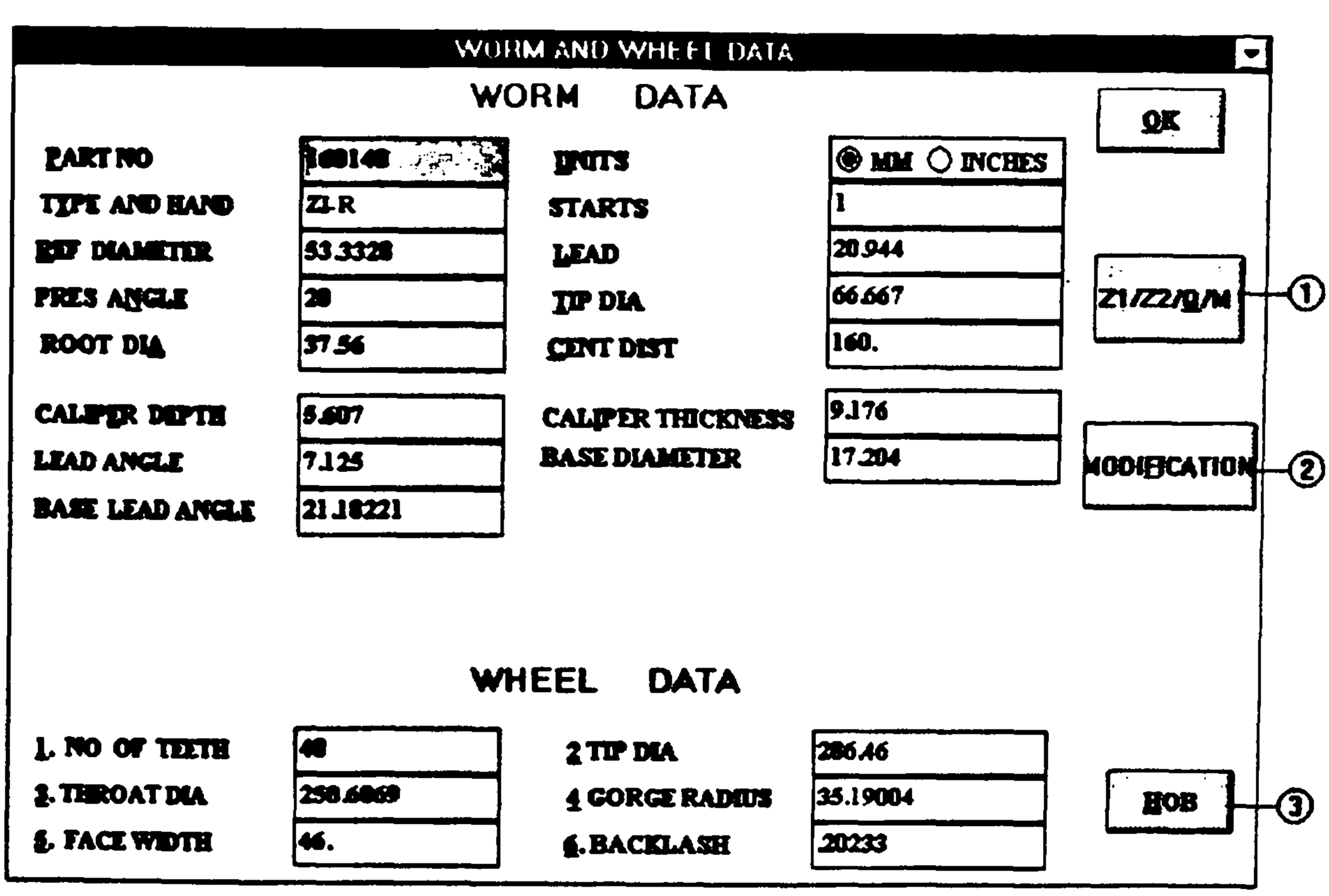


Figure 5.1 Logical structure of the worm gearing analysis software package



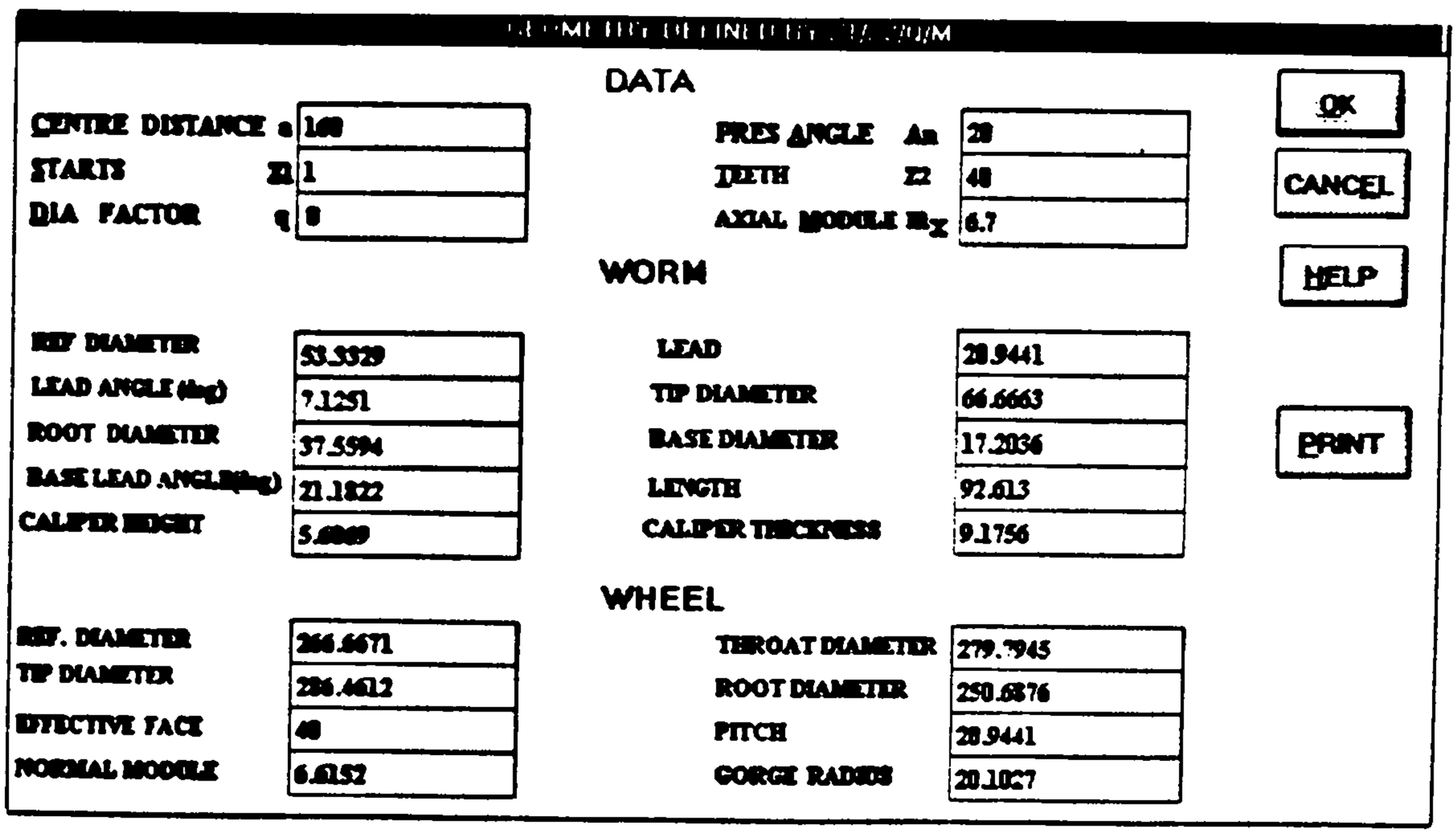
Main menu of the worm gearing analysis software

Figure 5.2



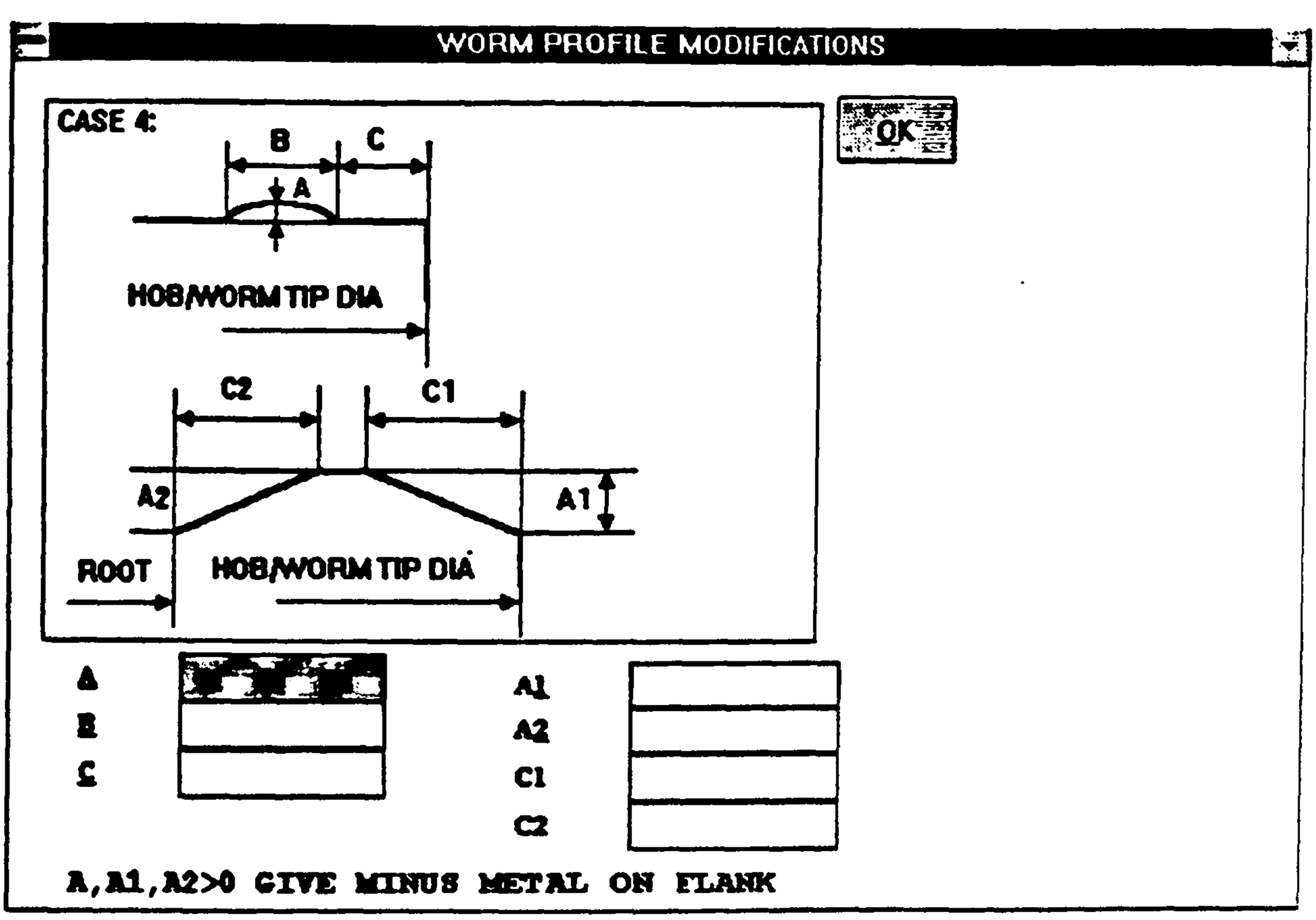
Worm and wheel data input of the worm gearing analysis software

Figure 5.3



BS721 Worm and wheel geometry

Figure 5.4



Worm profile modification input

Figure 5.5

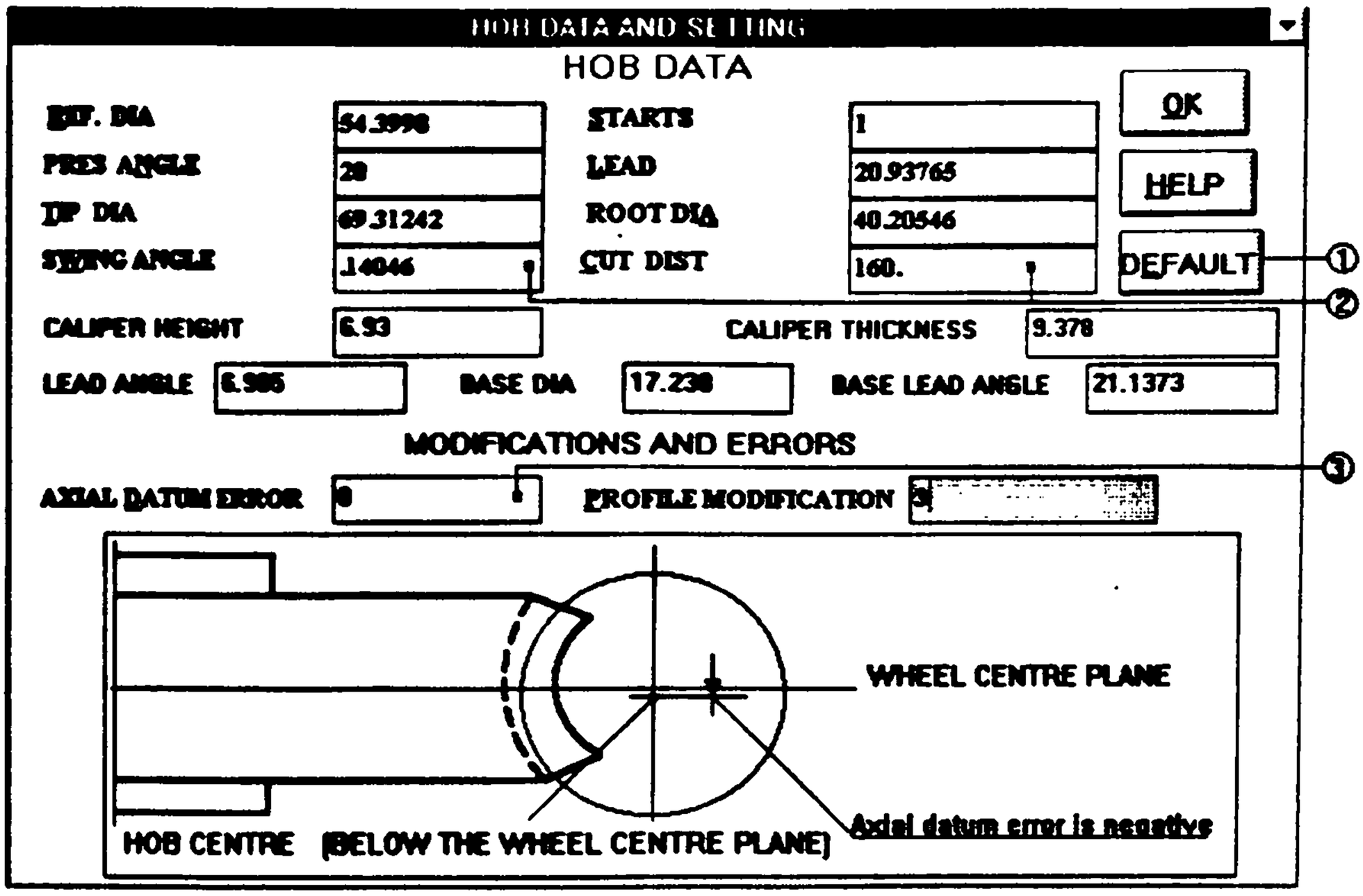


Figure 5.6 Hob data and settings

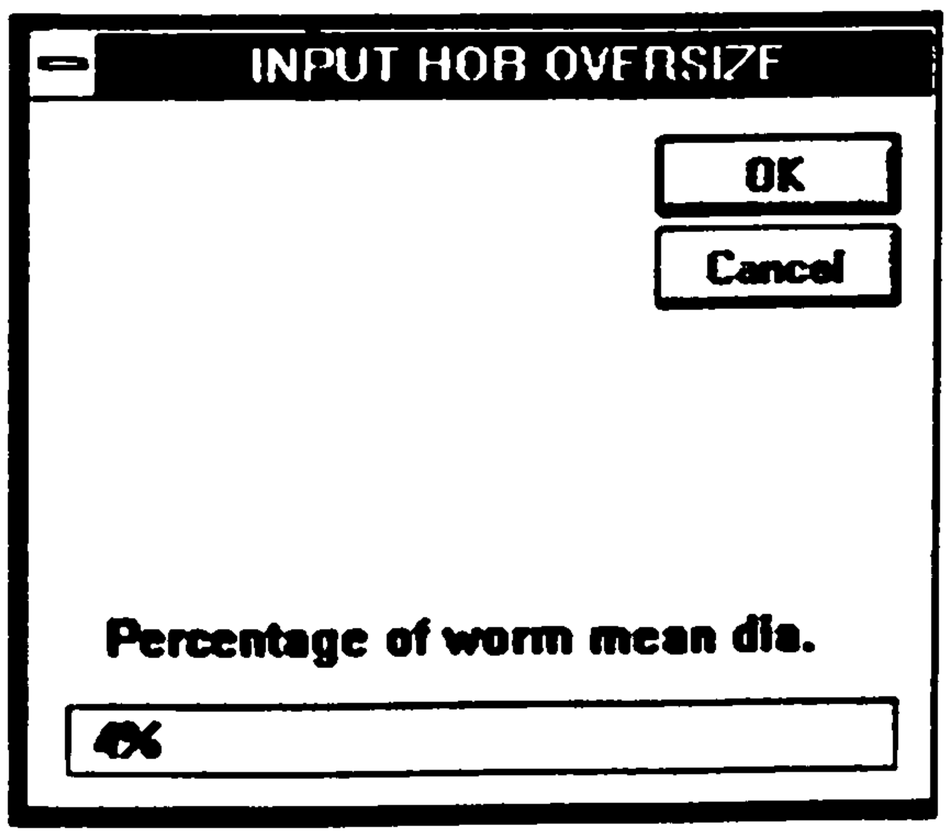
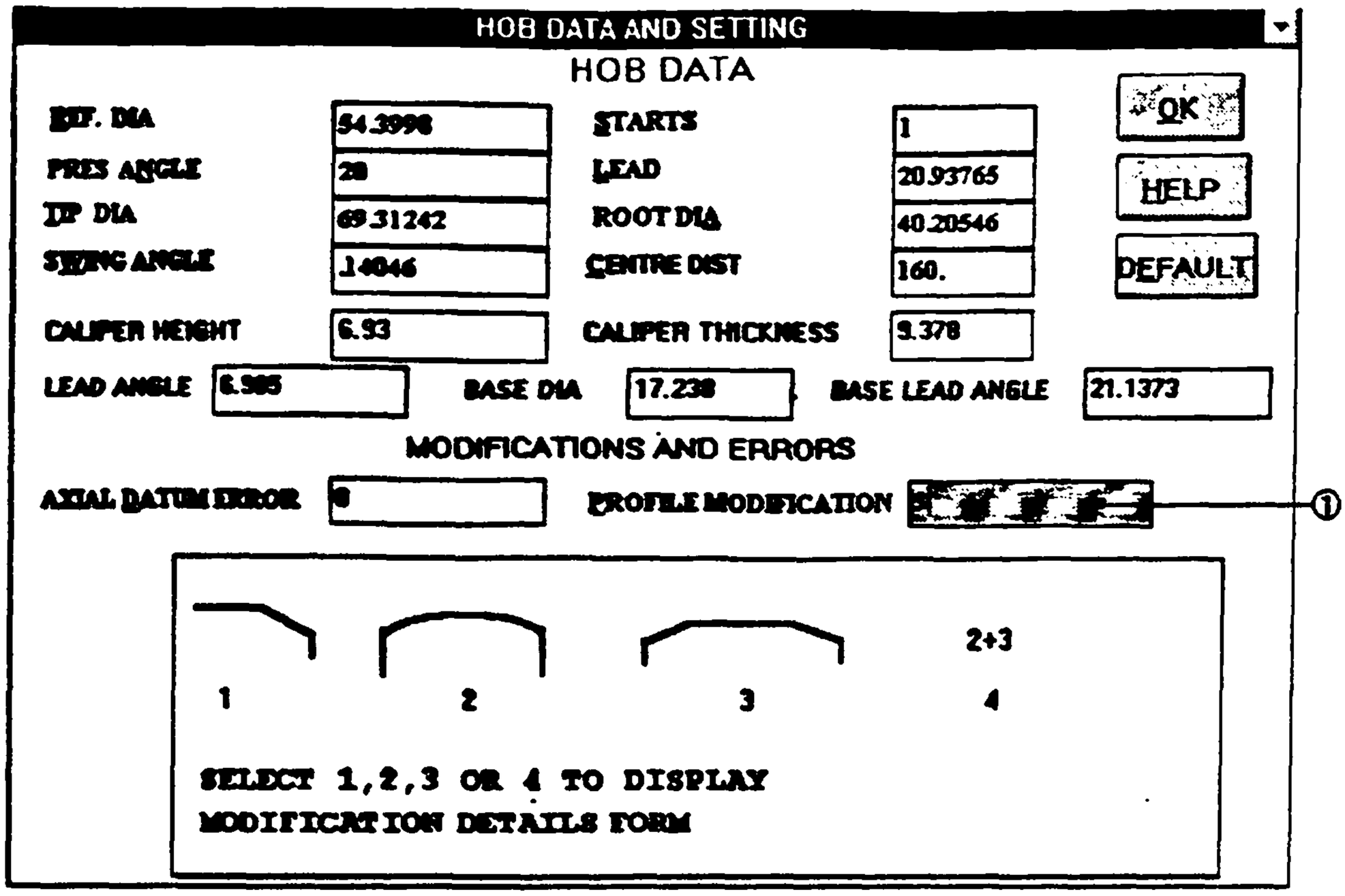
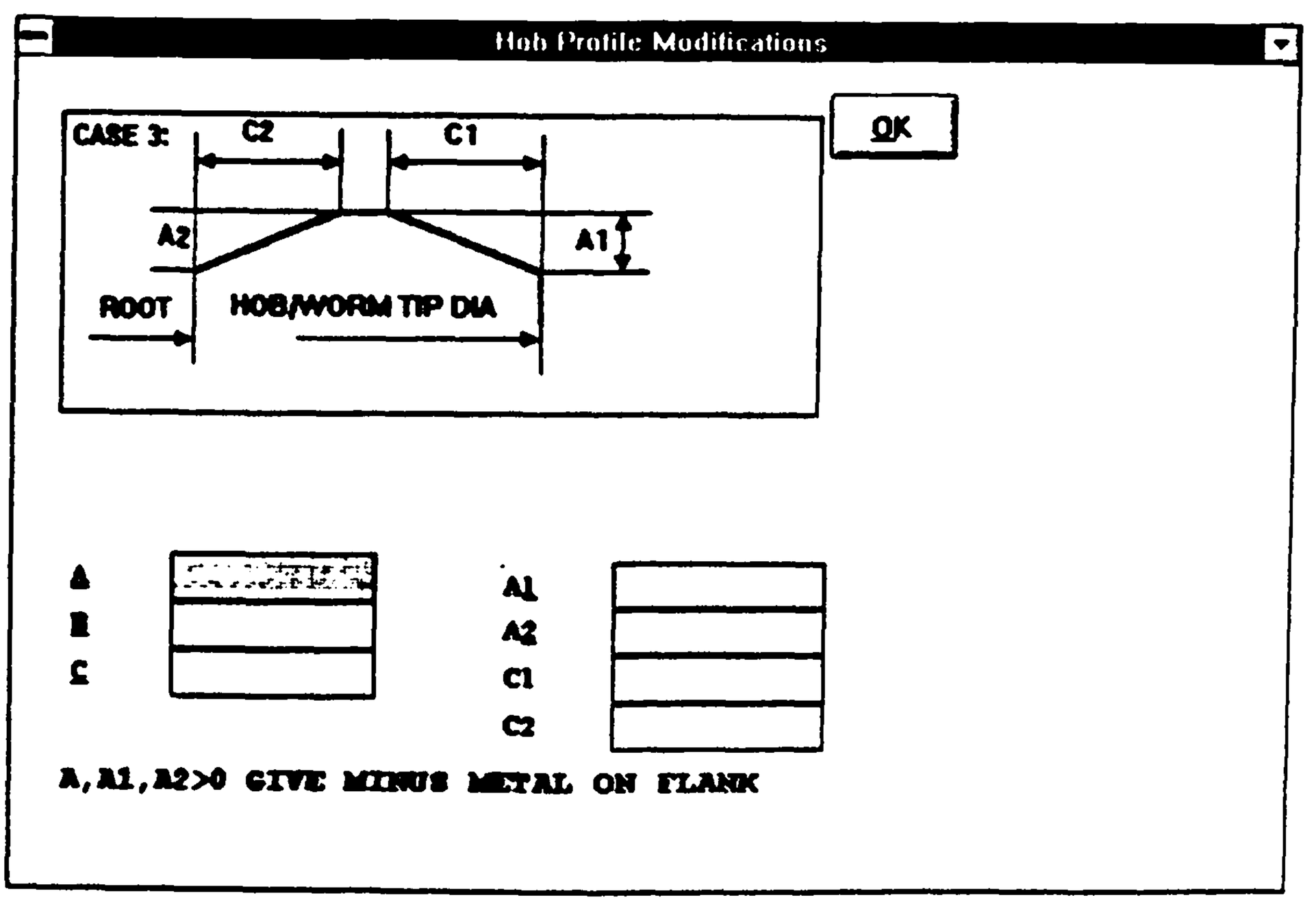


Figure 5.7 Hob oversize input dialogue box

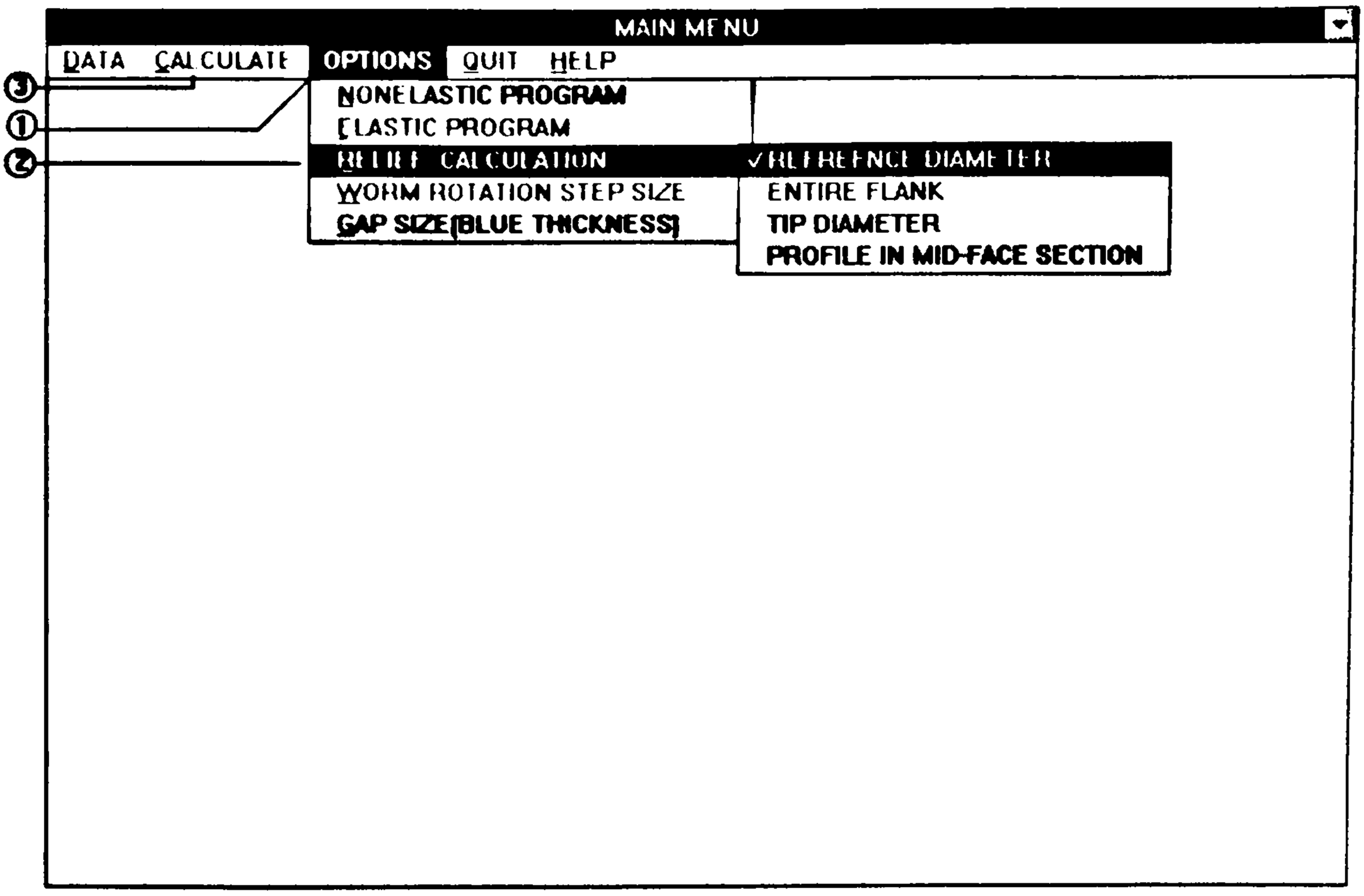


(a)



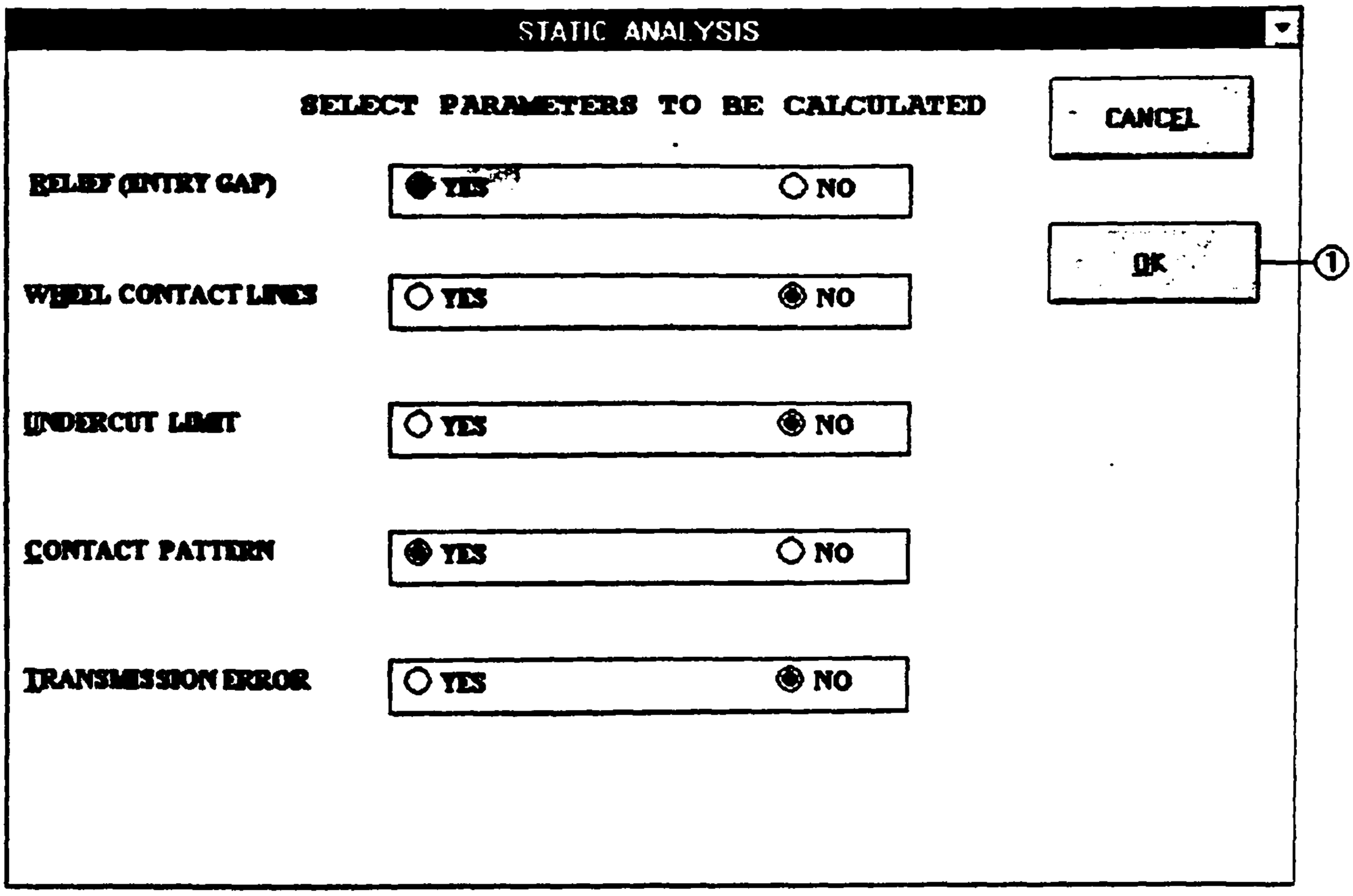
(b)

Figure 5.8 Hob modification input



Analysis options

Figure 5.9



Calculation parameters

Figure 5.10

Step	Task	Command and input form	Refer to
1	Data management (input new data or load data from existing file)	1) DATA command on the main menu; 2) READ DATA command on pull-down menu displays all existing files; 3) NEW DATA command on pull down menu shows the worm and wheel input form.	①,②,③ in Fig. 5.2
2	Design of standard (BS 721) worm gears	1) Z1/Z2/Q/M command on worm and wheel input form; 2) BS 721 input form.	① in Fig. 5.3 and Fig. 5.4
3	Design of user-defined worm gears	Worm and wheel input form.	Fig. 5.3
4	Modifications of worm flank	1) MODIFICATION command on worm and wheel input form; 2) Worm mods text/graphics input form.	② in Fig. 5.3 and Fig. 5.5
5	Display hob input form	HOB command on worm and wheel input form	③ in Fig. 5.3
6	Design of standard hob based on oversize	1) DEFAULT command on hob input form; 2) Oversize input dialogue box.	① in Fig. 5.6 and Fig. 5.7
7	Design of user-defined hob	Hob input form	Fig. 5.6

Table 5.1

8	Modifications of hob flank	1) PROFILE MODIFICATION command on Hob input form (chose one of the types shown graphically to give the appropriate hob modification form); 2) Hob mods text/graphics input form.	① in Fig. 5.8 (a) and Fig. 5.8 (b)
9	Default hob settings	Auto set by software	② in Fig. 5.6
10	Actual hob settings	Change auto set values	② in Fig. 5.6
11	Check input and report input errors	Auto check by software when leaving input forms	④ in Fig.5.2
12	Data management (data file name is given by software as the Part No.)	1) DATA command on main menu; 2) SAVE command on the pull down menu.	① and ④ in Fig. 5.2
13	Chose calculation options	1) OPTIONS command on main menu; 2) pull down and sub menu commands	① and ② in Fig. 5.9
14	Select items (parameters) to calculate	1) CALCULATE command on main menu shows the calculation form; 2) Items on Static analysis selection form.	③ in Fig. 5.9 and Fig. 5.10
15	Do analysis	OK on the selection form of static analysis parameter.	① in Fig. 5.10

Table 5.1 (continued)

16	Save calculation results	Auto save by software (result files are organized by combining Part No. and the calculated item).	
17	Presentation of current results or previous results	Option is given by the software to display previous results detected or do analysis again.	

Table 5.1 (continued)

5.2 Basis of investigation

5.2.1 Objective of the theoretical study in this chapter

Extensive experience has already been gained over many years in finding the sources of errors in spur and helical gears. In contrast, due to the complexity of worm gear geometry, worm gear manufacturers still rely, as explained in chapter 3, on the results of simple contact marking tests and the effectiveness of the initial bedding-in process. Since it was impossible to identify individual manufacturing errors before the appearance of CNC measuring machines, these limitations have played an important role in restricting the accuracy achievable in worm wheel manufacture and dictating the use of soft materials for worm wheels.

However, the possibility has now opened up for producing hard worm wheels accurately with optimised geometry instead of relying on the random bedding-in results obtained with soft wheels, and current research [Hofmann, 1993][Snidle and Evans, 1997] aims to optimise worm gear geometry in this way to improve its performance. Investigation into the individual geometrical error sources of soft worm wheels has nevertheless also been proven very helpful (chapter 7), since it would clearly be a good practice to get the worm wheel made accurately in the first place when high performance is wanted, even though the wheel flank will later be regenerated by its mating worm. These demands

for producing accurate worm wheels make it necessary to understand the effects of hob errors and hob setting errors on the accuracy of worm wheel flanks.

The objectives of this study are thus as follows:

1) Investigation into the effects of the hob flank errors (or modifications) and hob settings on the worm wheel flanks;

2) Proposal of a simple practical method for carrying out these investigations and designing tooth form modifications.

3) Investigation into the causes of typical contact pattern errors;

Efforts have been made to give easy-to-use guidelines giving satisfactory approximations for routine work.

5.2.2 Hobbing error sources

When a worm wheel is hobbed, its tooth form errors will mainly be introduced by the following hobbing errors:

1) Cutting tooth form errors:

The way in which cutting tool profile modifications and/or errors are transferred to the hobbed wheel is discussed in section 5.3.

2) Cutting tool mounting and spindle errors:

Two cutter spindle errors (see Figure 5.14(b)) are significant: those due to eccentricity e_c of the hob, which will introduce a sine-wave error into the hobbing centre distance, and and cyclic axial 'float' of the hob arbor by e_1 (due to, eg. to imperfections in the hob arbor thrust bearings), which will cause a sine-wave axial position error with phase

angle e_{φ_1} relative to axis O_1X_1. These two cutter spindle errors are modelled in section 5.4.

The effect of 'skew' mounting of a fly cutter - eg, mounting it at the wrong angle on its arbor - need not be considered separately, since it is easily treated as part of the hob profile slope error.

3) Machine setting errors:

Three machine setting errors studied here are: hobbing centre distance, swing angle and wheel axial datum errors. These are discussed in sections 5.4, 5.5 and 5.6.

4) Gear blank errors:

These errors will mainly cause gear swash and runout error. They are not discussed in this study, therefore, should be controlled to reduce difficulties when applying the results presented in this chapter. It would be a good practice that the worm wheel datum (reference) surface used for manufacturing is also used for worm drive assembly.

5) Kinematic and other errors in the hobbing machine:

The errors caused by modern CNC hobbing machines are normally negligible, as discussed in the proposal for this research program. But it is still recommended that the swash error of the machine rotary table and the machine transmission error should be checked regularly, otherwise, the sources of worm wheel errors will be very confusing.

5.3 Effect of hob cutting edge profile errors on a wheel flank

An understanding of the mechanism whereby hob flank errors (or modifications) are transferred onto the worm wheel flanks is essential for improving wheel geometry and investigation into worm gear manufacturing accuracy. However, the effect of hob flank errors may vary significantly from one case to another, because a hob can be single-

start or multiple-start with many cutting edges and the transmission ratio can be even or odd. The causes of geometric errors in an involute spur/helical gear are still not fully understood, although its geometry is much simpler than that of a worm wheel.

To simplify the discussion, the concept of 'effective cutting edge' is introduced here. The 'effective cutting edge' is a virtual cutting edge which will generate the same wheel flank as the hob. Some CMMs with hob measurement software inspect errors of the effective cutting edge by measuring all cutting teeth. In this case, the hob measurement results can be used straight away to predict the effect on the wheel flank, with similar results to those obtained in this chapter. Furthermore, study of the effects of errors of a single cutting edge on the wheel flanks is adequate for the worm gear research activity, since, except for the mass production of standard worm drives, fly tools are normally used to cut the worm wheel.

This section will only discuss the effects of errors of a single cutting edge on the wheel flanks, which is the basis of the solution.

In this chapter, the wheel flank is generated by the actual hob cutting edge. This is obtained by adding axial profile deviations (errors measured or modifications designed) to the nominal hob cutting edge profile as represented in form (5.1) and Figure 5.13. The deviations are defined in the hob axial section. This definition is chosen because actual hob axial profiles can be obtained by most hob measuring machines and the generating curve used in this thesis is also defined in the axial direction (see Chapter 4). The shape of the generating curve is given by

$$S_{Act} = S + \Delta S \quad \ldots \quad (5.1)$$

where S and S_{act} are the nominal and actual generating curve respectively, and ΔS is the axial profile deviation.

Once the actual hob cutting edge has been obtained, the actual hob helicoid can be formed with the method described in Chapter 4. The worm wheel can then be

generated by this actual hob helicoid using the method described in Chapter 2. The following points must be noted:

1) When the theory described in chapter 2 is applied to an actual hob helicoid, only the calculation of the actual worm wheel surface is carried out with the actual hob helicoid and the rest (calculation of parameters such as sliding/rolling velocities, avoidance of undercut, etc.) is carried out with the nominal hob helicoid. This is adequate because we are here concerned only to improve the geometric accuracy of worm wheel.

2) Any discontinuities in the actual cutting edge profile must be smoothed down before modelling the actual hob because the generating surface of a worm wheel must be regular (continuous derivative to the 1st order) to apply the conjugate analysis theory.

To simplify the calculation for practical use as a routine work technique, an approximate method is also proposed here. Instead of using the actual hob to generate the wheel, the method uses the nominal hob to generate the wheel flank and then simply transfers the errors of the effective hob cutting edge onto the wheel flank. The following procedure is used:

1) Generation of the wheel flank with the nominal hob profile using the conjugate analysis theory described in Chapter 2.

2) Transfer of the deviations of the effective hob cutting edge onto the generated wheel flank. The transfer scale is approximately 1:1. Plus/minus metal at a particular point on the hob flank produces minus/plus metal at the corresponding point on the wheel flank.

The accuracy of this approximation has been estimated both by numerical analysis and by comparison between hob and wheel measurement results.

The numerical analysis presented here is based on the hob cutting edge profile illustrated in Figure 5.11 in which a profile angle error is simulated. The nominal normal pressure angle of the hob is 22.5°. The actual normal pressure angle is 22.4027° since the measured hob axial profile angle error is 20 μm plus metal at the tip. The actual hob profile is used to generate a wheel flank with the conjugate analysis theory and the results are compared with approximate values obtained from the simplified method proposed above. The maximum difference over the entire wheel flank is less than 3 μm, (less than 1/6 of the hob error) which is acceptable for practical investigation. The calculation of the actual normal pressure angle is given in Table 5.2.

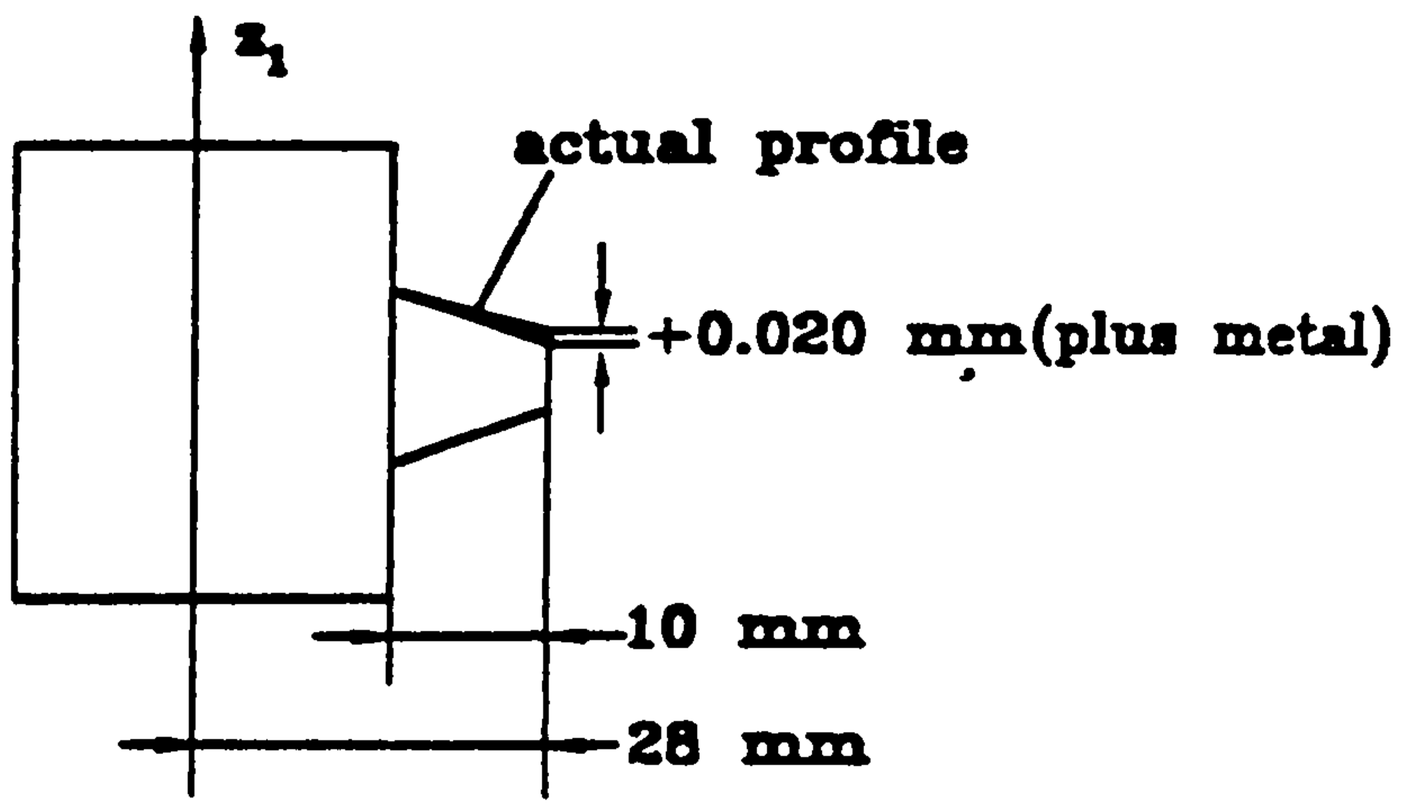


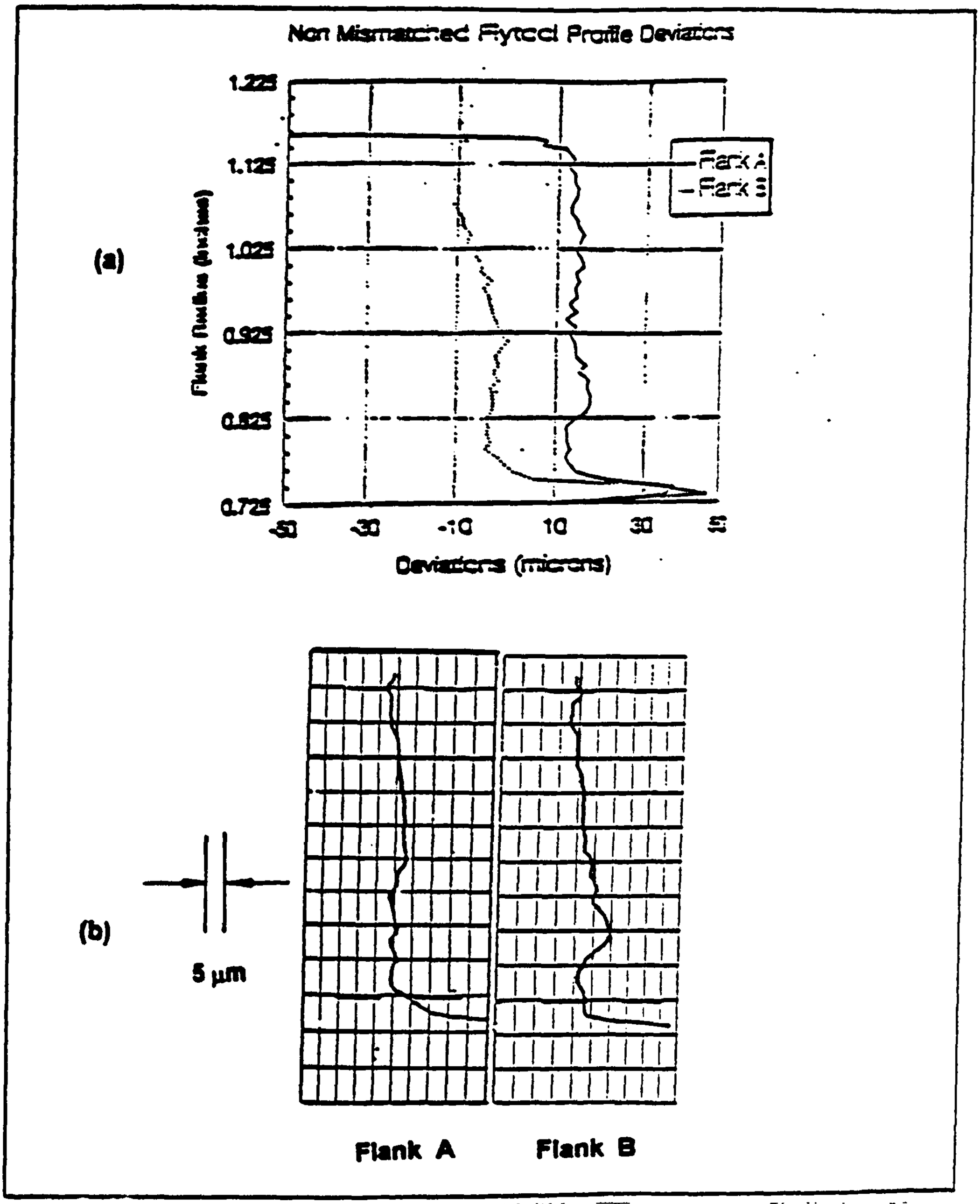
Figure 5.11 Calculation Example of Hob Profile Error

Procedure	Value
Nominal normal pressure angle α_n	22.5°
Lead angle γ	6.25°
Slope of nominal axial profile: $\tan \alpha_x = \tan \alpha_n / \cos \gamma$	0.41669
Axial profile error	+ 0.020 mm
Correction for slope change: 0.02 / 10	0.002
Slope of actual axial profile: $\tan \alpha_x' = \tan \alpha_x - 0.002$	0.41469
Actual normal pressure angle: $\alpha_n = \tan^{-1} (\tan \alpha_x' \cdot \cos \gamma)$	22.4027°

Actual Normal Pressure Angle by Given Axial Profile Slope Error

Table 5.2

The accuracy of the approximation was further checked by comparing hob profile measurement results with the corresponding worm wheel measurement results. An example of these comparisons is given in Figure 5.12.



(a) result of fly tool measurement (b) result of wheel measurement

Figure 5.12 Transfer of fly tool form error onto wheel mid-face profile

The example used a no-mismatch fly tool and the wheel profile was measured in the mid-face section. Figure 5.12 shows that the fly tool cutting edge form error was directly transferred onto the wheel mid-face profile, as predicted. In this test, the fly tool and the worm wheel were measured by Holroyd Machine Tools and the author respectively, and the comparison was carried out by Fish at Huddersfield University. It would, however, be better to use the 'exact' analysis in this case, although the fundamental relationship is applicable.

If measured results are not available, typical worm/ hob profile errors and modifications can be simulated by linear or parabolic deviations or a combination of the two as given in Figure 5.13. These types of profile modification can be used to simulate either designed (intentional) worm or hob profile modifications or measured (unintentional) errors. In practice, tip and/or root relief may be applied on worm and/or wheel flank to reduce the sensitivity to manufacturing errors or to reduce noise level (refer to Chapter 7). A barrelled profile modification or error can be approximately calculated as a parabolic form error, as case 2 of Figure 5.13. The combination of linear and parabolic modification can be calculated with case 4 in Figure 5.13. A local profile form error can also be simulated with case 2 of Figure 5.13 with its depth and location determined by A and C respectively.

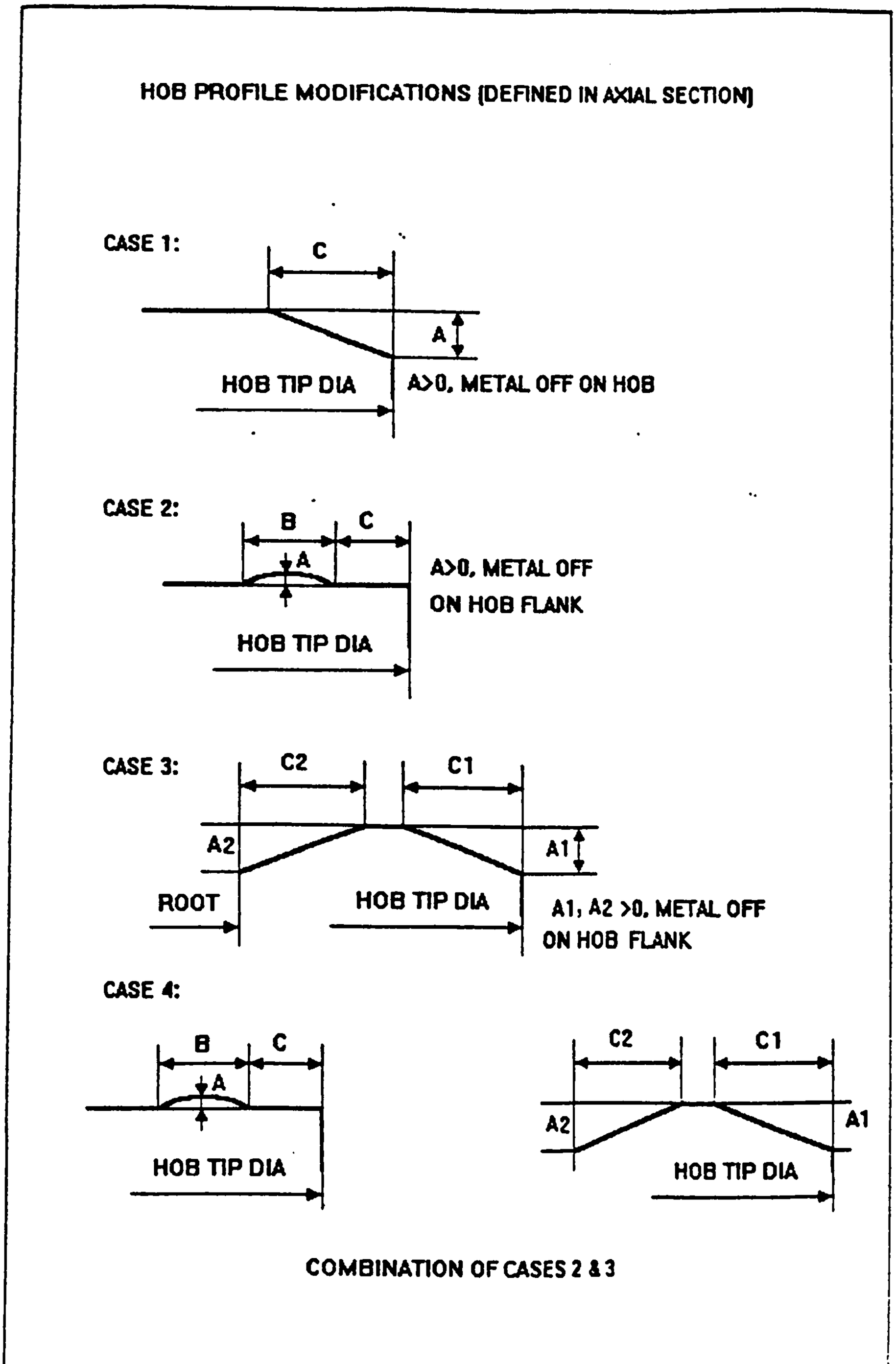


Figure 5.13 Profile modifications used

5.4 Mathematical models of hobbing error

To include typical hobbing setting errors in the analysis, the coordinate systems in Figure 5.14 (a) are set up, based on the coordinate systems in Figure 2.1 (see Chapter 2). The new parameters and features introduced to simulate the hob and wheel setting errors are:

L_1 - the constant axial position error of the hob/worm thread;

L_2 - the constant error of wheel axial datum position (see Figure 5.16);

C - the actual hobbing centre distance with a constant error included;

η - the actual swing angle with a constant error included.

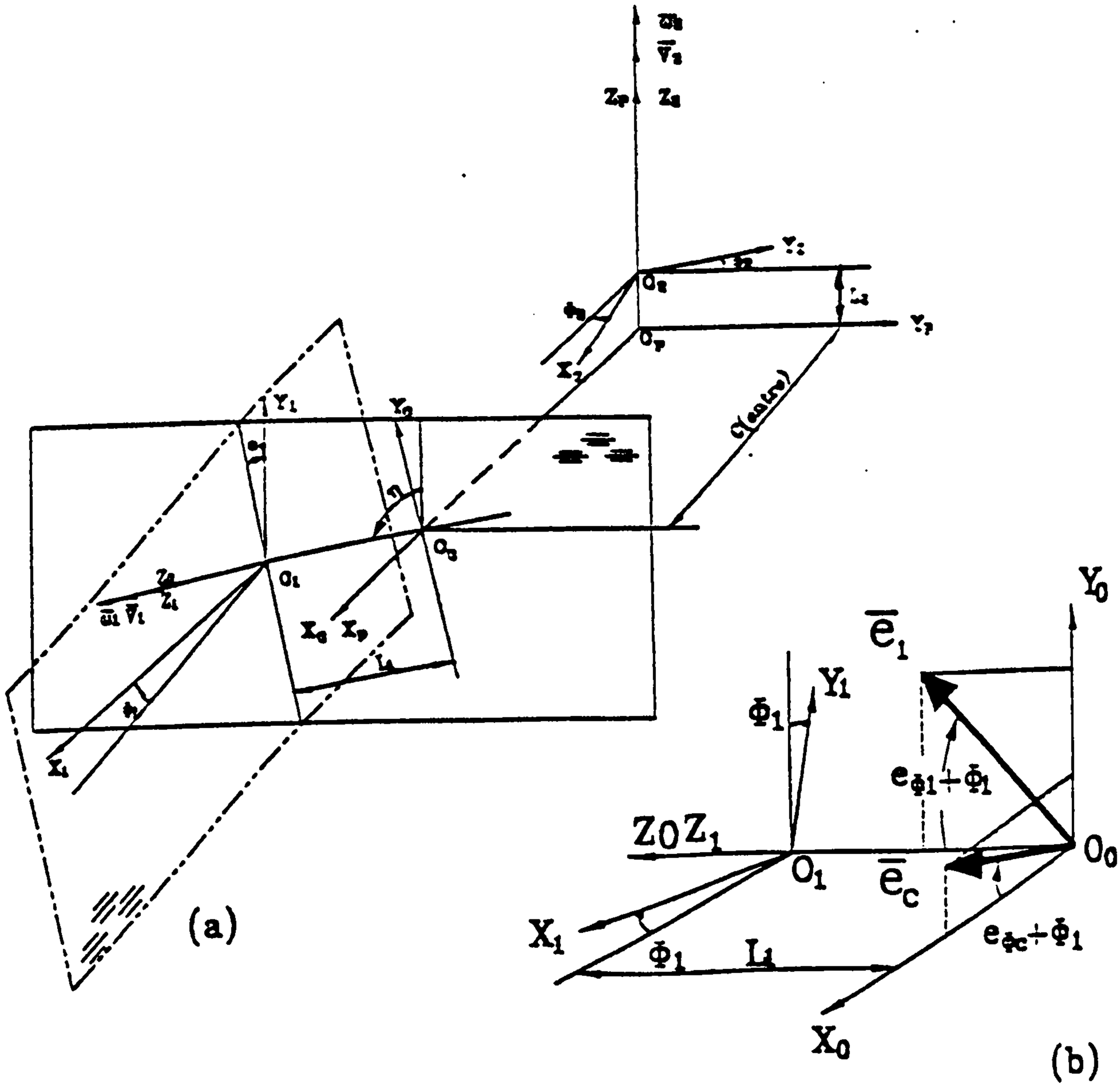


Figure 5.14 Coordinate systems including Hob Setting Errors

The mathematical models presented in chapter 2 are used directly without major changes apart from the errors which are introduced as follows [Zhang and Hu, 1989]:

1) The coordinate transformation matrix from coordinate system S_1 to S_2 has been modified to give

$$T_{21} = \begin{bmatrix} \cos\phi_1\cdot\cos\phi_2+\sin\phi_1\cdot\sin\phi_2\cdot\sin\eta & \cos\phi_2\cdot\sin\phi_1+\sin\phi_1\cdot\cos\phi_1\cdot\sin\eta & -\sin\phi_2\cdot\cos\eta & C\cdot\cos\phi_2-L_1\cdot\sin\phi_2\cdot\sin\eta \\ -\sin\phi_2\cdot\cos\phi_1+\cos\phi_1\cdot\sin\phi_1\cdot\sin\eta & \sin\phi_1\cdot\sin\phi_2+\cos\phi_2\cdot\cos\phi_1\cdot\cos\eta & \cos\phi_2\cdot\cos\eta & -C\cdot\sin\phi_2-L_1\cdot\cos\phi_2\cdot\sin\eta \\ \sin\phi_1\cdot\cos\eta & \cos\phi_1\cdot\cos\eta & \sin\eta & L_1\cdot\cos\eta-L_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \ldots (5.2)$$

2) The B-Matrix has included L_1 and L_2, so that

$$B = \begin{bmatrix} 0 & i\cos\eta-1 & -i\sin\eta\cos\phi_1 & -iC\sin\phi_1\cos\eta-i\cdot(L_1\sin\eta\cos\phi_1-L_2\sin\phi_1\cos\eta\frac{\partial L_1}{\partial\phi_1}) \\ 1-i\cos\eta & 0 & i\sin\eta\sin\phi_1 & -iC\cos\phi_1\cos\eta+i\cdot(L_1\sin\eta\sin\phi_1-L_2\cos\phi_1\sin\eta\frac{\partial L_1}{\partial\phi_1}) \\ i\sin\eta\cos\phi_1 & -i\sin\eta\sin\phi_1 & 0 & iC\sin\eta+L_1\frac{\partial L_1}{\partial\phi_1}-L_2\cos\eta\frac{\partial L_2}{\partial\phi_1} \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \ldots (5.3)$$

3) The total variable hobbing error is given

$$e=e_1+e_c=\begin{bmatrix} e_c\cdot\cos(e_{\phi c}+\phi_1) \\ e_c\cdot\sin(e_{\phi c}+\phi_1) \\ e_1\cdot\cos(e_{\phi 1}+\phi_1) \end{bmatrix} \quad \ldots\ldots \quad (5.4)$$

where:

- e vector of total variable error
- e_1 vector of variable hob axial error
- e_c vector of hob variable concentricity error
- $e_{\phi 1}$ phase angle of the hob axial error vector, measured relative to axis O_1X_1
- $e_{\phi c}$ phase angle of the hob concentricity error vector, measured relative to axis O_1X_1

These errors are all represented in terms of components of coordinate system $S_0(O_0, x_0, y_0, z_0)$, as shown in Figure 5.14 (b).

4) Actual hob settings relative to the wheel are then given by equation (5.5)

$$\begin{bmatrix} C_e \\ L_{2e} \\ L_{1e} \end{bmatrix} = \begin{bmatrix} C + e_c \cdot cos(e_{\phi c} + \phi_1) \\ L_2 + e_c \cdot sin(e_{\phi c} + \phi_1) \\ L_1 + e_1 \cdot cos(e_{\phi 1} + \phi_1) \end{bmatrix} \quad \cdots\cdots \quad (5.5)$$

where:

- L_{1e} actual axial position of the hob with the variable and constant errors
- L_{2e} actual axial position of the wheel with the variable and constant errors
- C_e actual hobbing centre distance with the variable and constant errors

L_1, L_2 and C need to be replaced by L_{1e}, L_{2e} and C_e respectively in formulas (5.3) and (5.4) when the variable errors are included in the analysis software. The variable error of swing angle has second order effect and therefore is not included, bearing in mind that our discussion has been confined to the single 'effective' cutting edge. But the actual swing angle might be set very differently from nominal value and thus the actual swing angle must be used in the analysis.

5.5 Effect of individual hobbing settings on a wheel flank

Once the models are ready for carrying out a parametric study of the effects of machine settings, the question to answer is how individual setting errors and composite setting errors affect the wheel flanks. The following influences have been investigated:

1) How the swing angle η affects the wheel lead traces;
2) How the wheel axial datum (hob setting height) L_2 affects the wheel lead traces;
3) How the hobbing centre distance C affects the wheel lead traces;
4) How composite errors affect the wheel lead traces;
5) How the swing angle η and hobbing centre distance C affect the wheel profile traces;

Figures 5.15 to 5.21 illustrate the terms used for hobbing settings and the corresponding results. The investigation results are also summarized in Table 5.3.

Error	Effect	Approximation
Effect of swing angle η on wheel lead traces	1)The slope of the lead trace is equal to that of the swing angle; 2) The lead trace variation is nearly linear.	$f_{H\beta} \approx b_{eff} \cdot \eta$ $f_{H\beta}$ - lead slope error; b_{eff} - effective face width; η – swing angle.
Effect of axial datum error L_2 (hob setting height) on wheel lead traces	1) The slope error of the lead trace is about half the axial datum error; 2) The lead trace variation is nearly linear.	$f_{H\beta} \approx 0.5 \cdot L_2$ L_2 - the constant error of wheel axial datum position.

Summary of Effects of Hobbing Errors on the Wheel Geometry

Table 5.3

Effect of hobbing centre distance on wheel lead traces	1) The lead trace is nearly a parabolic curve; 2) The high (low) point is at worm base radius.	
Effect of composite setting errors on wheel lead traces	The composite effect can be estimated by adding each individual effects.	$f_{H\beta} \approx \sum f_{H\beta}(i)$ $f_{H\beta}(i)$ - individual lead slope errors.
Effect of swing angle on wheel profile traces	1)The swing angle is critical to mid-face profile slope; 2)The slope error is approximately proportional to the swing angle.	$f_{H\alpha} \propto \eta$ $f_{H\alpha}$ – profile slope error.
Effect of hobbing centre distance on wheel profile traces	The profile angle error is approximately proportional to the hobbing centre distance error.	$f_{H\alpha} \propto \Delta C$ ΔC - hobbing centre distance error.
Effect of axial datum (hob setting height) on wheel profile traces	Insignificant, negligible	$f_{H\alpha} \approx 0$
Effect of composite setting errors on wheel profile traces	1)The composite effect can be estimated by adding each individual effects; 2) Increase of cut distance will reduce the profile error caused by swing angle.	$f_{H\alpha} \approx \sum f_{H\alpha}(i)$ $f_{H\alpha}(i)$ - individual profile slope errors.

Notes: 1) The lead trace is on the wheel reference cylinder;

2) The profile trace is in the wheel mid-face section;

3) The slope errors are defined as in DIN 3960 (see Figure 6.2);

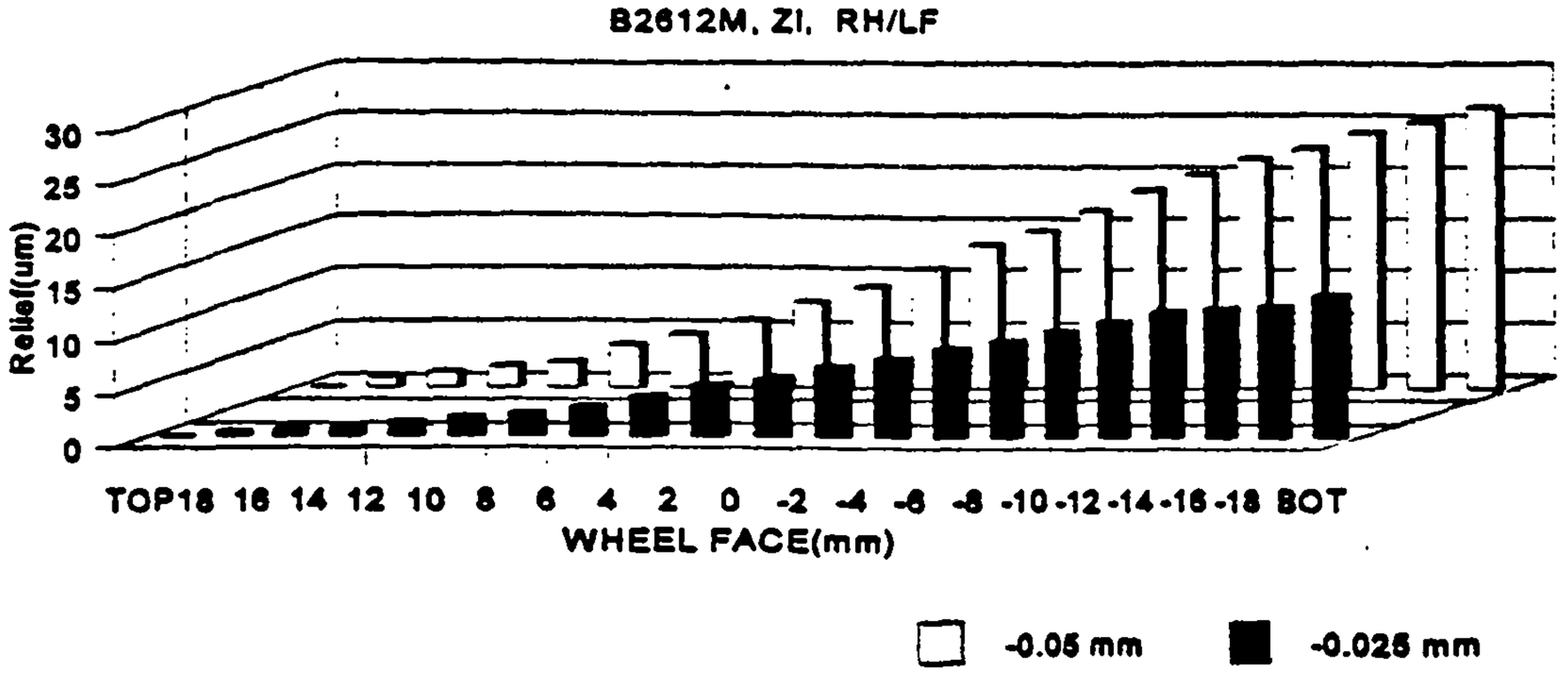
4) The sign of errors depends on the flank and hand of the worm.

Table 5.3 (continued)

This study makes it possible to compare actual (measured) traces with theoretical values. This direct comparison makes it much easier to identify the sources of worm wheel errors.

Of particular significance is the result (shown in Figures 5.18 and 5.21), that the composite effect of several setting errors can be calculated, with sufficient accuracy, simply by adding the effects of the individual errors, which appear not to interact significantly. This suggests that it may be possible to devise relatively simple rules for 'correcting' the wheel flank when necessary. It is noteworthy that the effect of an increased hobbing centre distance on wheel profile is opposite to that of an increased swing angle, which may be a useful guideline for cutting wheels of accurate profile.

The effects of swing angle and hobbing centre distance on wheel profiles are obviously significant (see Figures 5.19 and 5.20). For instance, 10' swing angle of hob setting will cause a profile error of 15 μm in worm wheel mid-face section for the sample. To achieve the required contact patterns, 0.4 mm variation of cut distance and 10' variation of swing angle from the nominal settings are often used by hobbing machine operators. That is the reason why measurements of worm wheels relative to hobs must use reliable real hobbing settings, as mentioned in chapter 3.



Notes: 1. "-0.05 mm" means that hob is 0.05 mm lower than the nominal height (refer to the figure below) .
2. The relief at bottom is nearly half of the axial datum shift for this geometry.
3. Variation of relief is nearly linear.
4. The entry is on bottom side.
5. Positive relief is equivalent to minus metal.

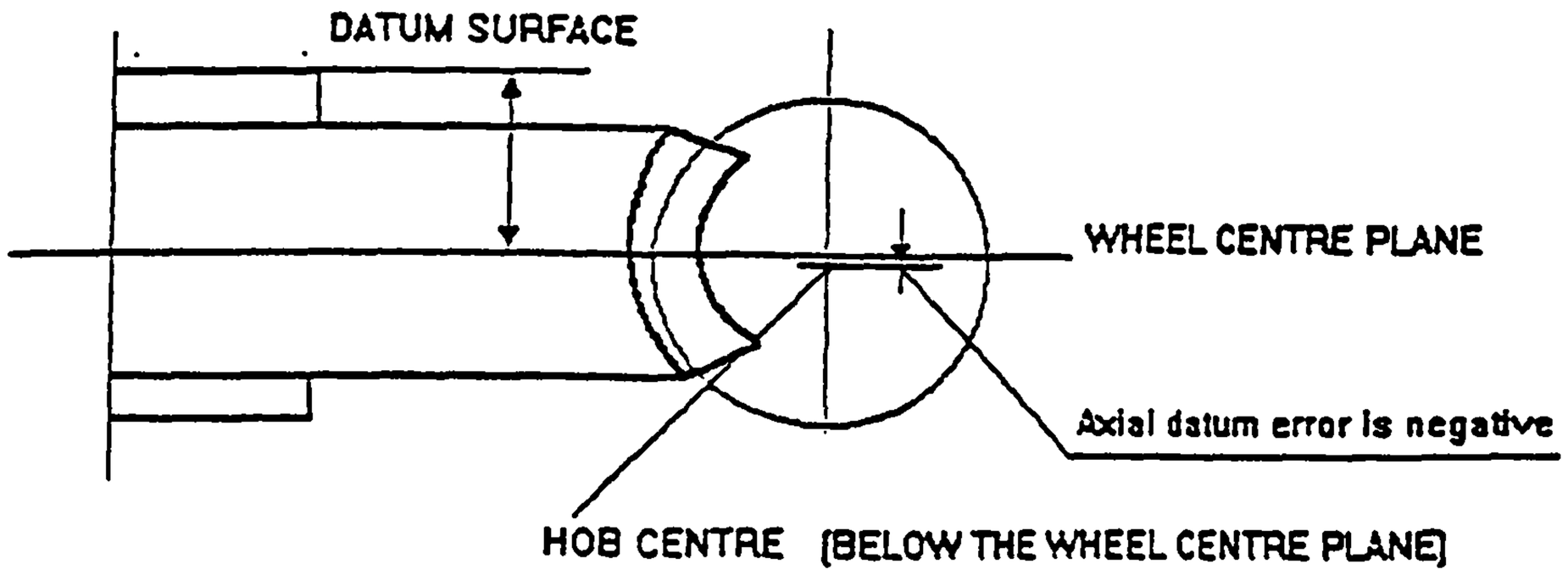
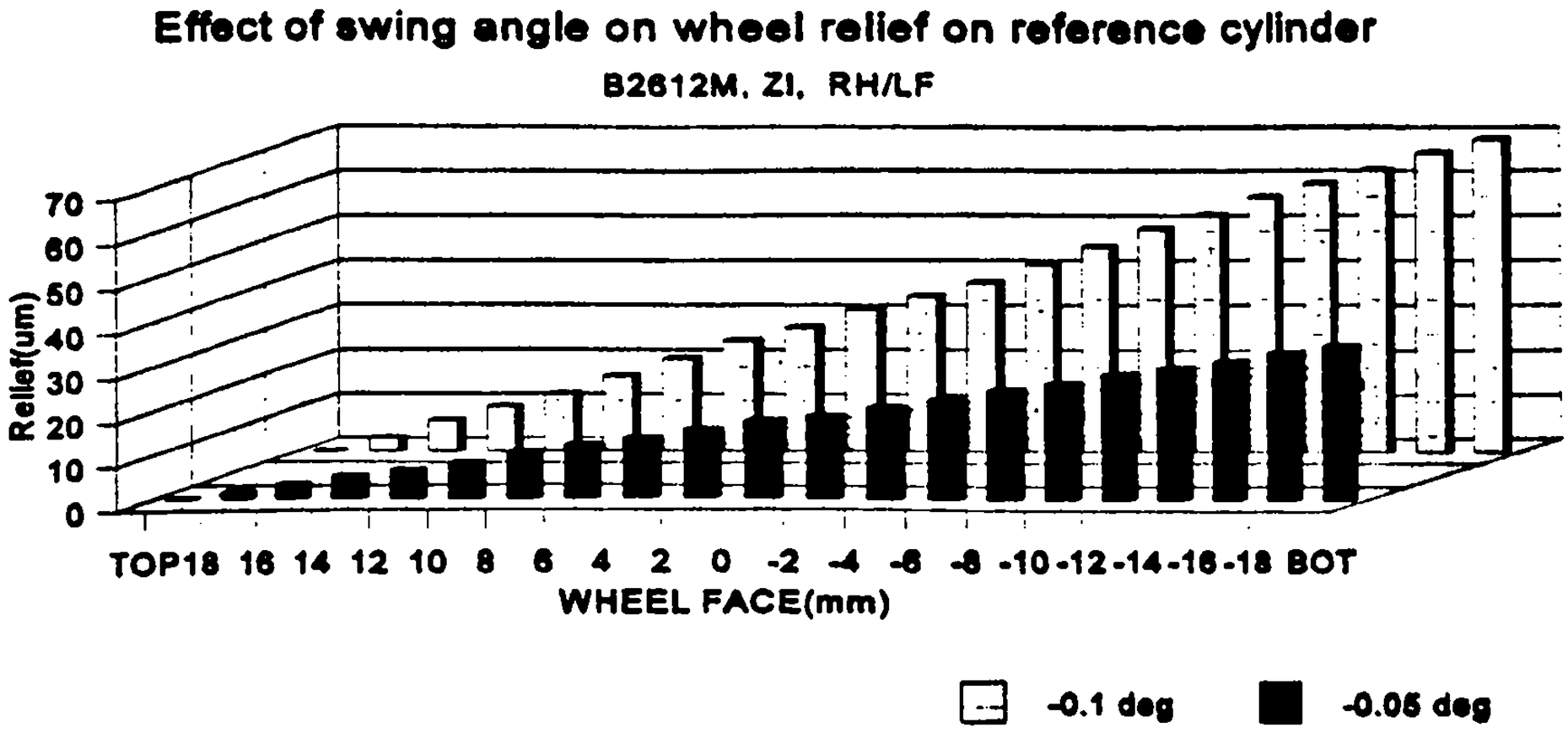


Figure 5.15 Effect of axial datum (hob setting height) on wheel lead trace



Notes: 1. The swing error signs refer to the figure below.
2. The slope (relief/face) equals to the swing (in radian).
3. The variation is nearly linear.
4. The entry is on bottom side.
5. Positive relief is equivalent to minus metal.

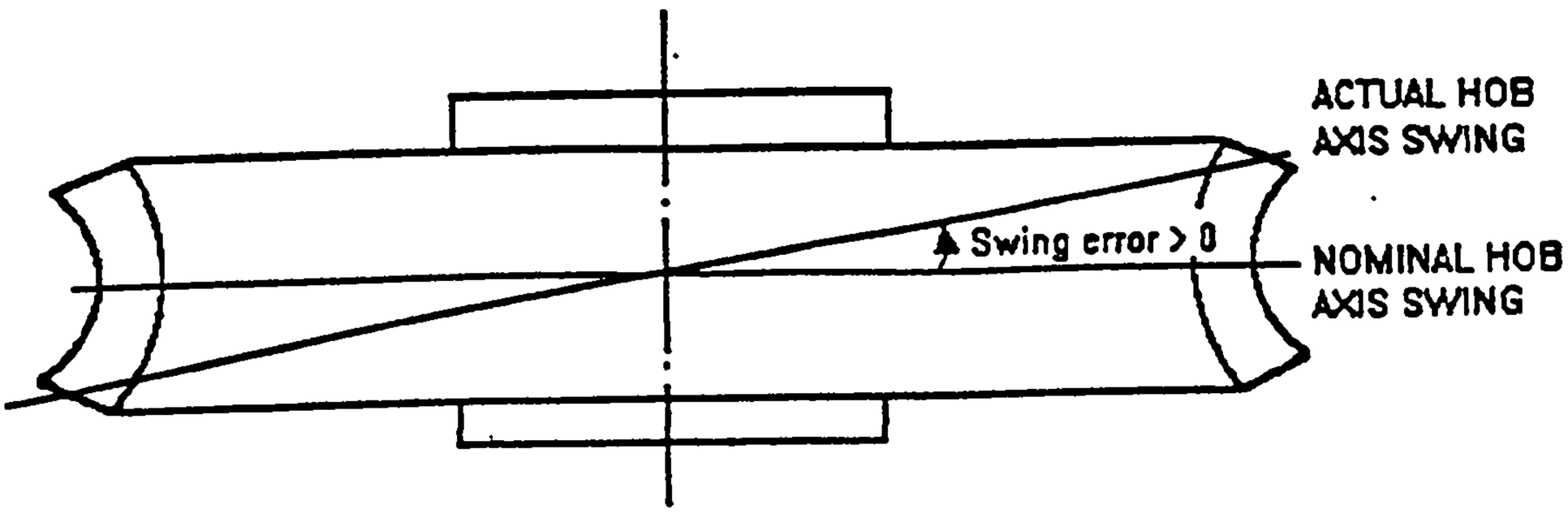
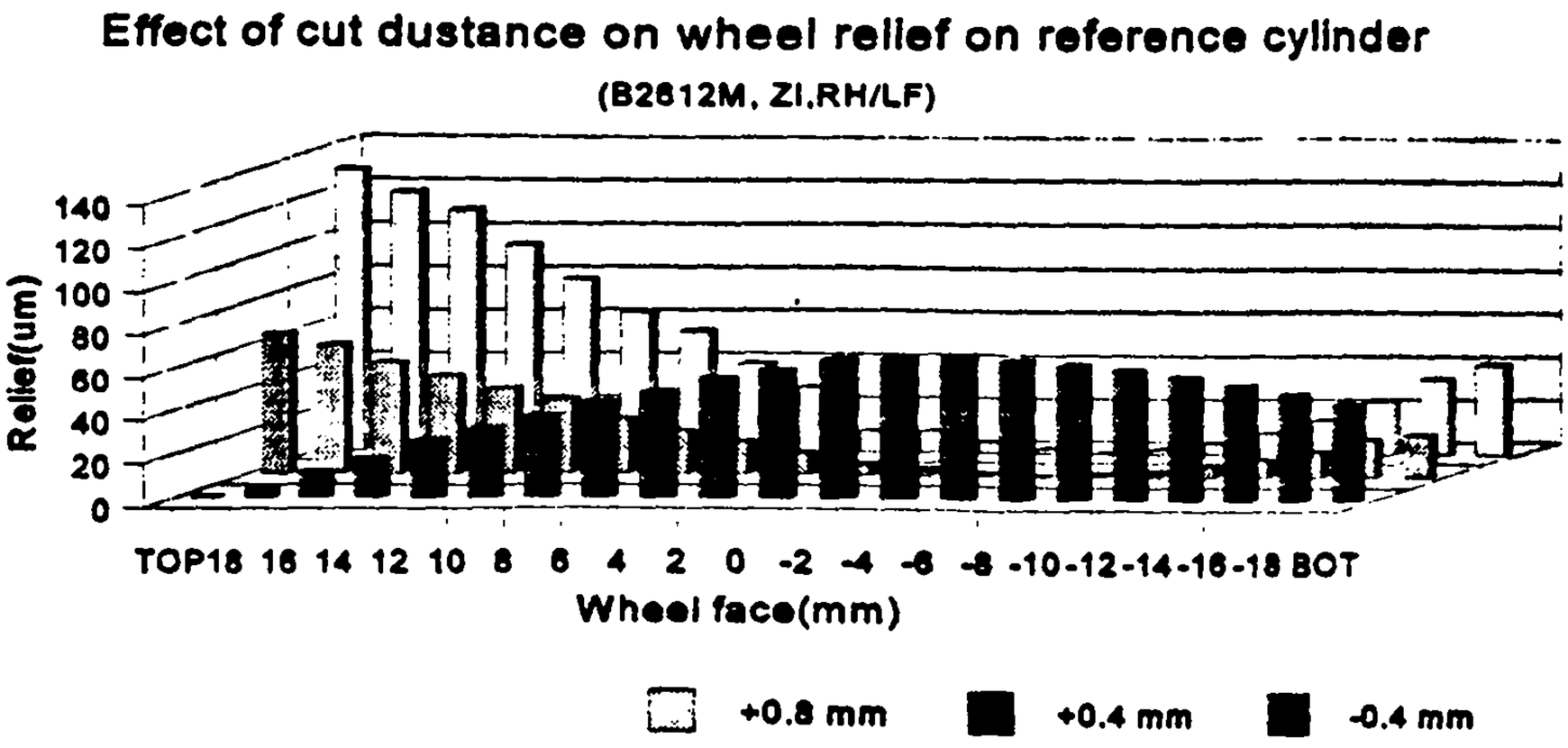


Figure 5.16 Effect of swing angle on wheel lead trace



Notes: 1. "+0.4 mm" means that the actual cut distance is 0.4 mm larger.
"-0.4 mm" means that the actual cut distance is 0.4 mm smaller.
2. Variation is approximately parabolic, but with an offset.
3. The high(low) point is at the hob base circle, for this sample, Rb=5.983 mm.
4. When centre is smaller, split contact might happen at the two ends.
5. The entry is on bottom side.
6. Positive relief is equivalent to minus metal.

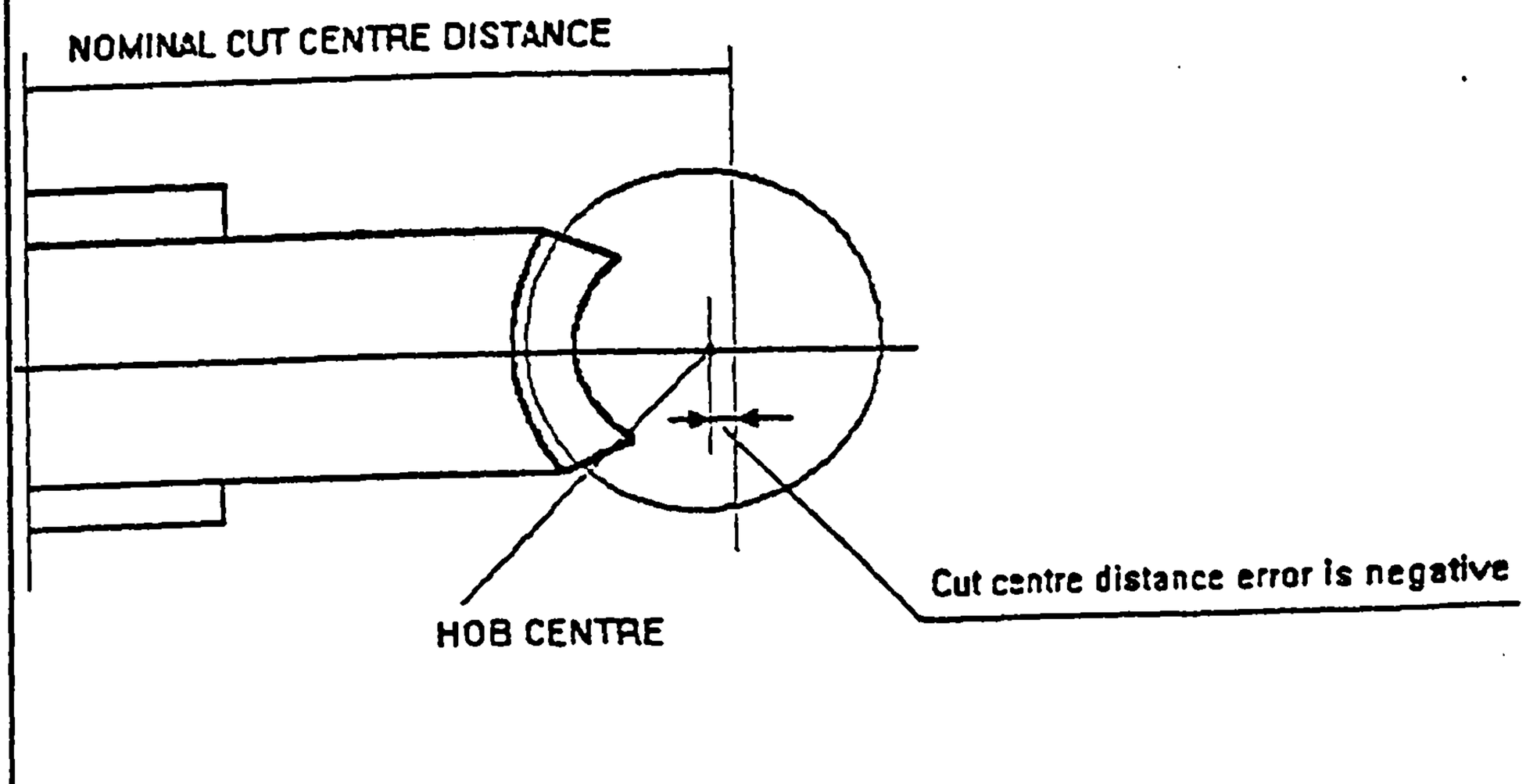
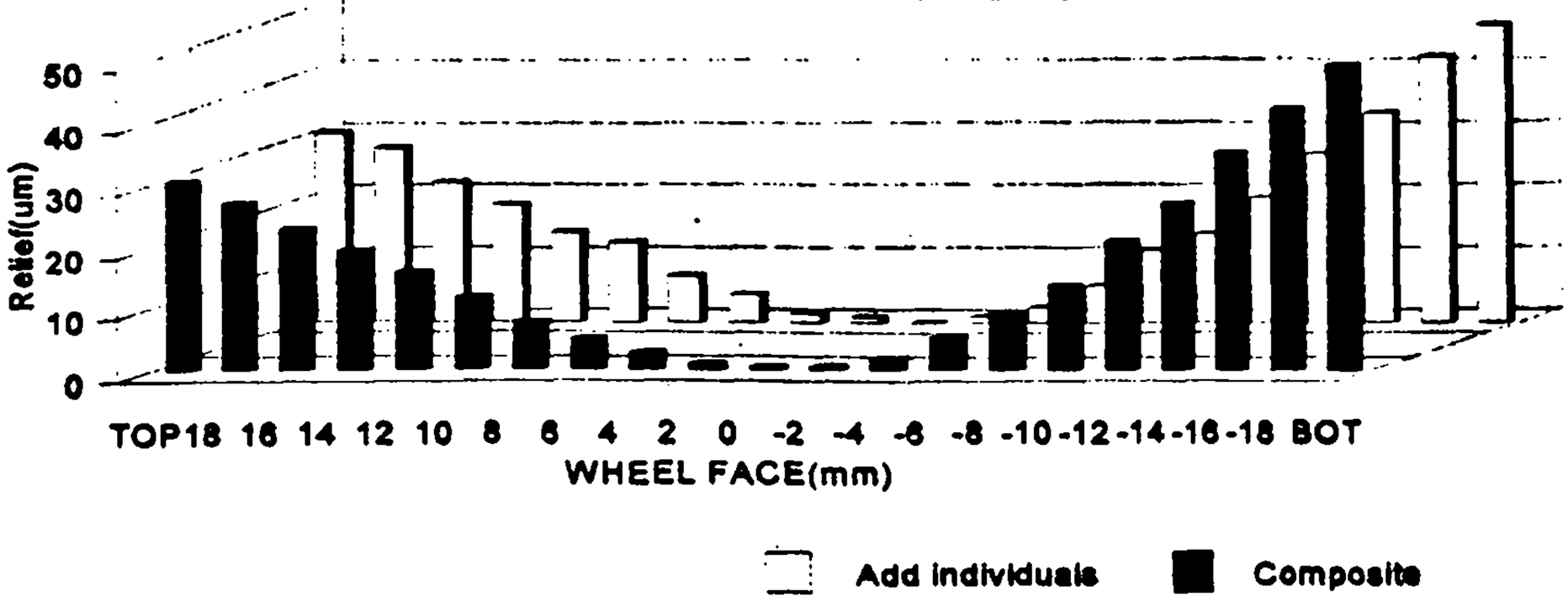
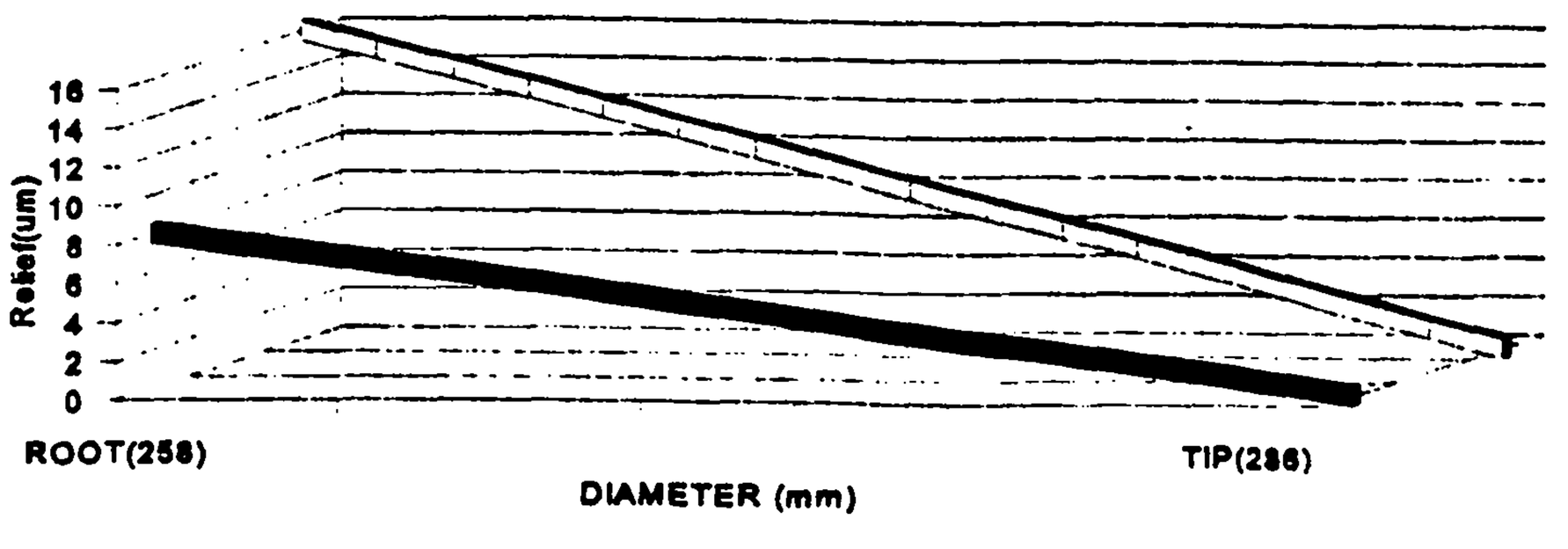


Figure 5.17 Effect of hobbing centre distance on wheel lead trace



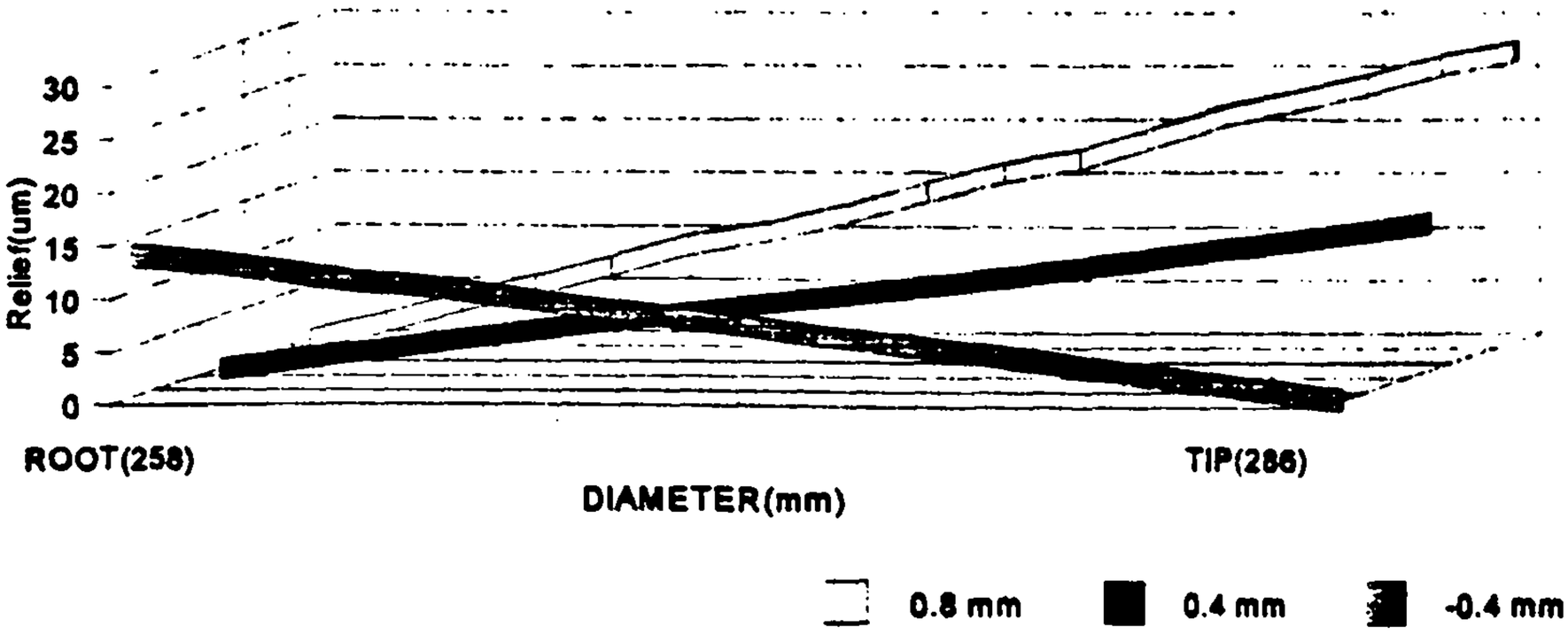
Notes: 1. Composite effect can be simply calculated by adding each individual effect .
2. Individual setting errors are: centre +0.4mm, axial datum -0.05 mm. and swing 89.95 deg.
3. The entry is on bottom side.
4. Positive relief is equivalent to minus metal.

Figure 5.18 Effect of composite hob setting errors on wheel lead trace



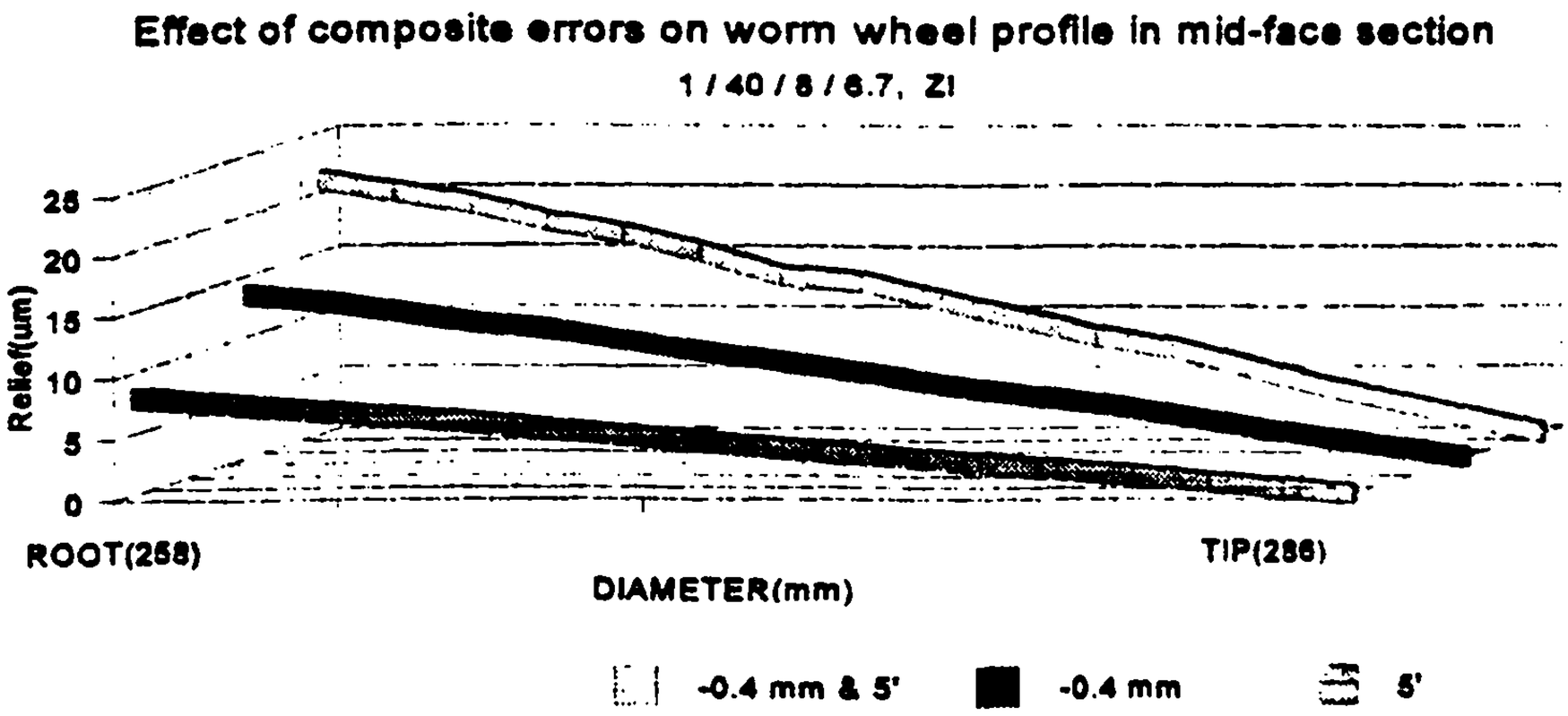
Notes: 1. "10'" denotes the hob swing angle in degrees.
2. Positive relief is equivalent minus metal.
3. The relief trace is nearly a straight line.
4. The relief is almost proportional to the swing angle.

Figure 5.19 Effect of swing angle on worm wheel mid-face profile



Notes: 1. "0.4 mm" denotes that actual cut distance is 0.4 mm larger than C.D.
2. Positive relief is equivalent minus metal.
3. The relief trace is nearly a straight line.
4. The relief is almost proportional to the deviation in cut distance.
5. Positive deviation in cut distance causes minus metal at wheel tip.

Figure 5.20 Effect of hobbing centre distance on worm wheel mid face profile



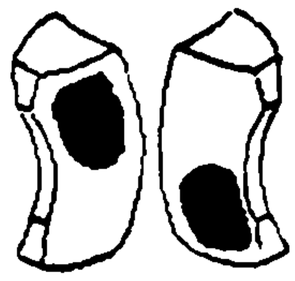
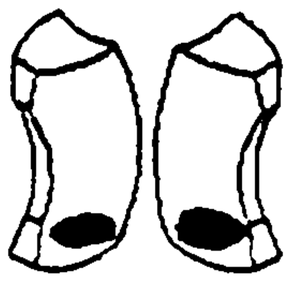
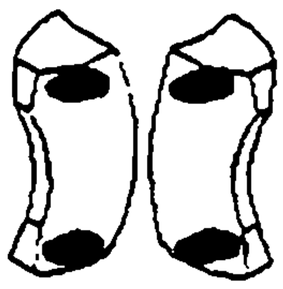
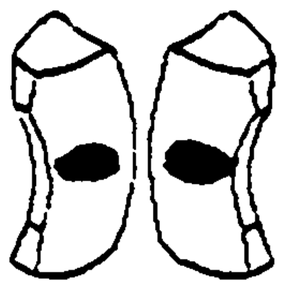
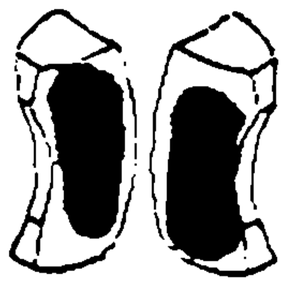
Notes: 1. "-0.4 mm" denotes that actual cut distance is 0.4 mm smaller than C.D.
2. "5'" denotes the hob swing angle in degrees.
3. Positive relief is equivalent to minus metal.
4. The relief trace is nearly linear.
5. The total profile deviation is almost the simple sum of the two individual deviations.

Figure 5.20 Effect of composite errors on worm wheel mid face profile

5.6 Effect of individual errors on a contact pattern

Effects of various errors on contact patterns have been explored with application of the non-elastic analysis software. Typical results are given in Table 5.3. Validation against results obtained from manufacturing practice [Tang, 1984] has been carried out with good agreement. The purpose of this study is to supply another set of basic guidelines to assist in improving worm wheel accuracy. Especially, it has been proved very helpful when used to cross-check the results obtained from the analytical measurement of worm wheels.

Combinations of more than one error result in composite errors on a wheel flank. The theoretical study provides guidelines for investigation into worm wheel manufacturing errors although a practical situation can often be much more complicated than the results obtained by theoretical study in this chapter. The advantage of the simplistic method is that a quick input of the hob profile deviation can be applied to give a satisfactory approximation of the actual wheel flank. The solutions of this type are a simplification of inaccuracy but serve as useful references for routine daily work.

Case	Contact Pattern	Possible Causes
1		Expected
2		Hob is set too high
3		Cutting centre distance is too small
4		Cutting centre distance is too large
5		1. The worm profile pressure angle too small or the hob profile pressure angle too large: 2. The worm pitch too large or the wheel pitch too small.
6		The hob size is too small
7		1. The hob size is too small 2. Design problem
8		1. Hob has large pitch error 2. Worm has large pitch error 3. Wheel has large pitch error
9		Large variation along hob shaft axis

Examples of incorrect contact patterns

Table 5.3

CHAPTER 6: IMPLEMENTATION AND APPLICATION II

STUDY OF WORM WHEEL MEASUREMENT RESULTS FOR ZI WORM GEARS

6.1 Introduction to the worm wheel measurement software

The worm wheel metrology software has been implemented on a dedicated gear measuring machine - the Gleason GMS 430. This gear measuring machine is equipped with a HP (Hewlett Packard) 9000 / 300 desktop computer with programming language HP Basic 4.0. The total RAM is 1 MB and the processor is a MC68010. Therefore the programmer must bear in mind that:

1) the RAM resident program has restriction on its size and

2) the speed is very low.

To overcome these problems, the worm wheel metrology software has been divided into modules and the computation procedure has been simplified. A module is resident in memory only when it is being used, otherwise it is purged from the memory. However, the dominant factor for achieving an acceptable speed for routine work is to simplify the mathematical models, as discussed in chapter 4.

The Gleason GMS 430 was supplied with spur/helical gear measuring software, but the worm wheel metrology software has been written separately, and is divided into 5 modules. They are the common module, the worm data input module, the wheel data input and calculation module, the measurement module and the standard and evaluation module. Each module is purged from memory as soon as it has finished its functions and loaded into memory just before it is used. This operation does not need any medium for communication between modules, since all the shared information is kept in the memory in the common data.

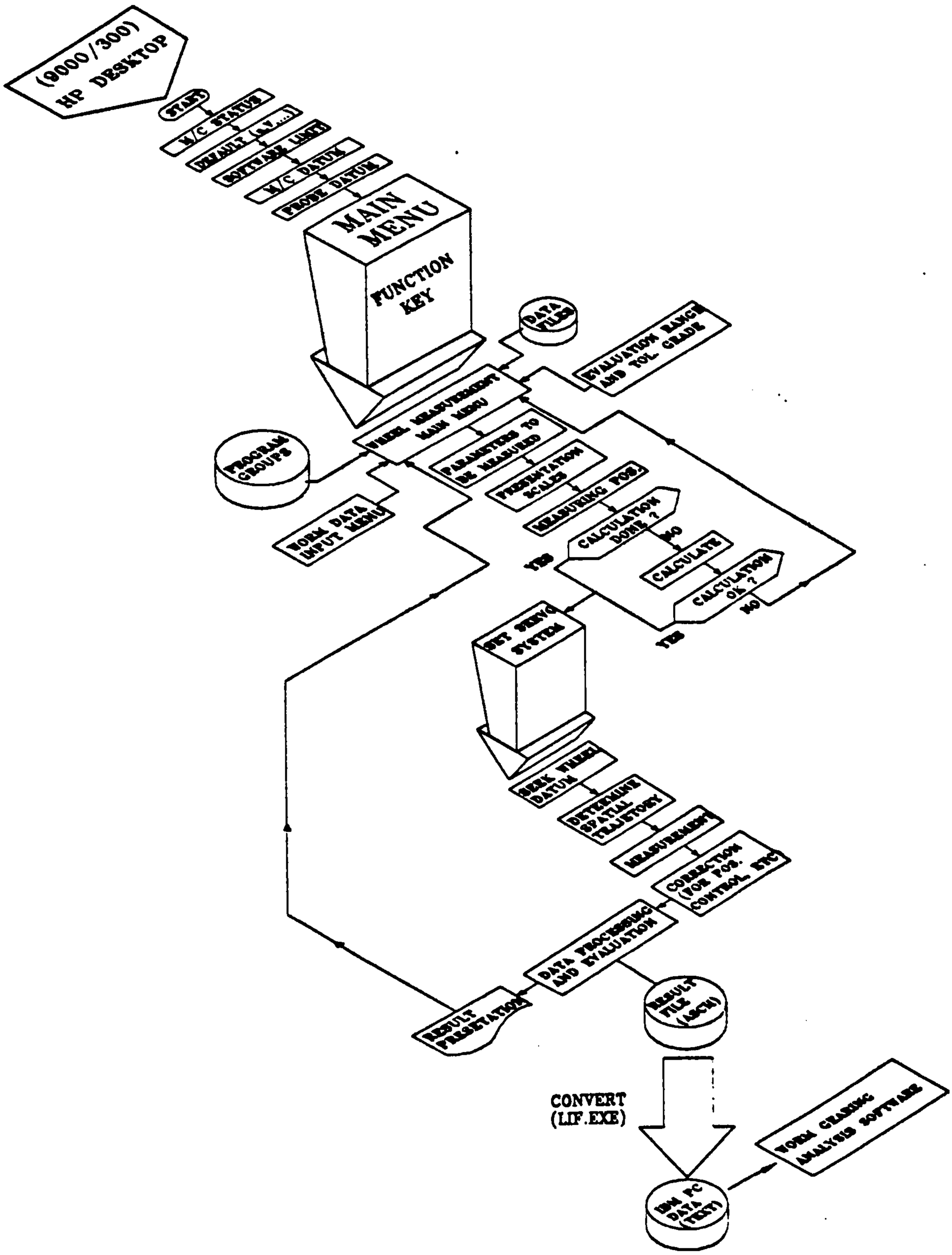


Figure 6.1 Logical structure of the worm wheel measurement software

The software is robust and prevents the user damaging the machine by mistaken operation or pressing the wrong keys. This machine is protected by frequent checking of its status by the high level software, which is not ideal, but a good solution in this case, since no interrupt method is available.

The logic structure of the worm wheel measurement software is given in Figure 6.1.

The symbols and terms used for spur and helical gear analytical measurements specified by DIN 3960 have been adapted for worm wheels, because no standards are currently available for analytical measurement of worm wheels themselves. The tolerances given by DIN 3962 for spur and helical gears have thus been 'borrowed' for the worm wheels in the software developed - using DIN grade 8 and grade 10 for classes C and D of BS 721 respectively. This is largely a matter of convenience, since DIN 3962 is widely used and the Gleason GMS 430 was already provided with DIN evaluation software. These tolerances given on the result sheets are mainly of interest in allowing the wheel measurement results to be compared with the tolerances of low quality cylindrical gears of similar size. For the sake of convenience, the symbols and terms used by DIN 3960 are listed in Table 6.1 and defined in Figure 6.2.

Symbol	Term
f_f	profile form error
$f_{H\alpha}$	profile slope error
F_f	total profile error
f_β	helix form error
$f_{H\beta}$	helix slope error
F_β	total helix error

Table 6.1 Symbols and Terms Used for Wheel Measurement

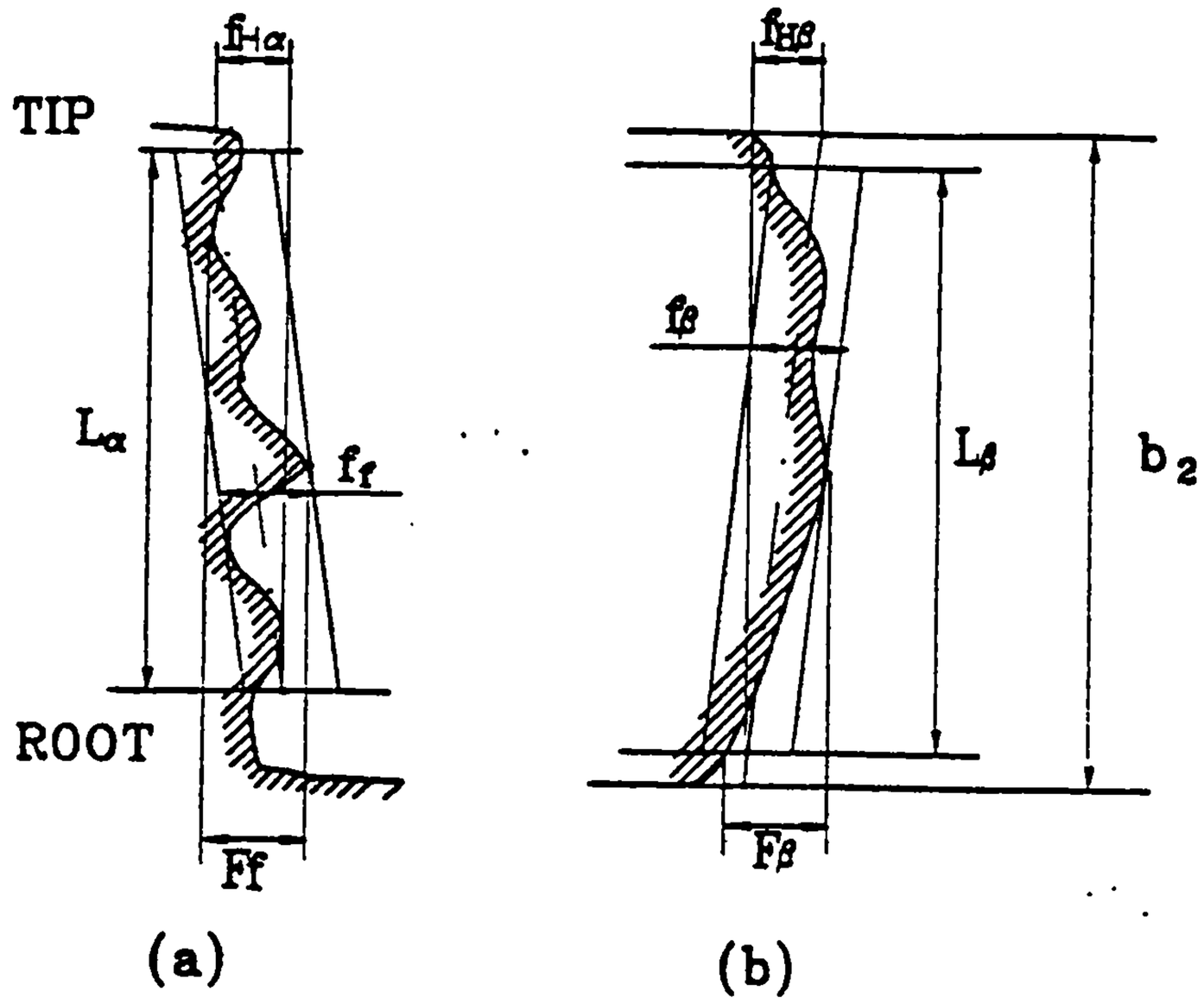


Figure 6.2 Definition of (a) profile and (b) lead errors

The measuring ranges are determined by the software as discussed in chapter 3. The evaluation ranges are user-specified. The helix slope error $f_{H\beta}$ labelled on the result sheet is the effective error over the face width b_2 as shown in Figure 6.2(b), even though it is evaluated over the user-specified range. Sufficient information about the flank geometry can be obtained by measuring several 'profile' errors, in user-defined transverse sections, and a single 'lead' error on the reference cylinder, which is determined by the software. Normally, more than one transverse section needs to be measured, since worm wheel profiles may vary significantly from one section to another. This is different from the method adopted when inspecting parallel axis spur and helical gears, whose profiles normally vary very little across the face width.

6.2 Prediction of the contact pattern using measurement results

The contact marking pattern and analytical measurement of a worm wheel relative to the worm are related. Generally speaking, plus metal on a trace from an analytical measurement indicates a heavier contact. Of course, the area of contact actually marked is also related to the amount of blue dye used for the testing and may also depend on the pitch accuracy. For a given worm wheel, the thicker the blue dye is, the

larger the contact pattern is, so that quantitative interpretation of contact marking needs sufficient information.

To show how the contact pattern can be predicted using the analytical measurements, examples of the interpretation of worm wheel measurement traces (the contact patterns associated with several measurement traces) are demonstrated in Table 6.2.

More experience needs to be gained in the interpretation of worm wheel measurements, bearing in mind that even the measurement results of a spur gear can be confusing. However, the following suggestions can be useful even at this stage:

The profile errors in a section on the exit side, in the mid-face section and in the involute section are important. This is because they contain the most informative messages. The contact between worm and wheel flanks is likely to extend across from the exit side to the mid-face section, while the hob profile errors in its axial section are approximately directly transferred onto the wheel mid-face section. The involute section, too, is helpful for identifying some error sources, since experience obtained from spur gear measurement can be used. However, measurement of profile errors in the involute section only is certainly not sufficient, since it is on entry side and the effects of some setting errors cannot be seen in this section.

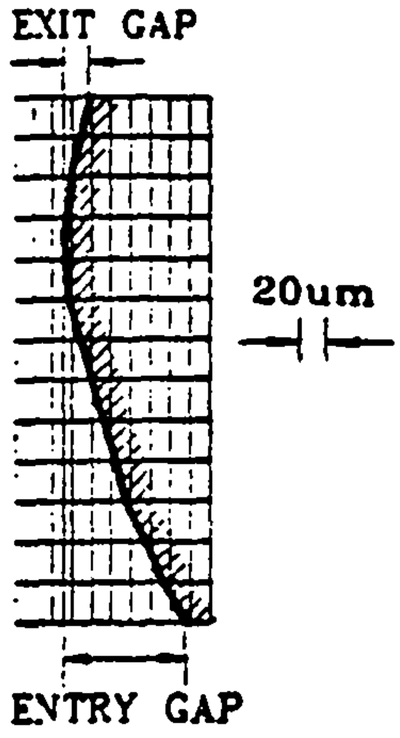
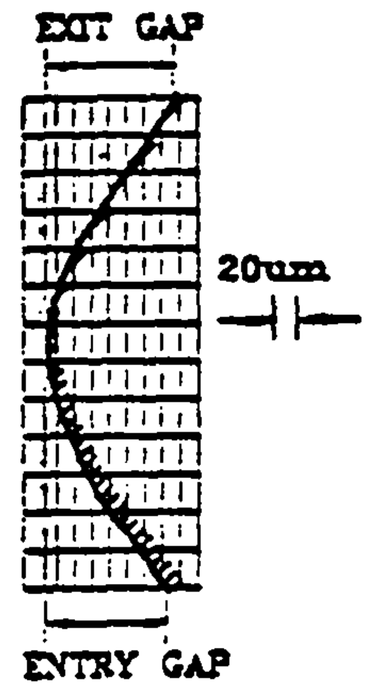
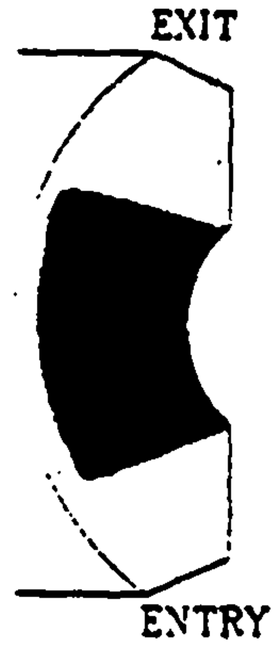
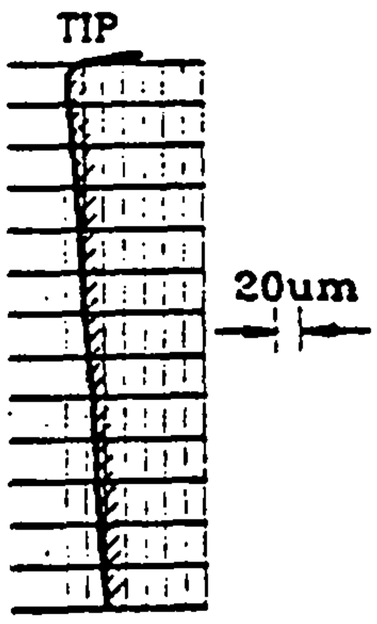
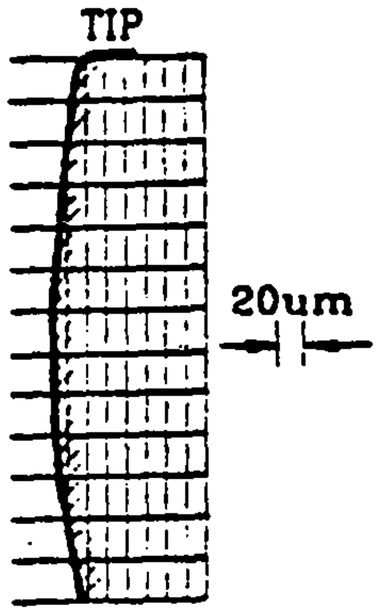
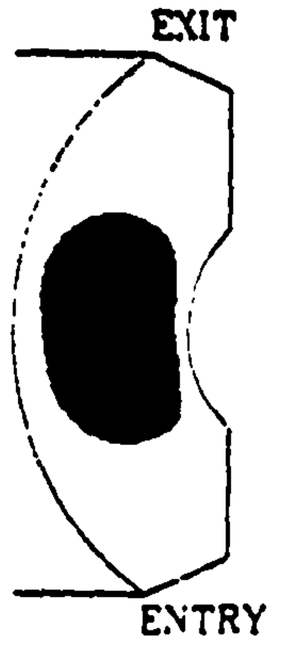
Case	Trace of result	Comment	Contact pattern
1 (lead trace)		1. Lead trace with high point on exit side; 2. The entry gap is 125 μm and the exit gap is 25 μm.	
2 (lead trace)		1. Lead trace with high point near middle; 2. The entry gap and the exit gap are both150 μm.	
3 (profile trace)		1. Profile trace with plus metal near tip; 2. Lead trace as case 2.	
4 (profile trace)		1. Profile trace with high point at middle of tooth depth; 2. Lead trace as case 2.	

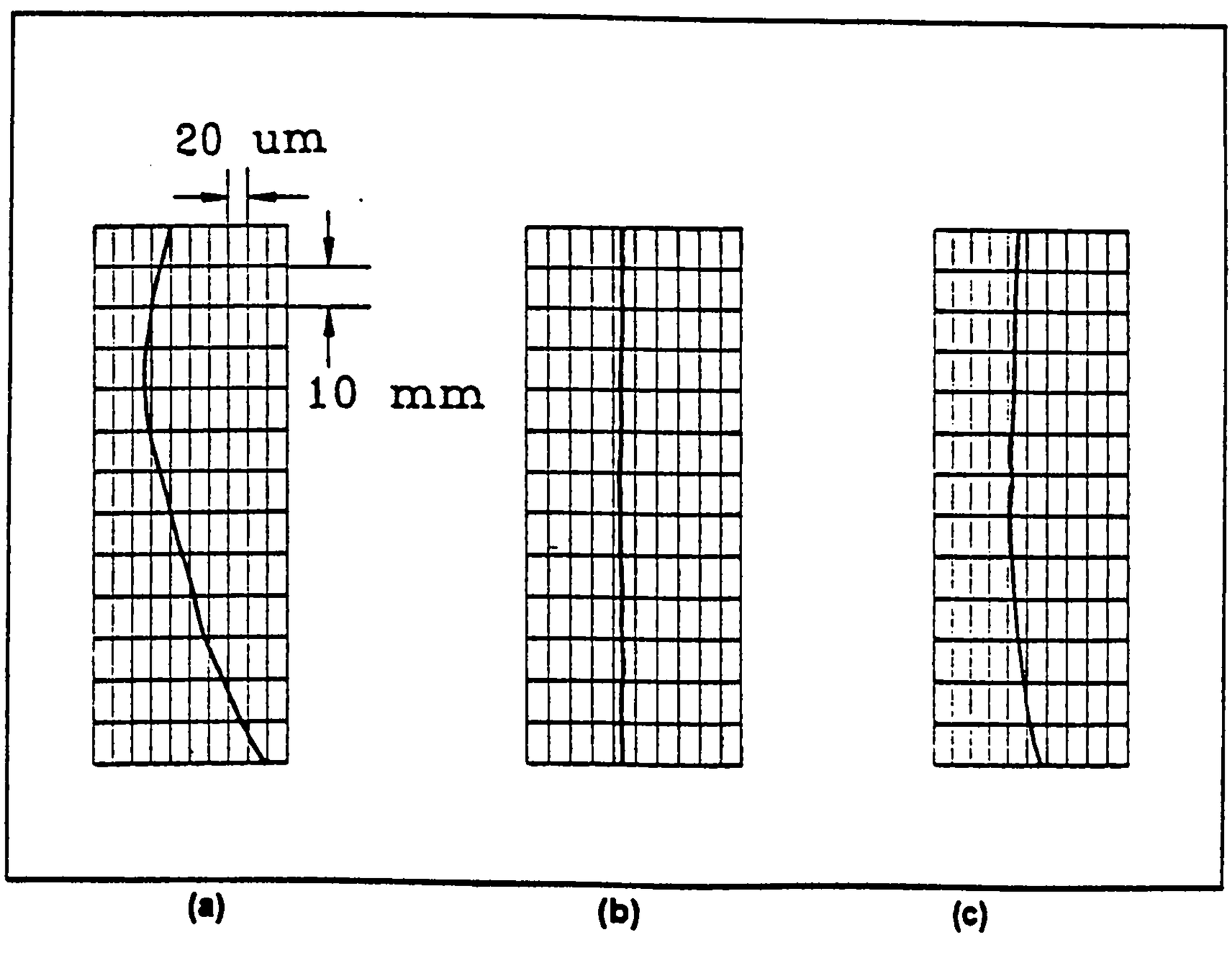
Table 6.2 Interpretation of worm wheel measurement results

6.3 Typical measured lead traces

Many worm wheels have been measured for both research and commercial purposes since the worm wheel measurement software was developed. The worm wheels are measured with respect to their mating worms and some typical measurement results will be discussed in the following sections.

6.3.1 General results

In general, three types of measured lead traces have been found to be most interesting. They are from 1) mismatch worm wheels; 2) no-mismatch worm wheels and; 3) bedded-in worm wheels. Typical examples of these three types of lead trace and their interpretation are shown in Figure 6.3.



(a) typical worm wheel with mismatch
(b) typical worm wheel without mismatch
(c) typical bedded-in worm wheel

Figure 6.3 The three types of worm wheel lead trace (exit side at top)

The lead trace of a standard mismatch worm wheel as in Figure 6.3 (a) shows that the contact pattern is intended to be near the exit side of the wheel face. This is demonstrated by the facts that the high point is near the exit side and the entry gap is much larger than exit gap. The lead trace of the no mismatch worm wheel in Figure 6.3 (b) shows that full face contact will occur, since the lead trace is almost straight, showing that the wheel is completely conformal to its mating worm. The main disadvantage of a no mismatch worm wheel is that the contact pattern is very sensitive to the manufacturing errors due to the high degree of conformity. The lead trace of a bedded-in worm wheel is given by Figure 6.3 (c). The main interesting result from the bedded-in worm wheel is that the contact region of the wheel flanks appears to have been regenerated by the mating worm in the gearbox. This results in a straight portion of the trace which indicates that the wheel in this region had line contact with the mating worm during the bedding-in process.

6.3.2 Worm wheels used for lifts

A new worm wheel designed for use in a lift tends to have a different shape of lead trace. The high point is normally found at the middle of face width, as shown in Figure 6.4 (compare with Figure 6.3 (a)). This type of lead trace, which gives central contact, is advantageous for applications where the worm wheel drives the worm and for reversing drives in which the torque may not reverse. This type of worm gear normally has such large entry and exit gaps that full face contact is seldom obtained even after many hours of running.

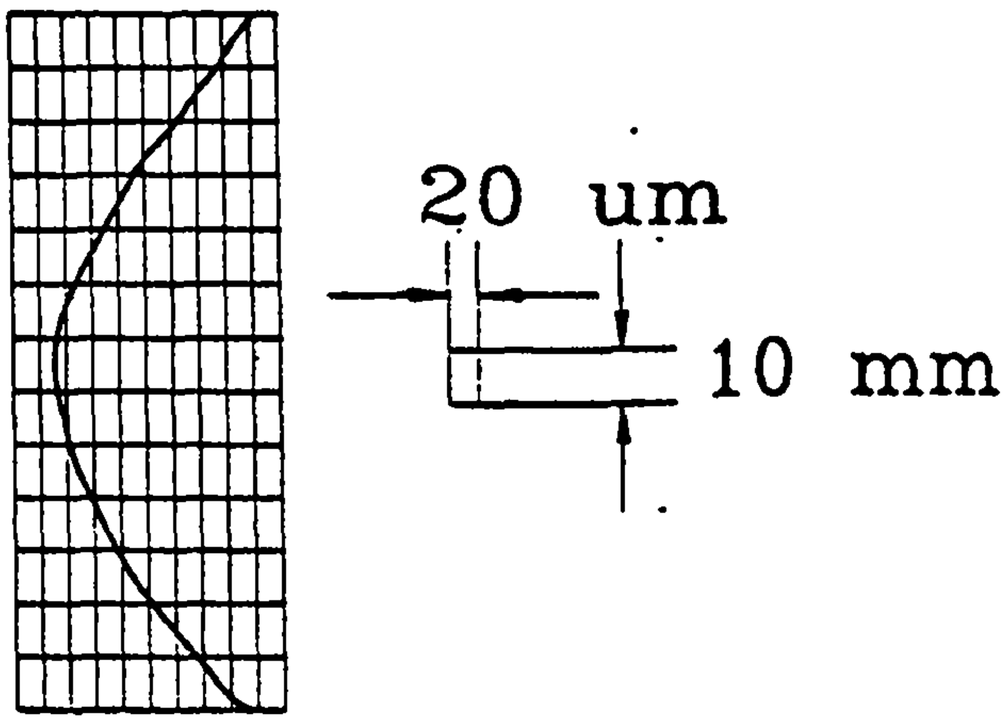


Figure 6.4 Lead trace of typical worm wheel for lift drive

6.3.3 Difference between bedded-in and no mismatch wheels

When measured, the no-mismatch wheel is found to be completely conformal with the mating worm, because the hobbing data is the same as that of worm gear box. However, the lead trace of a bedded-in worm wheel does not tend to be a complete straight line even if the full face width of the wheel flank is completely worn. The author is not very sure what causes this difference, but some possible reasons may explain it. One possibility is that the wheel flank wear is generated when load is applied and the tooth load and tooth deflections are not uniform across the face width. Therefore the lead trace measured after the load has been released does not show complete conformity, even through there was contact between worm and wheel flanks under load. Another possible cause is that the relative position between the worm and wheel in the worm box is not exactly the same as that of the manufacturing process due to defections, clearances and errors etc. Of course, both effects may contribute at the same time. A further possibility is that load variations in service, with corresponding variable worm shaft and tooth deflections, mean that the wheel never fully beds in at any load level.

6.3.4 Double and triple cut wheel flanks

A double or triple cut may be applied when hobbing a worm wheel to achieve the required contact pattern. The purpose of a double cut is to bring the contact pattern to the exit side and that of a triple cut is to bring the contact pattern to the middle. An example of double cut lead traces is given by Figure 6.5.

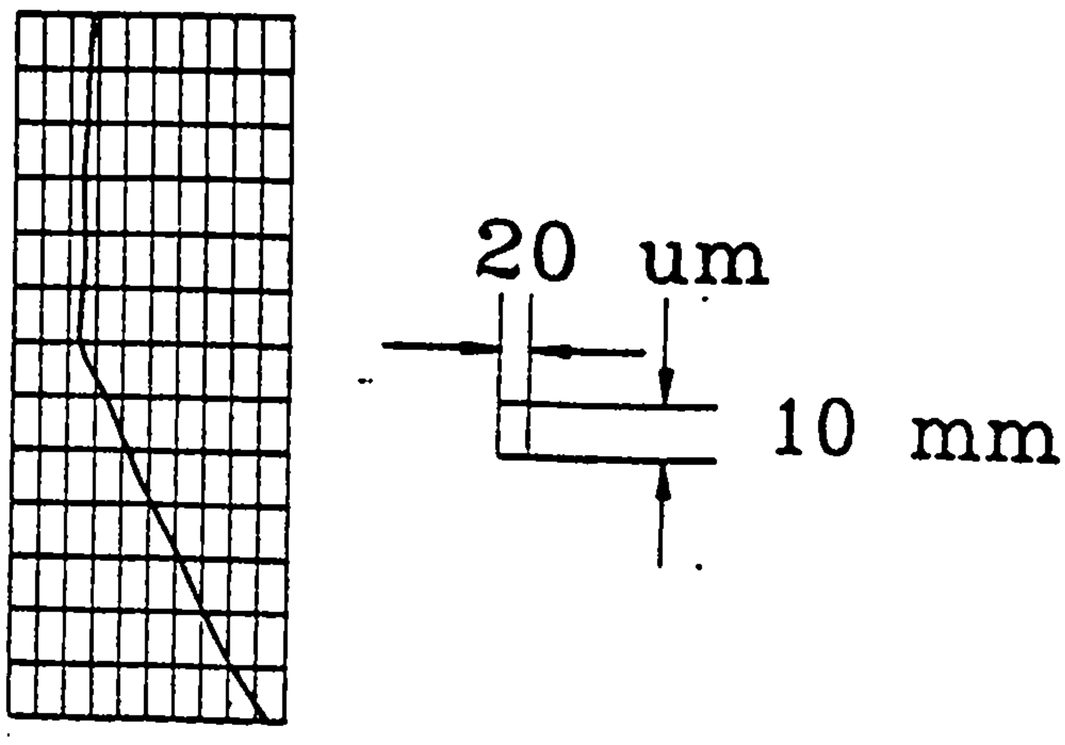


Figure 6.5 Lead trace of a double-cut worm wheel flank

The characteristic of a double cut lead trace is that a sharp change in slope occurs at the joint of the two surfaces. Industrial practice shows that double cut operations do not cause any major problems. This may be explained by the fact that the double cut reduces the sensitivity to manufacturing errors and the bedding-in process will eventually smooth down the wheel flank at the discontinuity. A triple cut is intended to be used in applications where a reversing drive is needed. A double or triple cut may have to be applied when the hob oversize has been significantly reduced by extensive resharpening.

6.3.5 Entry and exit gap

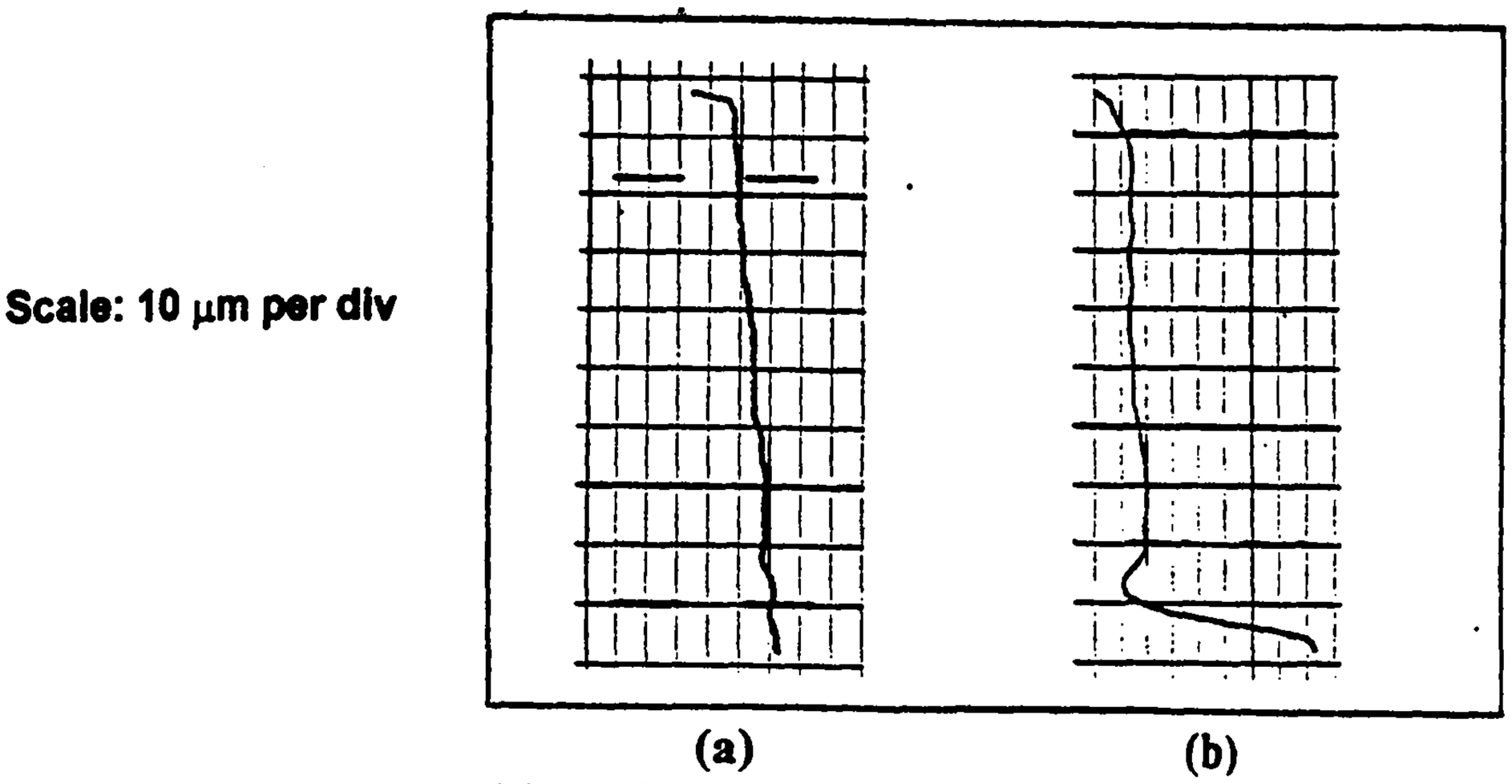
The entry gap helps of course to provide easy entry of lubricant to the contact zone and it is in fact widely understood that the major purpose of an entry gap is to obtain a better lubrication condition. This is however wrong, because the entry gap normally used is much larger than required for effective lubricant entrainment. The exit/entry gaps of the wheels measured went up to 6 μm per mm face width or even larger. In fact, the most important function of the entry gap is to allow for relatively large tolerances in worm wheel manufacturing. This is necessary, since almost every single setting parameter is critical to worm gears. With a large entry gap, when wheel deflects axially and the teeth and worm bend under load, the contact tends to become more central but still leaves an adequate entry gap for lubricant, even after bedding-in.

6.4 Extent of wear on worm and wheel flanks

Substantial wear has been found on most of the wheel flanks measured. Figure 6.6 shows an example of wear. Figure 6.6 (a) is the profile measured before running and (b) is that measured after running. As can be seen, the wear depth measured is up to 70 μm at the reference diameter after 4 hours of running under load at half the catalogue load rating. Badly worn wheel flanks can have wear depths up to a few millimetres, which certainly is of great concern although it does not represent the normal case.

The profiles of one worn wheel can be very different from another and from what is expected, just as one cylindrical gear's profile errors can be very different from others.

For example heavier wear of both flanks may occur on the same side of the face width rather than, as would normally be expected, on only the exit side of each flank. This has been confirmed by analytical measurement, by contact pattern testing and by taking an impression from the worm teeth. This phenomenon may cause problems on hard/soft worm gears, for instance very low efficiency and short life. Especially, it can ruin hard/hard worm gears since it indicates a very poor load distribution across the face.



(a) profile measured before running.
(b) profile measured after running.

Figure 6.6 Wear on worm wheel flanks

It is generally assumed that a worm stays unchanged once it has been made, since the worm flanks are hardened. However, analytical measurement shows that wear on even a carburized worm flank is normally quite significant. Many measurement results show that 20 ~ 30 μm wear on worm flanks is very common and even 60 μm wear has been found on examples. Wear on a worm flank is normally not uniform and heavier wear usually appears on the dedendum of the worm flanks. Possible reasons for the wear that occurs and its distribution are:

1) the large ratio between the number of teeth on the wheel and worm;
2) the worm flanks are not hardened enough (refer to chapter 7);
3) lubrication conditions are poor when contact occurs near the wheel tip, close to the envelope of contact lines on the worm flank (refer to chapter 2);
4) the edge of the wheel tip is sharp.

CHAPTER 7: IMPLEMENTATION AND APPLICATION III

EXAMPLES OF APPLICATION

The objectives of this PhD project were to develop analytical techniques and metrology of worm gears for the DTI research project 'Improving Worm Gear Performance', carried out by the Design Unit, and the EPSRC research project 'Improving the Stress Analysis of Worm Gearing', carried out by the Department of Mechanical, Materials and Manufacturing Engineering at the University of Newcastle upon Tyne. However, the actual work done has extended over a much wider range of activities. Apart from its academic achievements - the mesh analysis models developed to match the rapid progress of computer technology and the metrology methods and software which allow, for the first time in the world as far as we know, analytical worm wheel measurement - the project has been industry-driven. The theory and its implementation have been used to support other major worm gearing research projects carried out over recent years in other British universities. In addition, the software developed and expertise gained have been used in commercial projects carried out by the Design Unit. This chapter presents some examples of these applications which have been chosen from investigations which cover various different aspects of worm gearing.

7.1 Parametric study on effect of setting errors on Cavex wormgears

The analysis software has been used to investigate ZC (Cavex) worm gearing. The details are discussed in references [Hu, Pennell and Hofmann, 1996]. The main results of the investigation are summarised as follows:

1) A balance between the conformity of the two mating flanks (which increase load capacity) and the sensitivity of the design to setting errors is needed. The price

to pay for a better conformity is tighter tolerances on manufacturing and setting errors, or alternatively, longer bedding-in procedures.

2) The effects of setting errors on the contact pattern can be quantified. The centre distance errors appear to be most critical, as demonstrated by Figure 7.1. Unless the centre distance of ZC worm gears under running conditions can be held to tighter tolerances than that of typical ZI designs, the higher conformity and better contact line orientation of Cavex worms may actually give few advantages.

(3) The ZC worm gear profiles do create a favourable orientation of contact lines, which extend mainly at right angles to the sliding direction. It is widely accepted that large angles between the contact lines and the sliding direction favour better lubrication.

The investigations have also shown that, normally, a relatively large positive addendum modification must be chosen to establish the advantageous lubrication conditions and avoid undercut, as shown in Figure 7.2(a). ZC worm wheels are easily undercut The reason for this seems to be its small pressure angle at hob tip.

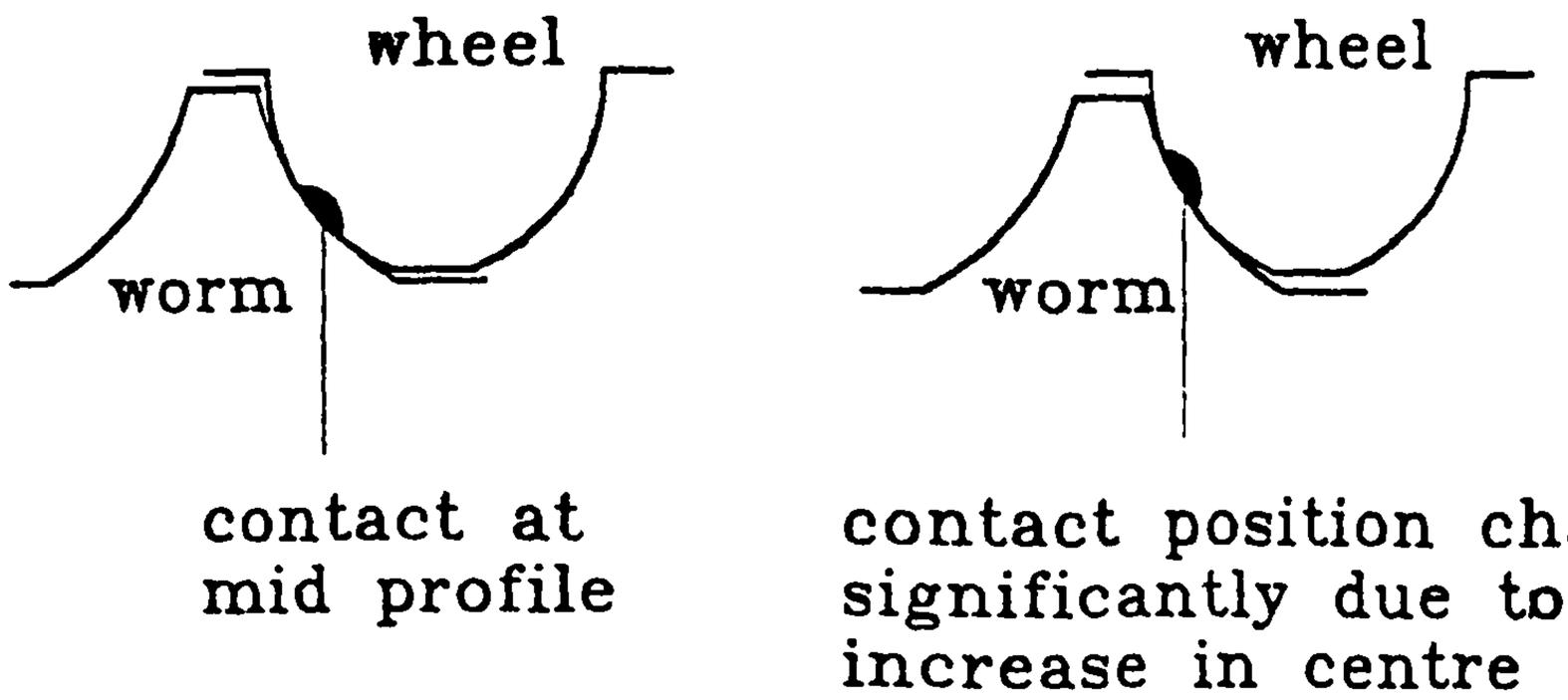


Figure 7.1 The effects of centre distance on the contact of ZC profile

7.2 Investigation into the causes of Cavex wheel surface pitting damage

Laboratory testing under load of even extensively-pitted bronze worm gears has shown that the accuracy of motion transmission, gearing efficiency, noise level, peak load capacity and the heat generated are not significantly affected by the presence of such pitting [ANSI/AGMA 6022-C93]. However the presence of pits in some applications is of concern.

The exact mechanism which causes pits to develop in a tooth flank is not yet fully understood, although several theories have been published. It is, however, generally agreed that pitting is a fatigue phenomenon, caused mainly by cyclic contact stresses. Pitting is thus most likely at points of greatest load concentration, whether caused by profile errors, pitch errors, shaft misalignment errors, poor design or incorrect bedding-in procedures. Many publications show that the highest contact stress is on the exit side and claim that this is the reason why pitting normally occurs on this side. This is not surprising, since worm gear contacts are normally designed to have the load concentrated on the exit side, since an entry gap is always needed. The relative absence of pitting on the entry side is thus inevitable.

Cavex worm gears appear, in service, to be more susceptible to pitting surface damage than ZI (involute) worm gearing, so an attempt has been made to find the reason. One possible reason could be an increased likelihood of localized load concentrations on a Cavex worm wheel flank.

It is a characteristic of Cavex worm gearing that a relatively large positive addendum modification is normally chosen to establish the advantageous lubrication conditions and avoid undercut [Vos, 1996]. The negative effect of this choice is a narrowing of the conjugate contact area on the wheel flank, as illustrated in Figure 7.2 (a). On the area outside conjugate contact, interference can occur between the meshing teeth. This can be seen on some occasions as hobbing marks, shown in Figure 7.2 (b). This interference zone is characterized by intensive rubbing between the worm and wheel flanks, especially if wear has reduced the exit and entry gaps so that heavy contact

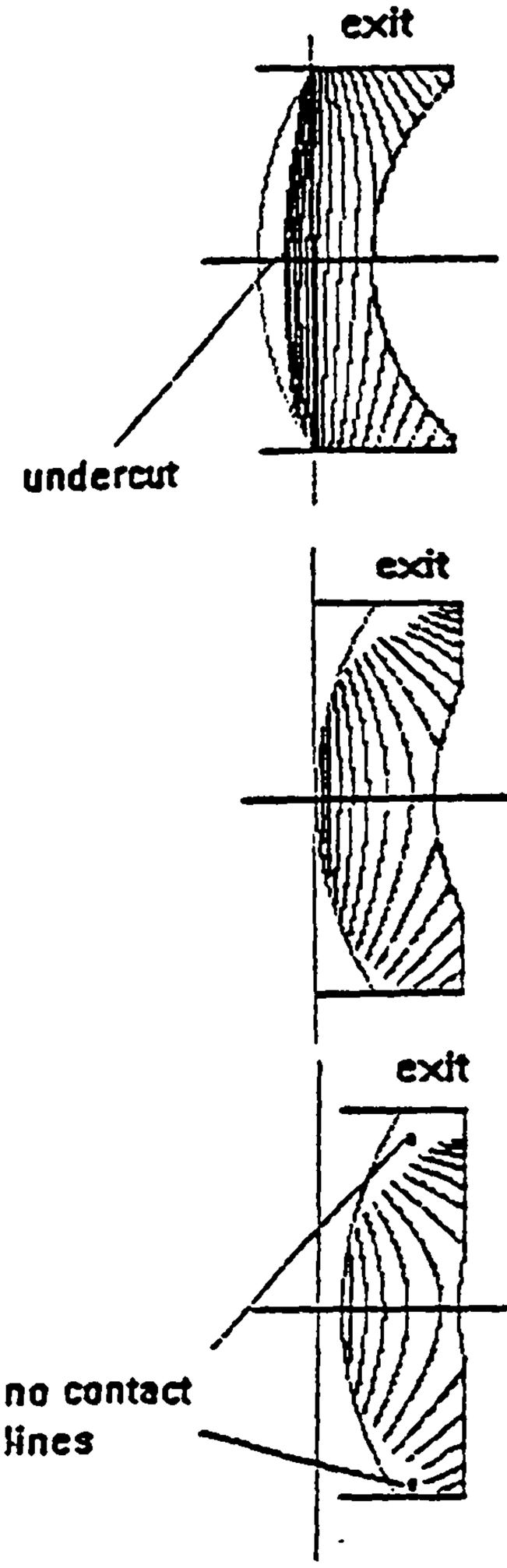


Figure 7.2 Contact lines on ZC work wheel flanks

occurs in this region.

The border line between the conjugate contact zone and the interference zone has often been discovered to corresponding closely to the hob tip (during cutting) or to the worm tip (during running). Here, there will be a blunt `corner' between the two surfaces, so that the contact stresses near these border lines could be very high. The software also shows that the exit side normally has a larger interference zone than the entry side, as well as a smaller gap than the entry side. On Cavex worms, intense localized contacts on the `corner' between the conjugate and interference surfaces is thus much more likely on the exit side. This type of localized `corner' contact rarely occurs on ZI worms because there is normally no interference zone. The sensitivity studies also show that the likelihood of this type of failure on Cavex worms will probably be further increased if close centre-distance tolerances cannot be held under running conditions.

These conclusions are supported by results from other publication [Winter, 1981] and by real running examples. We must recognise that the current design practice for the ZC worm gears produced by Flender may include techniques to eliminate or reduce the interference. However, as a research activity, we issue this preliminary discovery for discussion and further investigation if necessary.

7.3 Diagnosis of Design Errors

Design of worm gears is sophisticated and some design problems prove to be very difficult to solve without theoretical analysis. The following two examples are of problems which occurred in industry with which the Design Unit was asked to help. These two examples demonstrate how useful the analysis software developed is for solving industrial problems.

Case1: Cross contact

It is often found in manufacturing that a the contact pattern is split into two patches and

a damaging contact pattern appears at the wheel flank tip corner, as shown in Figure 7.3. This type of contact pattern is called "cross contact". A hobbing machine operator and engineer could spend hours- even days-trying to get rid of the corner patch and end up without success. A cross contact caused by the design shown in Figure 7.3 was diagnosed on site by the author by using the analysis software running on PC. The solution was found within a few minutes although the client had been struggling for days to cure the cross contact by trying different settings without being aware that the problem was inevitable due to a design error.

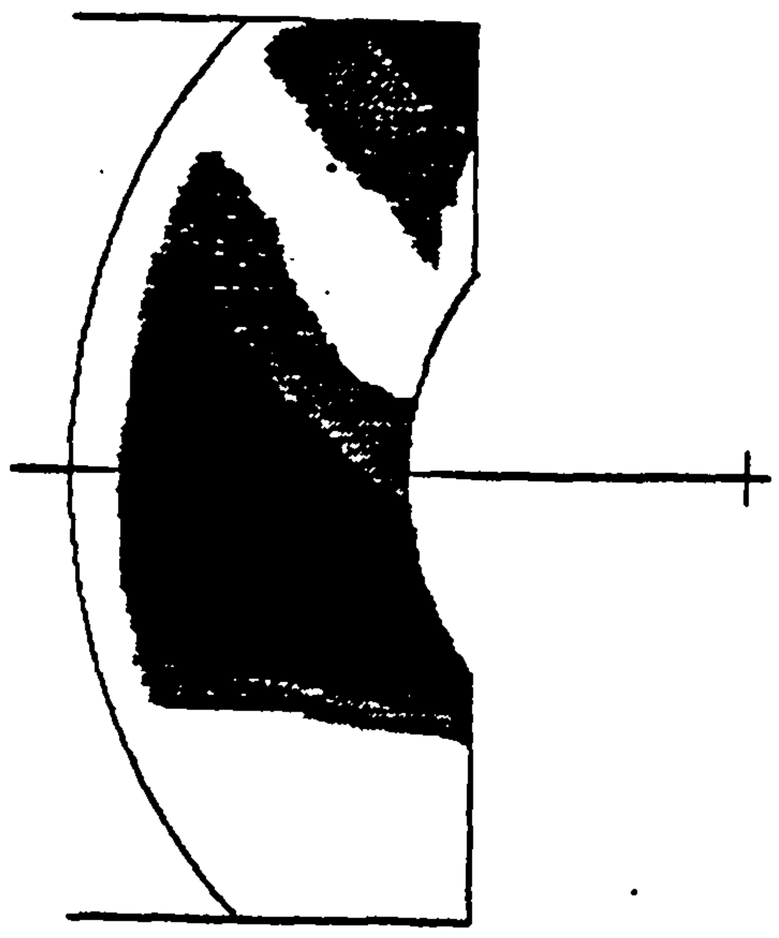


Figure 7.3 Cross contact on a worm wheel flank

Case 2: First surface interference

There are two possibilities of causing an area on the wheel dedendum to have no contact with the worm flank. The first possibility, unacceptable undercut, is well known and can be approximately estimated by simple calculations. But the second one, here called first surface interference, as shown in Figure 7.4, is more difficult. As can be seen, the contact lines will never extend into the dotted area. Actually, curve ABA is the limit of the conjugate contact surface on the wheel flank and often called "first boundary". This case could be extremely confusing since simple calculations suggest that there is no undercut, but increasing the hob size or reducing the hobbing centre distance makes no difference. The dotted area is simply "removed" by the hob rather than being generated. This is immediately shown up by the software.

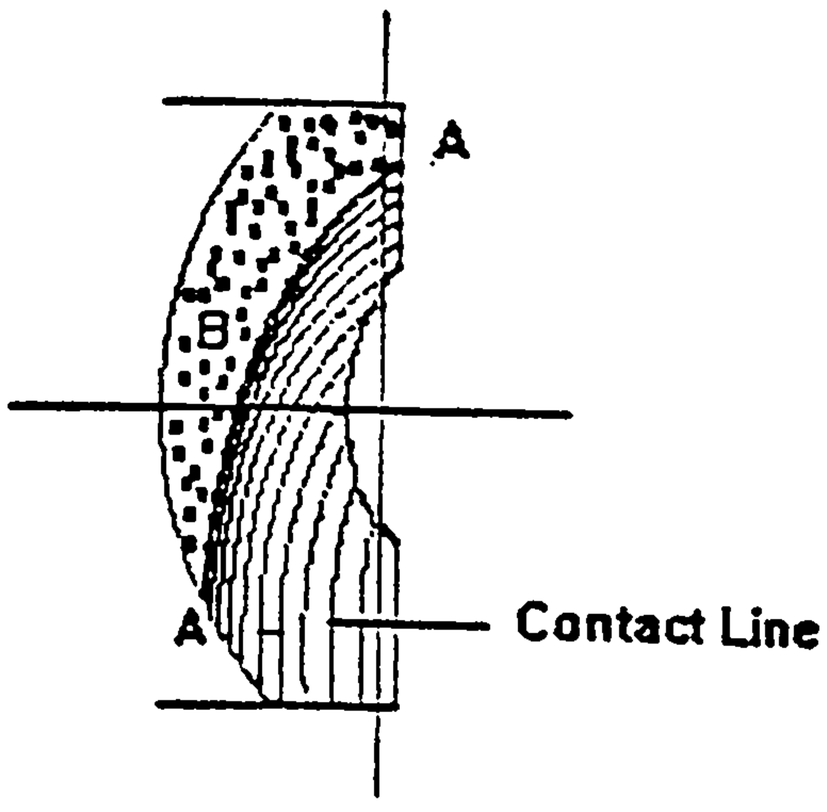


Figure 7.4 First surface interference

7.4 Collaboration with other worm gear research projects in the UK

The analysis software has also been used to provide data for the project "Worm Gear Design: Elatohydrodynamics" carried out at the University of Wales Cardiff (EPSRC Award GR/H 41546) and the worm gear metrology software has been used to measure test worm wheels for the DTI-founded project "The Analysis and Measurement of Transmission Error in Precision Worm Gears" carried out at University of Huddersfield. The collaboration with the University of Huddersfield and the University of Wales Cardiff has been very successful and mutually beneficial. For details, see [Snidle, 1996] and [Fish and Munro, 1995].

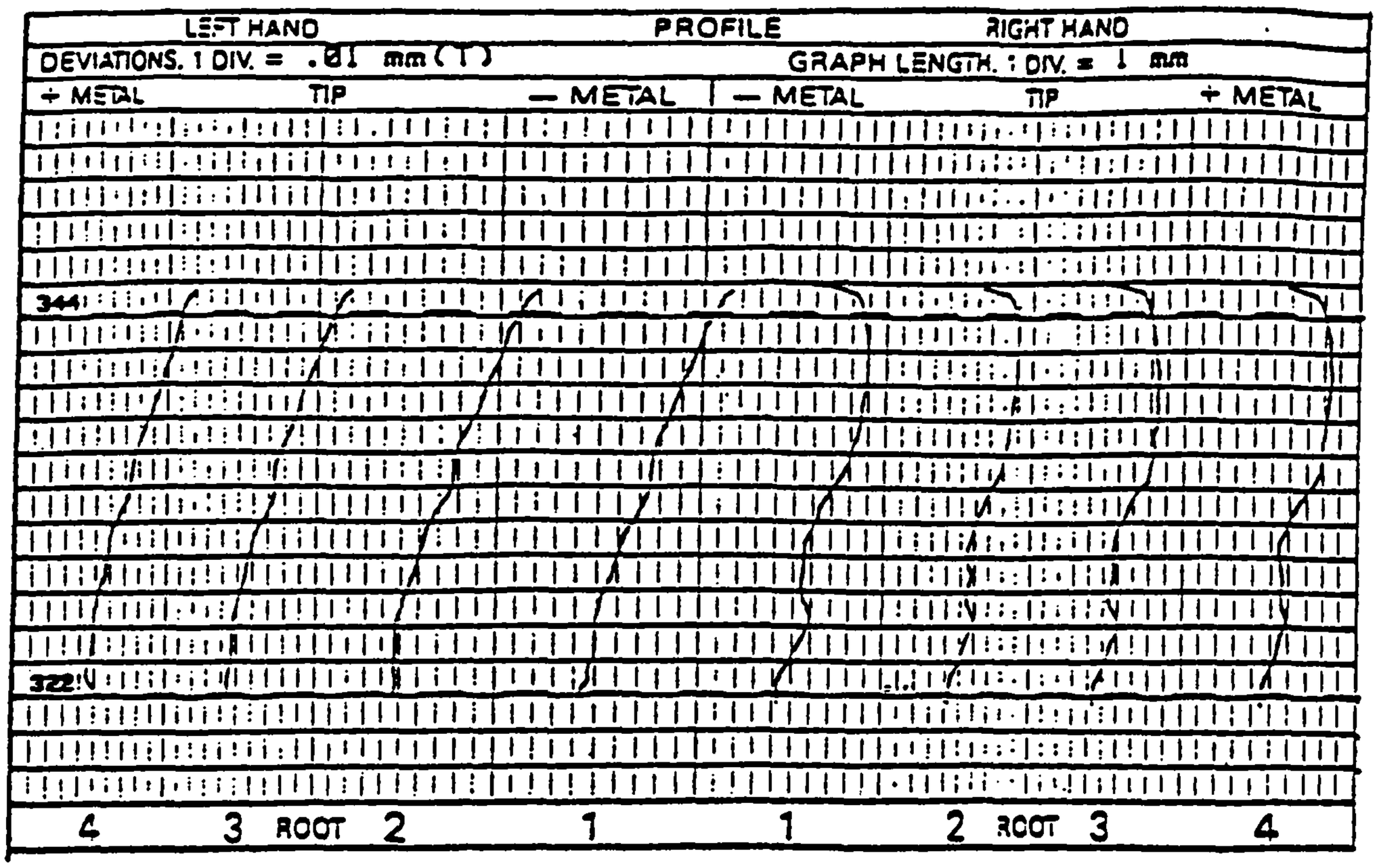


Figure 7.5 The worm wheel profiles before improvement

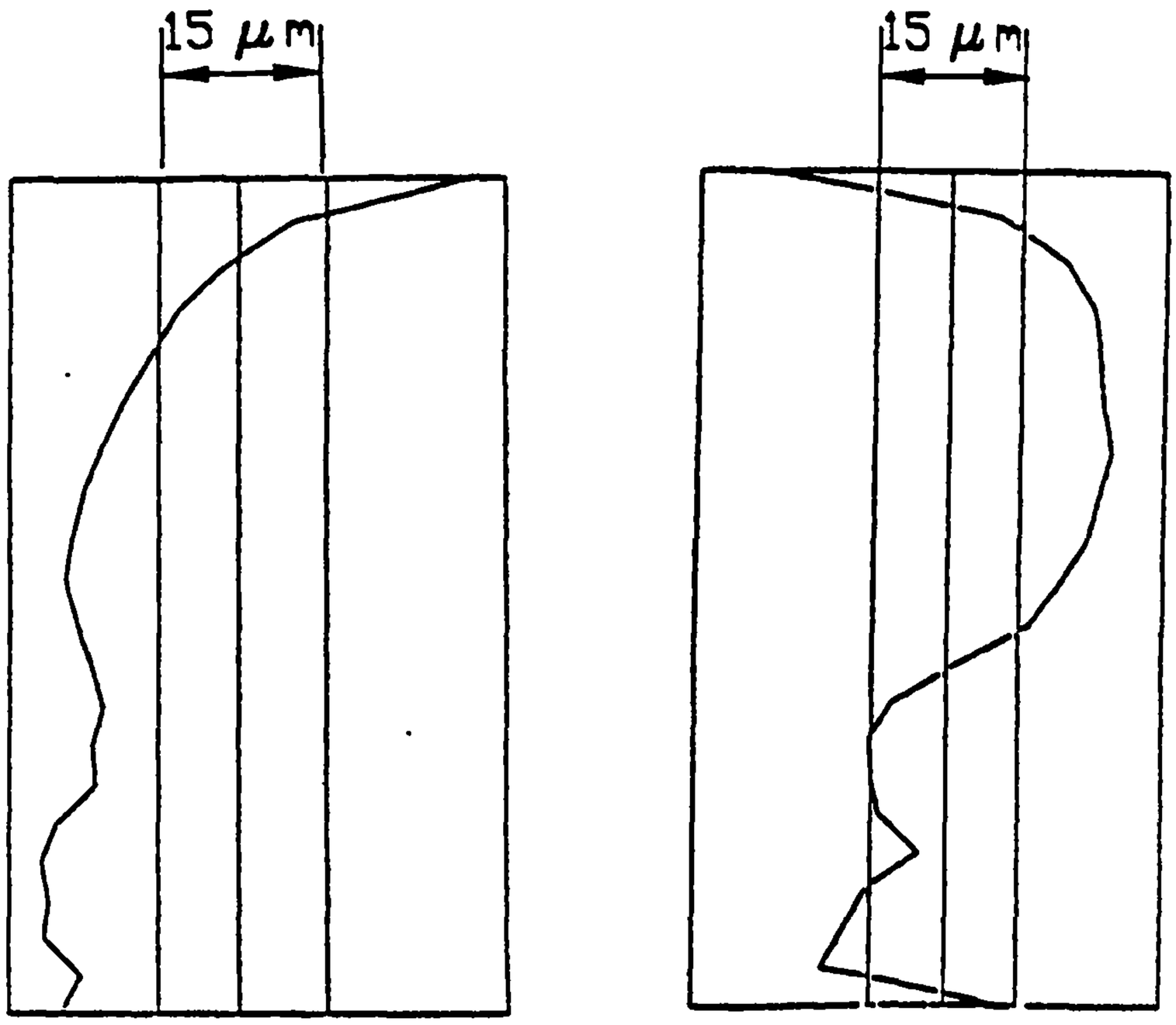


Figure 7.6 The hob profile before improvement

7.5 Improving manufacturing worm wheel quality [Hu and Frazer, 1996]

Design Unit was asked to check the accuracy of a worm, worm wheel and worm wheel hob from a passenger lift worm gear drive. The aim was to find the reasons why the expected contact pattern could not be achieved. Large profile errors were measured on the wheel and hob as shown in Figures 7.5 and 7.6. The similarity of the profile errors on of hob and wheel indicated that the hob profile error had probably been introduced into the worm wheel during hobbing. Then this was clarified by the analysis software which was used to predict the contact pattern with the actual hob and worm geometry measured. The prediction is shown in Figure 7.7. The prediction was checked against the actual contact pattern by the machine operator with very good agreement. Figure 7.8 shows the contact marking assuming there were no errors in the worm and wheel, which verifies that the design of hob and worm was satisfactory.

The conclusion was drawn that it was a tooling error, not a fundamental design error, so the Design Unit suggested that the quality of the hob should be improved to reduce the wheel profile errors.

The hob was then reground and used to cut a worm wheel, and the reground hob and worm wheel cut by it were again sent to Design Unit for measurement. The measurement results show a significant improvement in hob and wheel accuracy and the contact pattern was significantly improved. The improved results are demonstrated by Figure 7.9.

This type of diagnostic work would not have been possible without the software and measurement techniques developed during this project.

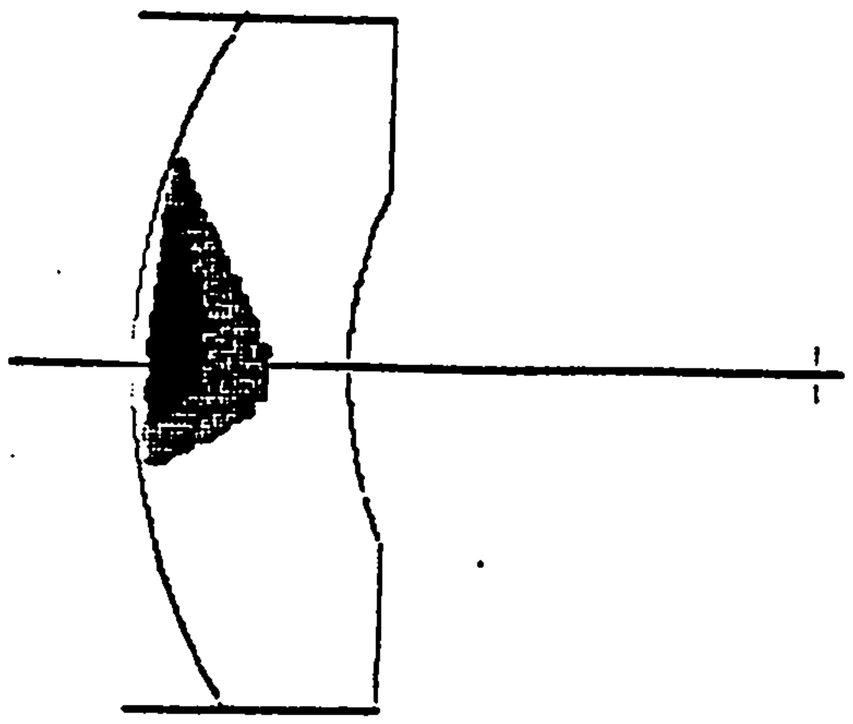


Figure 7.7 Contact pattern predicted with actual hob profile

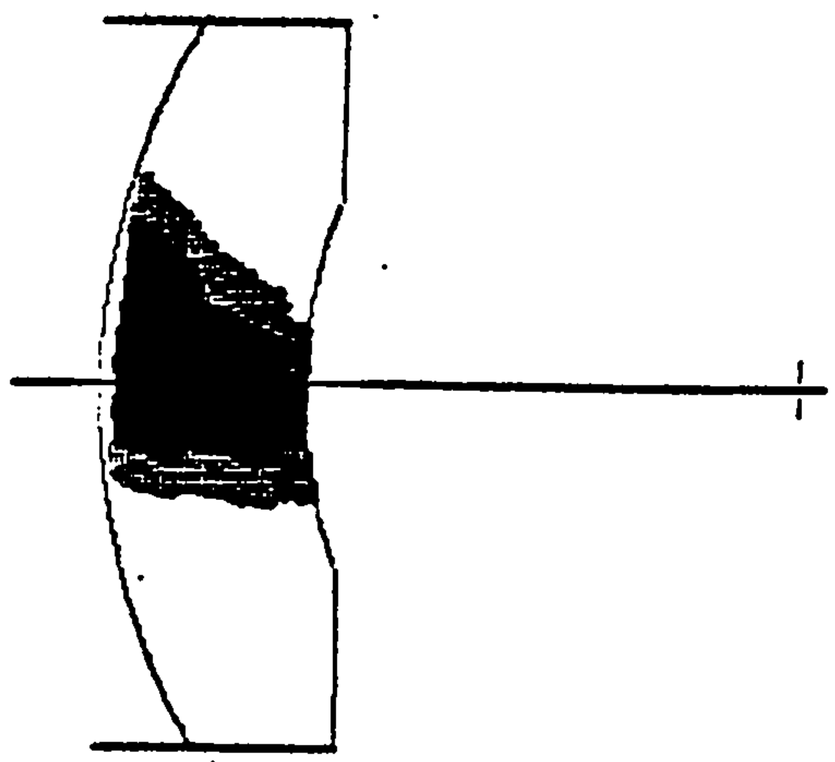


Figure 7.8 Contact pattern predicted with hob profile specified on drawing

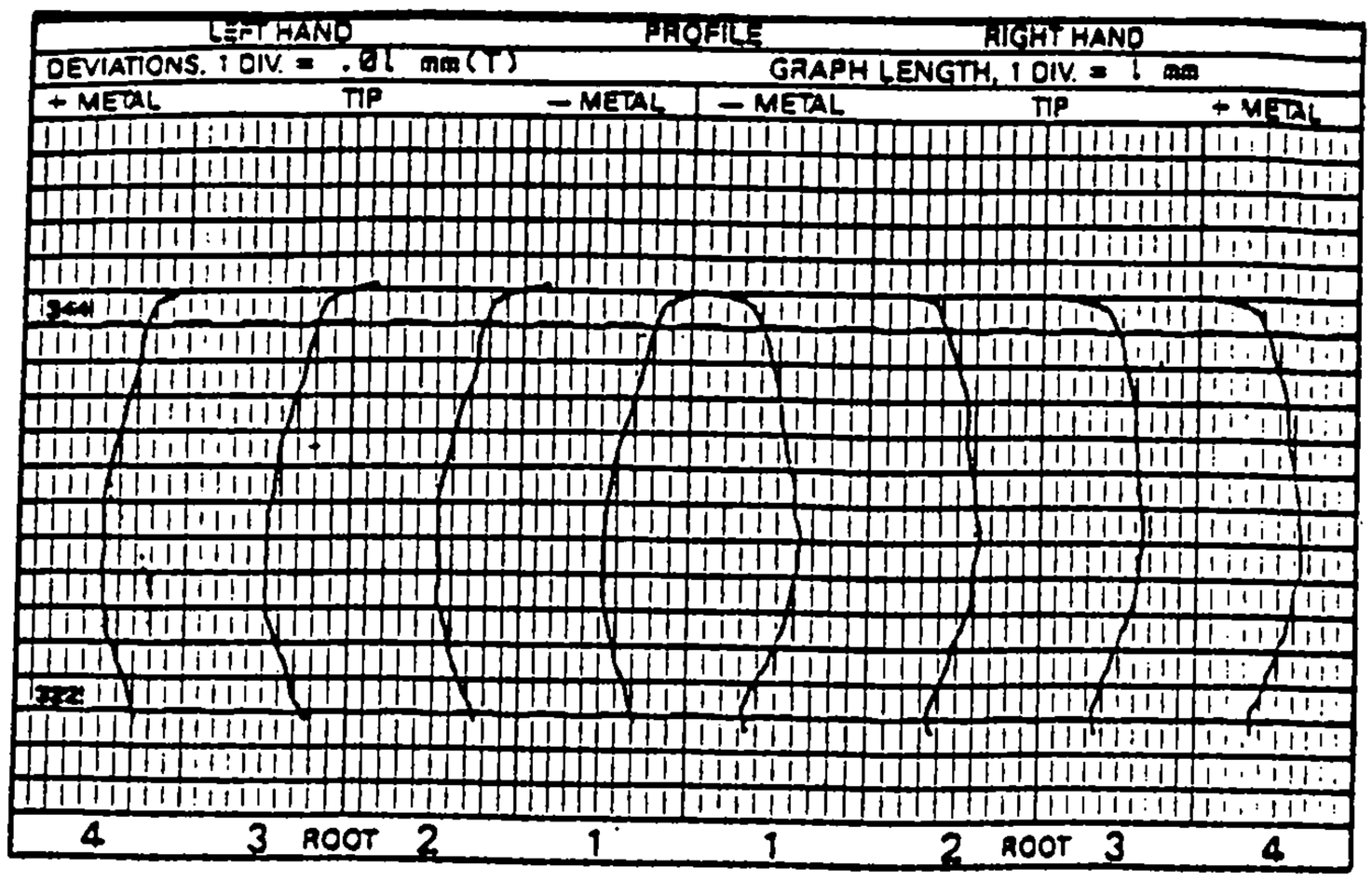


Figure 7.9 Hob profile after improvement

7.6 Reducing noise level of worm boxes [Frazer and Hu, 1995]

A worm gear set (designation 1/40/12.8/1.25 mm, 33mm CD) used in a stair lift produced unacceptable noise when running under load. The Design Unit was asked to determine the cause of this problem by measuring worm and wheel samples. The client reported that their prototype worms and wheels ran with acceptable noise level but that the production batch had proved to be excessively noisy. The worms and wheels supplied for measurement included both prototype and production examples. A gearbox assembly was also supplied to assist with the investigation.

The worms and wheels were measured. Profile measurement of the worm wheels showed 20 to 30 μm errors with (+) metal at the tip. This is a large error for a fine pitch worm gear and will cause significant mismatch between the worm gear pair which will generate noise. An example of the wheel measurement results is given in Figure 7.10. When the quiet worm was measured, a considerable amount of wear (15 ~ 20 μm) was found on the dedendum, while the noisy worm profile errors were within 5 μm (Figure 7.11).

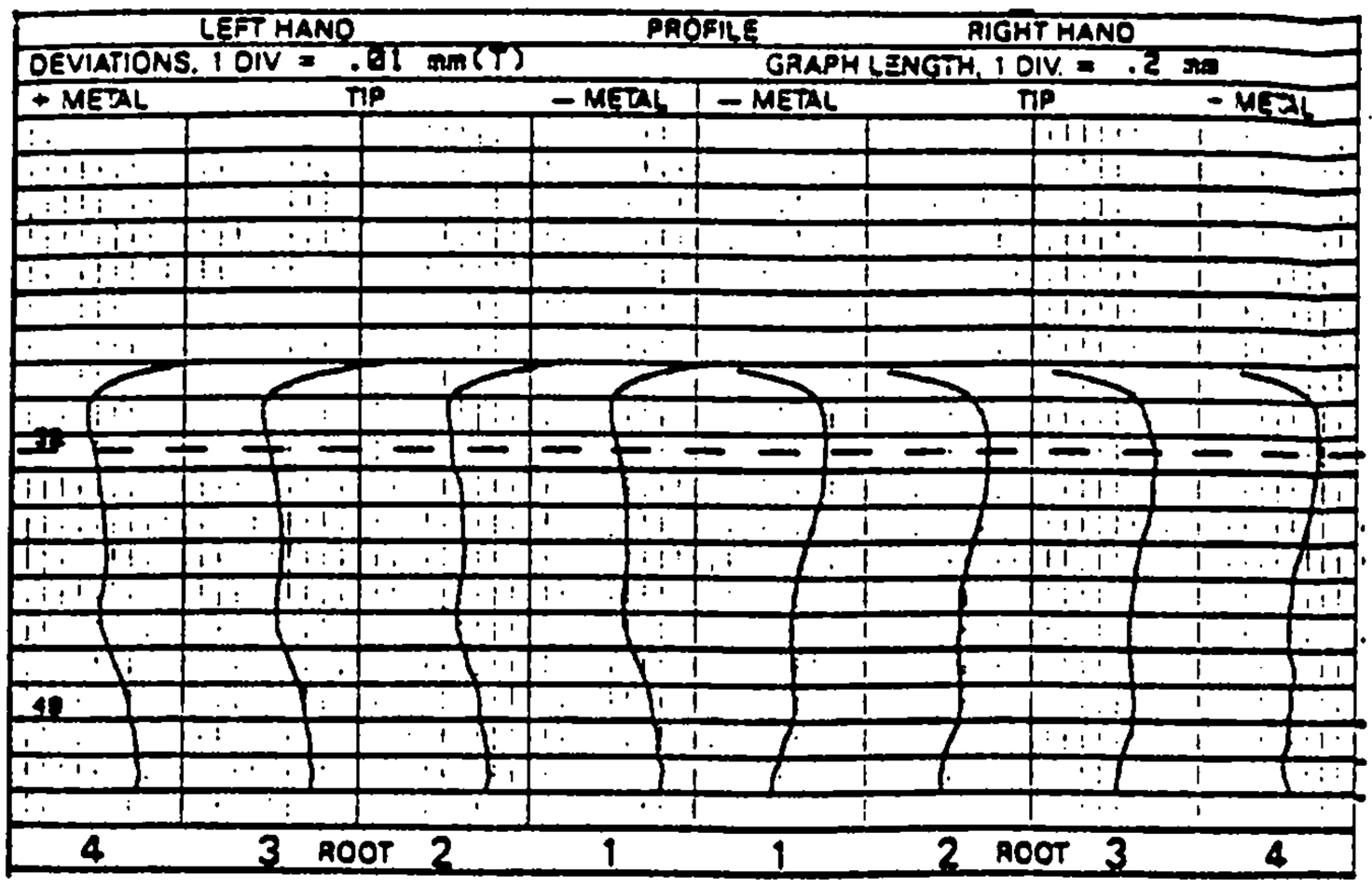
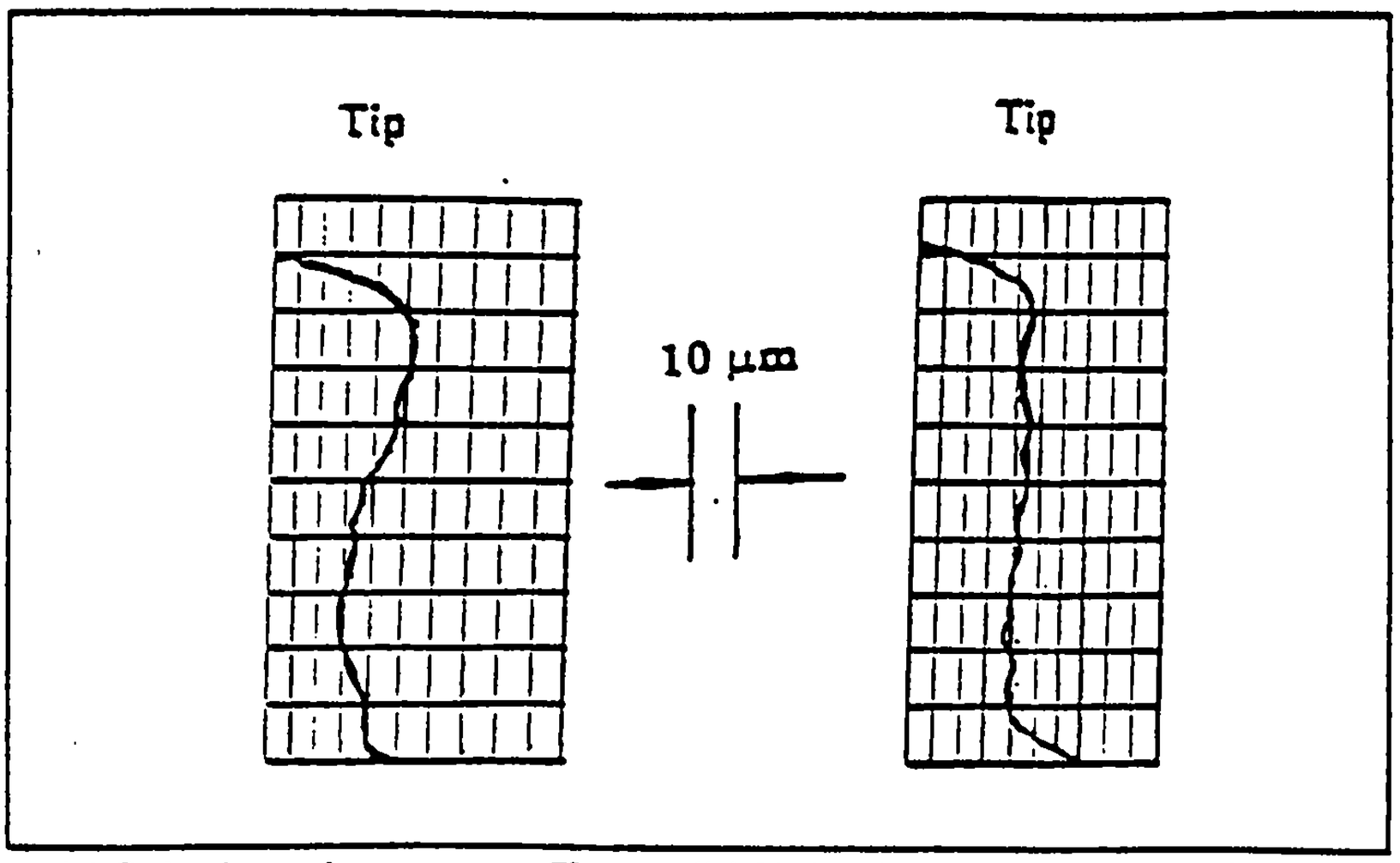


Figure 7.10 The wheel measurement results for example in section 7.6



(a) Quiet (worn) worm profile (b) Noisy (unworn) worm profile

Figure 7.11 Comparison between quiet and noisy worm profiles

The worm profile measurements suggest strongly that the profile of the prototype (quiet) worm was regenerated by the mating wheel. The wear on the worm dedendum had compensated for the mismatch due to the wheel profile error and helped to reduce the noise level. This diagnosis was confirmed by building an assembly comprising the prototype (worn) worm and a production wheel which produced a quiet gear box. The hardness of the worms was then tested and found to be only about 300 HV which is low for an induction hardened worm (specified heat treatment). This may explain the wear measured on the quiet worm.

The recommendation for improving the production quality was that the profile of the worm and wheel should be modified to get a smooth transmission and reduce the noise level. These recommendations have been proved to be very successful.

7.7 Lubrication and contact stress analysis

As mentioned above, the analysis software has been used to supply data to the Cardiff University for lubrication and contact stress analysis. One of the sample wheels

(designation 1/40/8/6.67 mm and 160 mm CD) is cut by a typical oversize hob and the no-load contact between the mating teeth is a point contact. The EHL analysis [Snidle and Evans, 1996] showed that the maximum contact pressure is 0.9 GPa when 75% of the catalogue rating load is applied to this sample. The materials of the sample are steel/bronze so a 0.9 GPa contact pressure is more than twice the maximum allowable value.

1) These results have well explained the reason why some companies start bedding-in load with a quarter of its catalogue rating and then increase the load gradually [Cone Drive, 1995].

2) The results have also convinced us that the elastic analysis of the unworn profiles for is not the best design procedure for hard/soft worm gears (see chapter 8).

CHAPTER 8

CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK

8.1 Conclusions

8.1.1 All the planned objectives have already been achieved. In addition, the results of this work have been used in all major worm gear research projects carried out in UK universities over recent years and have been a crucial contribution to most of them [Snidle and Evans, 1996]. Furthermore, the results have been extensively used for commercial purposes to investigate and improve worm gear design and manufacture (see chapter 7).

8.1.2 A new analytical mechanics of gearing - the B-Matrix method has been fully developed and successfully applied to worm gearing. The B-Matrix method has been proved to be advantageous in fully utilizing modern computers. It is a very compact, direct, accommodating and generalised 3D kinematic analysis technique, compared with polynomial models. These are tedious to develop and program and prone to errors, but are nevertheless the best solution for a slow computer. The author has found the B-matrix models very quick to develop and believes that the advantages of this method will become more obvious as computers develop.

In this project, the B-matrix method has been successfully implemented in software for non-elastic (no load) analysis of cylindrical worm gearing. The software has been written to be very user friendly with a GUI (Graphics User Interface) and is fully functional and robust.

8.1.3 The worm wheel measurement principle and associated mathematical models originated by the author have been proved to be a great success, and all the known theoretical and practical problems of implementation have been overcome without

sacrifice of measurement capability.

This measurement software has been implemented on a CNC measuring machine in the UK National Gear Metrology Laboratory, to allow, for the first time, analytical measurement of worm wheel tooth flanks. As far as we are aware, this worm wheel absolute measurement technology is unique to the UK National Gear Metrology Laboratory, although some leading CMM manufacturers (Tesa Metrology, Höfler, M&M) announced many years ago they were developing techniques for measuring worm wheels.

8.1.4 The combination of the mesh analysis software and CNC measurement of worm wheels has been proved to be a powerful tool in the investigation of industrial worm gearing problems in manufacture, quality control and design (see chapter 7). These applications of results from this project have shown clearly that this type of investigation and diagnostic work would have been virtually impossible without the availability of the analysis software and measurement technique.

8.2 Discoveries

8.2.1 The axial datum for worm wheel measurement is very critical to the lead measurement results. When measuring lead errors, a given error in axial datum will result in approximately 40% error in lead measurement. The profile measurement result is much less sensitive to the axial datum. A 50 μm error in axial datum hardly makes any difference in the profile measurement results.

The effect of axial datum on analytical measurement results has well explained the reason why the oversize hob technology is a great help in terms of reducing the sensitivity of worm gearing to machining and setting errors.

8.2.2 Wear of a wheel tooth normally extends over the entire flank, and even a few hours running at the nominal load rating can produce wear of 80 μm on a wheel flank; The worm measurement results have also confirmed that wear on worm flanks can be significant and is not negligible. Wear on a worm flank of up to 30 to 50 μm has been measured on many samples. The location of heavy wear on a worm flank is normally on its dedendum. However, a "perfect" worm flank has been widely assumed in previous worm gearing analysis programs. This is certainly at variance with what actually happens and has been misleading, since it has been found that worm flank deviations from ideal form can significantly affect worm gear performance, for instance its noise level.

8.2.3 A discrepancy has been found between the wear distribution and the contact marking predicted from lead traces measured after running. This is likely to be associated with elastic deflections of the worm, wheel, and worm gearbox assembly under load.

8.2.4 No-load contact analysis can be used to predict the contact pattern and no load performance of wormgears. Theoretical lead deviations (predicted entry / exit gaps) also show how the contact region will develop during bedding-in.

8.2.5 Cross contact tends to appear when hob oversize is reduced as a result of repeated resharpening, but this phenomenon may also occur with large oversize. A hob with 7% oversize produced very bad cross contact.

The contact marking can never go beyond the limits of the conjugate region, which can prevent achievement of the contact pattern wanted (see section 7.3).

Therefore, the worm gearing analysis software should be used to identify these problems when a new design is proposed, and certainly before manufacturing it.

8.2.6 Elastic "point contact" analysis is not meaningful for the analysis of worm gearing under load, since bedding-in will soon more or less conform the contact and substantially alter both the worm and wheel tooth form.

TCA was developed originally by Gleason for spiral bevel and hypoid gears which do have point contact between hard-hard surfaces. In this case, bedding-in wear is negligible and nominal point contact is maintained throughout the working life of these gears. Therefore it is not sensible to simply adopt this method for worm gearing analysis for soft worm wheels.

8.2.7 If the extent of wear or bedding-in is unknown, contact stress analysis and lubrication analysis remain uncertain. The reliable solution is to measure analytically the bedded-in flanks and carry out loaded analyses based on this geometry. Alternatively, an attempt must be made to develop an adequate wear model so that the bedded-in geometry can be predicted.

However, generally speaking, line contacts are probably closer to the actual situation than point contacts, since worm wheels normally show almost full face contact after bedding-in. Analytical measurements have shown that the bedded-in wheel flank is more or less conformal with the mating worm flank.

A loaded and lubrication analysis based on assumption of conformal contact may thus be more appropriate. This suggests that Wilkesmann's approximate loaded analysis

(which is the basis of the new draft worm gear load capacity rating procedure DIN 3996) is likely to be more reliable than procedures based on "exact" elastic analysis of the theoretical (unworn) profiles.

8.2.8 The effects of individual manufacturing errors on a worm wheel flank varies. The profile errors of a fly tool can be transferred directly onto the wheel flank. The link between hob profile errors and wheel geometry can also be seen from measurement results, but the interpretation of these results is much more difficult. More work to clarify the effect of hob manufacturing and mounting errors still needs to be done. The presentation of the effective cutting edge errors of a hob is certainly helpful.

The effects on the lead error of errors in axial datum and swing angle appear to be linear. The main effect on lead error of a centre distance error is to change the curvature of the lead trace and hence the entry/exit gaps. The combined effect of the three parameters on lead can be simply added together and successfully predicted by the software.

8.2.9 There appears to be a conflict between designing (a) to achieve favourable orientation of the contact lines for effective lubrication and (b) to prevent undercut.

8.2.10 ZC worm profiles may provide conformal contact and better lubrication. The price to pay for achieving these advantages is higher accuracy requirements (or a longer bedding-in procedure) and a relatively large positive addendum modification which tends to cause surface interference (see sections 7.1 and 7.2).

8.3 Discussion

During bedding-in, the wheel flank is regenerated by the worm flank, so that it becomes gradually conformal with the worm flank. Although it is very difficult to tell when the bedding-in stops precisely, point contact certainly no longer occurs after bedding-in.

For successful bedding in, it can be assumed that

1) The bedding-in wear procedure will continue until the contact stress has reduced to allow the oil film to build up sufficiently to separate the worm/wheel flanks under the transmitted load.

2) The contact should become almost full-face when worm gears are operated under full load, but should not be completely full-face, since the contact area may need to increase further under peak loads. The contact area nevertheless should not be a small area under full load, as compared with "almost full-face" contact, since this means the design is probably too conservative.

Therefore, it is necessary to predict the contact region after bedding-in. The contact area extends during bedding-in process, and clearly, will first extend to where the gap is the smallest, so that lead traces from theoretical analysis can be used to predict how the contact area will change during bedding-in. The prediction of the contact region after bedding-in can also be demonstrated from the predicted contact pattern plotted by gradually increasing the gap size (equivalent to gradually increasing the thickness of blue dye).

8.4 Suggestions for further work

8.4.1 The absolute measurement of involute worm wheels is a break through in the field, but it is only a start. More experience and further improvement are needed to take full advantage of involute worm wheel measurement;

8.4.2 The metrology software should be validated for ZA, ZN worm wheels and extended for other types of cylindrical worm wheels (for example, Cavex), as necessary.

8.4.3 The mesh analysis software should be modified to include functions to analyse hob runout and bearing errors, and hob "effective" profile and pitch errors. This will enable the complex interaction of these errors to be studied, so that the interpretation of analytical measurements of worm wheels can be improved.

8.4.4 Since these are still unexplained discrepancies between the calculated and actual worm wheel tooth forms, even when the kinematic hobbing errors have been minimized, the effect of hob/fly-tool deflection due to the cutting loads should be investigated. Because of the small hob/arbor diameter, these deflections are a much greater problem in worm wheel manufacture than in the conventional hobbing of spur/helical gears. On CNC machines, some measure of automatic compensation for these deflections may even be possible.

8.4.5 The effect of the original manufacturing errors on the wear process during bedding-in is still unknown and worthy of investigation, so as to be able to cut worm wheels to close to their bedded-in profile. This would greatly reduce the amount of wear and lubricant contamination that presently occurs.

8.4.6 It is very worthwhile considering the possibility of using hard worm wheels, since wheel manufacture with high accuracy should be achievable, so that the investigation can be based on known geometry and proper theoretical and experimental analyses. Although attempts to use hard/hard worm/wheel pairings failed in the past, the causes of these failures are still not clear. One of the reasons is probably that the actual wheel geometry used was unknown.

REFERENCES AND BIBLIOGRAPHY

Anand, S. Corporation, X. Houser, D. and Vijayakar, S. Study of Effect of Machining Parameters on Performance of Worm Gears, AGMA Technical Paper, 95FTM14, 1995.

ANSI/AGMA6022-C93, Design Manual for Cylindrical Wormgearing, ANSI/AGMA, 1993.

ANSI/AGMA 6034-B92, Practice for Enclosed Cylindrical Wormgear Speed Reducers and Gear Motors, ANSI/AGMA, 1992.

ANSI/AGMA 6030-C87, Design of Industrial Double-enveloping Wormgears, ANSI/AGMA, 1987.

Anon-1, Gear Design - Manufacturing and Inspection Manual, SAE(Society of Automotive Engineers) Gear and Spline Technology Committee, USA, 1990.

Anon-2, The Gearing Industry in the UK: Is It Healthy? Engg. Science Div. Gearing Working Party Report. Inst. Mech. Engrs. 1984.

Anon-3, Shougang Double-Enveloping Worm Gear Reducer Test (Model SG-71, 100 mm C.D., 10:1 Ratio), Test Report #545, Cone Drive Ltd., USA, December 12, 1995.

Anon-4, Gear Seminar, Tesa Metrology, September, Bedford, UK, 1993.

Anon-5, Tooth Damage on Worm Gears (Cavex, 686 mm C.D., 14.75 Ratio, Power Rating 170 KW, Speed n1=1000 RPM), Technical Information Sheet 1/77, GG-K23 D/EN 1.84, Flender Bocholt, German.

Anon-6, Angus Gear Book, 1st Edition, George Angus & Co. Ltd. Newcastle & London, 1959.

Anon-6, The Competitive Position of the Gearing Manufacturing Industry in the United Kingdom, Gearing Working Party Report, I. Mech. E. London, 1984.

Bagci, C. A New Theory for the Design of Helical Gear for Surface Fatigue Using Hertzian Point Contact stress, AGMA Technical Paper, 87 FTM1, 1987.

Barrett, C. The Current Status of Heavy-duty Open Gear Drive Lubrication, IEEE Transactions on Industry Applications, 1996, Vol.32, No.3, Pp.678-688

Beitz, W. & Küttner, K. H. Handbook of Mechanical Engineering, Springer-Verlag London Ltd. 1994.

Benthake, K. Tooth Damage on Worm Gears, Technical Information Sheet, GG-K 23 D/EN 1.84, Flender Bocholt.

Beyer, I. W. Traceable Calibration of Master Gears at PTB, AGMA Technical Paper, 96FTM4, 1996.

BS721, Specification for Worm Gearing, Part 2, BSI, London, 1983.

Buckingham, E. Analytical Mechanics of Gears, McGraw-Hill, New York 1949. (Reprinted by Dover, New York, 1963.)

Carnahan, B. Luther, H. A. and Wilkes, J. O. Applied numerical methods, John Wiley & Sons, Inc. 1969.

Colbourne, J. R. The Use of Oversize Hobs to Cut Worm Gears, AGMA Technical Paper, 89 FTM 8, AGMA, November 1989.

DIN 3975, Terms and Definitions for Cylindrical Worm Gears with Shaft Angle 90°, Beuth, October 1975.

DIN 3960, Concepts and Parameters Associated with Cylindrical Gears and Cylindrical Gear Pairs with Involute Teeth, Beuth, July 1980.

DIN 3962, Tolerance for spur and helical gears, Beuth, August 1978.

Dowson, D. and Higginson, G. R. Elastohydrodynamic Lubrication, Pergamon Press, Ltd., 1966.

Dowson, D. Elastohydrodynamics, Proceedings of the Institution of Mechanical Engineers, Vol. 182, Part 3A, 1967-68, pp.151-167.

Dudley, D. W. Gear Handbook, 1st Edition, McGraw-Hill Book Company, New York, 1962.

Dudley, D. W. and Poritsky, H. Cutting and Hobbing Worms and Gears, 26th Semi-Annual Meeting of AGMA, Chicago, 1943, p241.

Dudley, D. W. and Poritsky, H. On Cutting and Hobbing Gears and Worms, Trans ASME, Chicago, 1943. pp A139-A146, A197-A201, A247-A251

Errichello, R. The Lubrication of Gears - Part II, Gear Technology, May/June 1991, pp18 - 22.

Evans, H. P. and Snidle, R. W. A simple, Accurate Method for Calculation of Stresses and Deformations in Elliptical Hertzian Contacts, Proceedings of Institution of Mechanical Engineers, Vol. 206, pp139 - 141, 1992.

Fang, H. S. and Tsay, C. B. Mathematical and Bearing Contact of the ZK-Type Worm Gear Set Cut by Oversize Hob Cutters, Mech. Mach. Theory, Vol. 31, No. 3,

pp271-282, 1996.

Fish, M. and Munro, R. G. Analysis of Marking Patterns and Transmission errors in Worm Gears, BGA Annual Congress, UK. 1995.

Fish, M. and Munro, R. G. Transmission Errors in Worm Gearing, DTI Committee Progress Reports on the Programme of Collaborative Research Funded by DTI (Ref: REH/LMS/600.1), University of Huddersfield, 1994 to 1997.

Frazer, R. C. and Hu, J. Worm Noise Investigation, Internal Report, GM301, Design Unit, University of Newcastle upon Tyne, July 1995.

Frazer, R. C. and Hu, J. The differences between 3-axis and 4-axis gear measuring machines, Proc. of British Gear Association Seminar, Huddersfield, October 1996.

Frazer, R. C. and Hu, J. Methods of Testing Calibration Equipment in the UK's National Gear Metrology Laboratory, Proceedings of the Third International Conference on Laser Metrology and Machine Performance, Huddersfield, UK, 1997

Frazer, R. C. and Hu, J. Verifying the Accuracy of Involute Gear measuring Machines, Proceedings of the Third International Conference on Laser Metrology and Machine Performance, Huddersfield, UK, 1997.

Frazer, R. C. and Myers, E. J. The Code of Practice for Gear Measurement, BGA Annual Congress, UK, 1993.

Frazer, R. C. and Myers, E. J. A Code of Practice for the Specification of Drawing Data and Associated Tolerances for Involute Gears, DUCOP01, BGA Transmission Technology, Issue 1.1, UK, September 1994.

Frazer, R. C. and Myers, E. J. A Code of Practice for the Specification of Gear Blanks for Involute Gears, DUCOP02, BGA Transmission Technology, Issue 1.1,

UK, September 1994.

Frazer, R. C. and Myers, E. J. A Code of Practice for Involute Gear Measurement, DUCOP03, BGA Transmission Technology, Issue 1.1, UK, September 1994.

Frazer, R. C. and Myers, E. J. A Code of Practice for Checking the Reproducibility of Gear Measuring Machines, DUCOP04, BGA Transmission Technology, Issue 1.1, UK, September 1994.

Frazer, R. C. and Myers, E. J. A Code of Practice for Verifying the Accuracy of Gear Measuring Machines: Part 1, DUCOP05/1, BGA Transmission Technology, Issue 1.1, UK, September 1994.

Frazer, R. C. and Myers, E. J. A Code of Practice for Verifying the Accuracy of Gear Measuring Machines: Part 2, DUCOP05/2, BGA Transmission Technology, Issue 1.1, UK, September 1994.

Frazer, R. C. and Myers, E. J. A Code of Practice for Verifying the Accuracy of Gear Measuring Machines: Part 3, DUCOP05/3 (In preparation ?), BGA Transmission Technology, Issue 1.1, UK, September 1994.

Frazer, R. C. and Myers, E. J. A Code of Practice for Interpretation of Measurements, DUCOP06, BGA Transmission Technology, Issue 1.1, UK, September 1994.

Frazer, R. C. and Myers, E. J. A Code of Practice for Gear Calibrations (In preparation), DUCOP07, BGA Transmission Technology, Issue 1.1, UK, September 1994.

Geotz, A. Introduction to Differential Geometry, Addison Wesley, Reading, 1970.

Gu, A. Elastohydrodynamic Lubrication of Involute Gears, Paper 72-PTG-34,

ASME, New York, N.Y. 1972.

Heading, J. Mathematical Methods in Science and Engineering, 2nd Edition, UK, 1979.

Hiltcher, Y. and Guingand, M. Computed Contact Surface Topology of Worm Gear - New Approach and Cutting Parameters Influences, 1994 International Gearing Conference, Newcastle upon Tyne, UK, September 1994.

Hofmann, D. A. Drive Systems, BGA Gear Technology Teaching Pack, Vol1. 1992.

Hofmann, D. A. Improving Worm Gear Performance - The Development of New Gear Geometries and Worm Gear Materials for High Load Carrying Capacity, Proposal for Collaborative Research Program Funded by DTI and GRF, University of Newcastle upon Tyne, UK 1993.

Höhn, B. R. and Steingröver,K. DIN 3996: A New Standard for Calculating the Load Capacity of Worm Gears, 96FTM11, AGMA Technical Paper, 1996.

Höhn, B. R. and Winter, H. DIN 3996: A New Standard for Calculating the Load Capacity of Worm Gears, BGA Annual Congress, UK, 1993.

Hu, J. Theory and Priciple of Worm Wheel Measurement, Internal Reports, Design Unit, Newcastle University, 1992.

Hu, J. Measurement Results of Worm Wheels, Internal Reports DU 1664, Design Unit, Newcastle University, 1992-1997.

Hu, J. and Bin H. Z. Automatic Inspection of Double-enveloping Worm Lead Deviation, Journal of Mechanical Engineering, ISSN 1001-2192, China, July 1989.

Hu, J. Bin H. Z. Yang, C. and Zhang, F. The Communication Between Z80-single

Board Computer and Encoder Display Unit, Journal of Flexible Machine Tool & Automatic Manufacturing Technique, ISSN 1001-2265, China, March 1987, pp9-12.

Hu, J. and Frazer, R. C. Investigation into a 53-2 Worm Gear Drive, Internal Reports GM 349 and DU 1664, Design Unit, Newcastle University, 1996.

Hu, J. Frazer, R. C. and Pennell, J. A. The Measurement of Involute Worm Wheels, Proceedings of the Third International Conference on Laser Metrology and Machine Performance, Huddersfield, UK, 1997.

Hu, J. and Pennell, J. A. A Practical Method of Measuring the Worm Wheels of Cylindrical Worm Gears, Proceedings of the 1994 International Gearing Conference, Mechanical Engineering Publication Limited, London, September 1994.

Hu, J. Pennell, J. A. and Hofmann. D. A. Improving WormGear Performance, DTI Committee Progress Reports on the Programme of Collaborative Research Funded by DTI (Ref: MP 20/113), DU1664, October 1994 to May 1997.

Hu, J. Zhang, F. and Bin,H. Z. The Mathematical Model for Helicoid of Double-Enveloping Worms with an Improved Kinematic Method, Journal of Mathematics in Practice and Theory, China, No.3. 1991.

Hua, D. Y. and Khonsari, M. M. Application of Transient Elastohydrodynamic Lubrication Analysis for Gear Transmissions, Journal of Tribology Transactions, 1995, Vol.38, No.4, pp.905-913

Janninck, W. L. Contact Surface Topology of Worm Gear Teeth, AGMA Technical Paper, 87 FTM 14, AGMA, USA, October 1987.

Janninck, W. L. Our Experts discuss Hobbing Ridges, Crooked Gear Teeth, and Crown shaving, Gear Technology, March/April 1992, pp.41-43.

Janninck, W. L. Cutting Worm Gears with Standard Gear hobs, Gear Technology, January/February 1994, pp.45-47.

Kin, V. Computerized Simulation of Meshing and Bearing Contact for Single-enveloping Worm-gear Drives, Ph.D Thesis, University of Illinois at Chicago, 1990.

Kin, V. Computerized Analysis of Gear Meshing Based on Coordinate Measurement Data, USA., September 1994. Vol. 116 p739.

Lauer, D. Alternative Lubrication Methods for Large Open Gear Drives, Lubrication Engineering, 1996, Vol.52, No.7, pp515-523

Litvin, F. L. Theory of Gearing(In Russian), 2nd Edition, Nauka, Moscow, 1968. (Translation, Shanghai Publishing House of Sci. And Tech. China, 1984.)

Litvin, F. L. Theory of Gearing, NASA. Reference Publication 1212, AVSCOM. Technical Report 88- C- 035, 1989.

Litvin, F. L. Gear Geometry and Applied Theory, PTR Prentice Hall, USA, 1994.

Litvin, F. L. Theory of Gearing and Applications, In Dudley's Gear Handbook (D. P. Townsend (Ed.)), McGraw Hill, 1992.

Litvin, F. L. and Kin, V. Computerized Simulation of Meshing and Bearing Contact for Single- Enveloping Worm-gear Drives, Journal of Mechanical Design, USA, 1992, Vol.114, No.2, pp313-316

Loveless, W. G. Barlow, R. J. and Greening, J. H. Technical Aspects of Double Enveloping Worm Gears, Ex-Cell-O Corporation, Cone Drive Operation, 1980.

Lu, Y. Qi, S. Hu, J. and Yang C.M. Study on Principle and Method of Computer-Aided Inspection of Double-enveloping Worms, Proceedings of 4th International Conference on CAPE., Edinburgh, UK, September 1988, pp225-231.

Maw, S. A Review of Worm Gear Rating Techniques, BGA Congress, UK 1995.

Merritt, H.E. Worm Gear Performance, Proc. I. of Mech. E., Vol 129, 1935. p127.

Merritt, H.E. Gear Engineering, Halsted Press, 1971.

Munro, R. G. A Review of the Theory and Measurement of Gear Transmission error, IMechE, C404/032, pp3-10, 1990.

Munro, R. G. and Ling, P. A New Method of Measuring Involute Profile Deviations of Gear Teeth, Part C - Journal of Mechanical Engineering Science, 1996, Vol. 210, p63.

Narayan, A. Corportation, X. Study of Effect of Machining Parameters on Performance of worm Gears, AGMA Technical Paper, 95FTM14, October 1995.

Octrue, M. A New Method of Designing Worm Gears, AGMA Paper, 1989.

Octure, M. Study of the Geometry of Cylindrical Worms and Wormwheels (In French), CETIM., Technical Note No.22

Opalic, M., Husnjak, M. and Goja, Z. Determination of Contact Pressure Distribution in Worm Gear Drives, Strojarstvo, 1989, Vol.31, No.4-6, pp225-228

Pennell, J. A. Improving the Stress Analysis of Worm Gearing, Proposal for Collaborative Research Program, University of Newcastle upon Tyne, UK, 1990.

Pennell, J. A. Development of Worm Gearing Analyses, Unpublished Report, Department of MMM, University of Newcastle upon Tyne, UK, 1995.

Pressman, R. S. Software Engineering - A practitioner's Approach, Third Edition, McGraw-Hill Inc. 1994.

Qin, D. Zhang, G. and Kato, M. Study on Hourglass Worm Gearing with Good Contact State and Insensitivity to Manufacturing Errors, Proceedings of the 1994 International Gearing Conference, Mechanical Engineering Publication Limited, London, September 1994.

Samman, N. Lau, Sn. A Novel-Approach to Enclosed Gear Lubrication on Board Mining Equipment, Lubrication Engineering, 1996, Vol.52, No.1, pp84-93

Shen, Y. F. and Rong, E. Q. The Theory of Spatial Meshing and SG-71 Double Enveloping Worm Gear Drives, 2nd Edition, Metallurgical Industry Publishing House, China, April 1983.

Sheng, C. Q. Lei, J. R. Savage, M. A General-method for Computing Worm Gear Conjugate Mesh Property .1: The Generating Surface, Journal of Mechanisms Transmissions and Automation in Design-transactions of the ASME, 1989, Vol.111, No.1, pp143-147

Sheng, C. Q. Lei, J. R. A General-method for Computing Worm Gear Conjugate Mesh Property .2: The Mathematic Model of Worm Gear Manufacturing and Working Processes, Journal of Mechanisms Transmissions and Automation in Design-transactions of the ASME, 1989, Vol.111, No.1, pp148-152

Sheng, X. Design of High Load Capacity Gears, Mechanical Industry Publishing House, Beijing, 1981.

Shimokohbe, A. Toyama, A. and Lin, Q. J. Improvement in Load Capacity of cross Helical Gears, ASME, New York, 1993.

Simon, V. A New Worm Gear Drive with Ground Double Arc Profile, Mech. Mach. Theory, Vol 29. No.3 pp407-414, 1994.

Snidle, R. W. and Evens, H. P. Worm Gear Design: Elastohydrodylamics, Final Report on EPSRC Research Grant GR/H 41546, University of Cardiff, 1996.

Snidle, R. W. and Evens, H. P. Novel Worm Gear Designs, Proposal for EPSRC Research Program (draft), University of Cardiff, 1997.

South, D. W. and Ewert, R. H. Encyclopaedic Dictionary of Gear and Gearing, McGraw-Hill Inc. New York, 1995.

Stirzaker, G. Measurement Results of Worm and Wheel Samples, Internal Report, Department of MMM, University of Newcastle upon Tyne, June, 1994.

Tang, J. S. The Hertzian Load-Capacity and Lubrication of A Double-enveloping Wormgearing with Planar Generating Surface (In Chinese), PhD Thesis, Mineral University of China, May 1991.

Tang, Z. Study on Improvement for Worm Gear Contact Accuracy (In Chinese), Journal of Gear, China, Vol. 8, No.1, 1984. pp 20-24.

Townsend, D. P. Elastohydrodynamic Principles Applied to the Design of Helicopter Components, Journal of the American Helicopter Society, July 1974, pp10-16.

Tsay, C. B., Jeng, J. W. and Feng, H. S. A Mathematical Model of the ZE-type Worm Gear Set, Mechanism & Machine Theory, Vol. 30, No.6, 1995. pp777-789.

Vos, H. Design Features Increase the Load Capacity of Worm Gears, Maschinenmarkt 102 (8), 1996, pp56-59.

Walker, H., Worm Gear Deflection (A Consideration of Distortion Effects and Correcting Measures), Automobile Engineering, 1945.

Watson, H. J. Modern Gear Production, 1st Edition, Pergamon Press, Oxford,

1970.

Weber, C. The Deformation of Loaded Gears and The Effects on Load Carrying Capacity, British Dept of Science and Industrial Research-Report No. 3, 1949.

Winter, H. and Wilkesmann, H. Calculation of Cylindrical Worm Gear Drives of Different Tooth Profiles, Journal of Mechanical Design, Trans. ASME, USA, January 1981. Vol.1

Wu, C. X. and Tang, Y. Y. Differential Geometry, High Education Publishing House, Beijing, 1985.

Wu, H. Y. and Su, D. Z. Equation of Meshing of ZK Worm Gearing, Journal of Mechanical Engineering, China, No.4 1982.

Wu, H. Y. and Su, D. Z. Study on the Load Capacity of ZK Worm Gearing, Proceedings of Annual Congress of Chinese Mechanical Engineers Association, China, 1983.

Wu, H. Y. Zhang, Y. X. and Qi, L. Design of Worm Gear Drives, Mechanical Engineering Industry Publishing House, China, 1st Edition, December 1986.

Yang, C. Qi, S. and Hu, J. A multi-stage and Multi-function System for Measuring Manufacturing Complex Surface, Proceedings of International Computer Application Conference, Hong Kong, May 1989.

Ye, K. M. et al. Gear Handbook (In Chinese), The Mechanical Engineering Publishing House, Beijing, 1990.

Yuan, Q., Sun, D. and Brewe, D. Effect of Worm Gear Geometry on its Contact Properties, Tribology Transactions, 1996, Vol.39, No.1, pp139-147

Zhang, F. and Hu, J. An Improved Kinematic Method, Proceedings of 5th International Conference on CAPE., Edinburgh, UK, November 1989, pp232-239.

Zheng, Z. and Hu, J. Study on the Control of Electrohydraulic Speed Servo-System, Proc. 1st National Conference on Computer Control, Shengzhen, China, April 1986, pp234-241.

APPENDIX 1

CONVENTIONS USED

A 1.1 Worm gear types: ZI (involute) worm gears unless otherwise stated.

A 1.2 Hand and flank of worm gearing considered: left flank of a right-handed worm and wheel is studied in this thesis, as shown in Figure A 1.1.

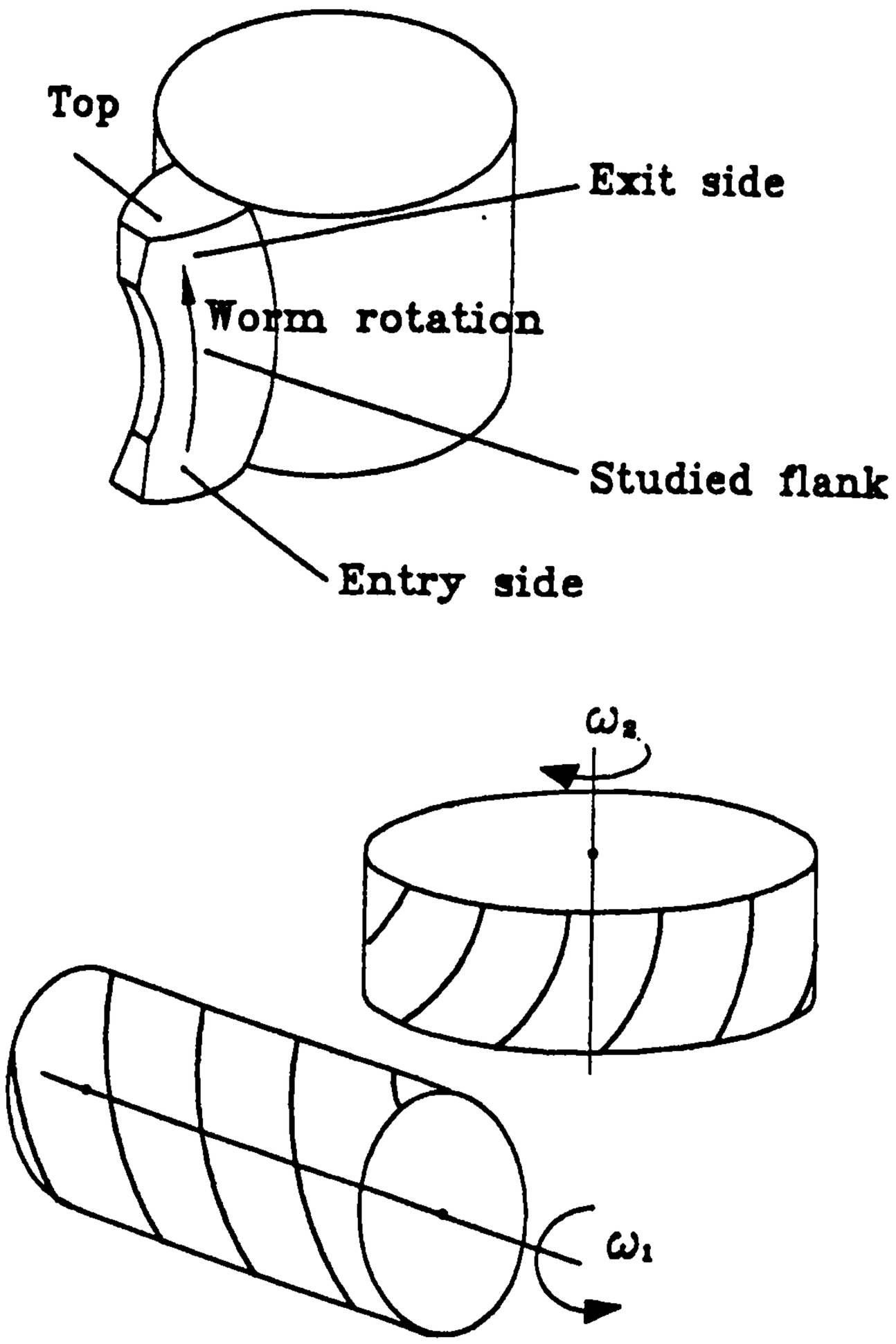


Figure A 1.1 Hand and flank of worm gearing considered

A 1.3 Surface conditions:

$$r^{(1)}(u,\theta) = \begin{pmatrix} x^{(1)}(u,\theta) \\ y^{(1)}(u,\theta) \\ z^{(1)}(u,\theta) \\ 1 \end{pmatrix}$$

$$r^{(1)}(u,\theta) \in C^2$$

$$\frac{\partial r^{(1)}}{\partial u} \times \frac{\partial r^{(1)}}{\partial \theta} \neq 0 \qquad (u, \theta) \in G$$

The superscript (1) denotes the worm surface. u and θ are curvilinear parameters. Symbol C^2 denotes that $x^{(1)}(u,\theta)$, $y^{(1)}(u,\theta)$, $z^{(1)}(u,\theta)$ have continuous derivatives to the second order. $(u,\theta) \in G$ denotes that the curvilinear parameters belong to the area G which is an open region (excluding part of its boundary).

$$r^{(2)}(u,\theta,\phi_1) = \begin{pmatrix} x^{(2)}(u,\theta,\phi_1) \\ y^{(2)}(u,\theta,\phi_1) \\ z^{(2)}(u,\theta,\phi_1) \\ 1 \end{pmatrix}$$

$$r^{(2)}(u,\theta,\phi_1) \in C^2 \qquad (u, \theta) \in G \qquad \phi_1 \in (a,b)$$

A 1.4 Rigidity: No elastic effects are included:

APPENDIX 2

WORM GEAR FOR THE EXAMPLES USED

All the following samples are ZI (involute) worm gears.

A 2.1 Data for validation of the B-matrix models and calculation procedures

Worm Data	C (mm)	z_1	z_2	m_x (mm)	q	α_n
	160	2	41	6.4	9	25^0

A 2.2 Data for validation of the non-conjugate analysis against other calculation methods

Worm Data	C (mm)	z_1	z_2	m_x (mm)	q	α_n
	152.4	1	50	5.1557	9.1267	22.5^0
Hob Data	C (mm)	z_1	η	m_x (mm)	q	α_n
	152.273	1	0.0667^0	5.1429	12.094	22.5^0

A 2.3 Data for validation of the non-conjugate analysis against other calculation methods

Worm Data	C (mm)	z_1	z_2	m_x (mm)	q	α_n
	203.2	2	53	6.1895	12.654	20^0
Hob Data	C (mm)	z_1	η	m_x (mm)	q	α_n
	206.16	2	0.7^0	6.1803	13.5678	20^0

Both the worm and hob flanks had been modified using parabolic profiles. The actual

(measured) profiles are simulated in the model by assuming the profile corrections shown in Figure A 2.1 and A 2.2.

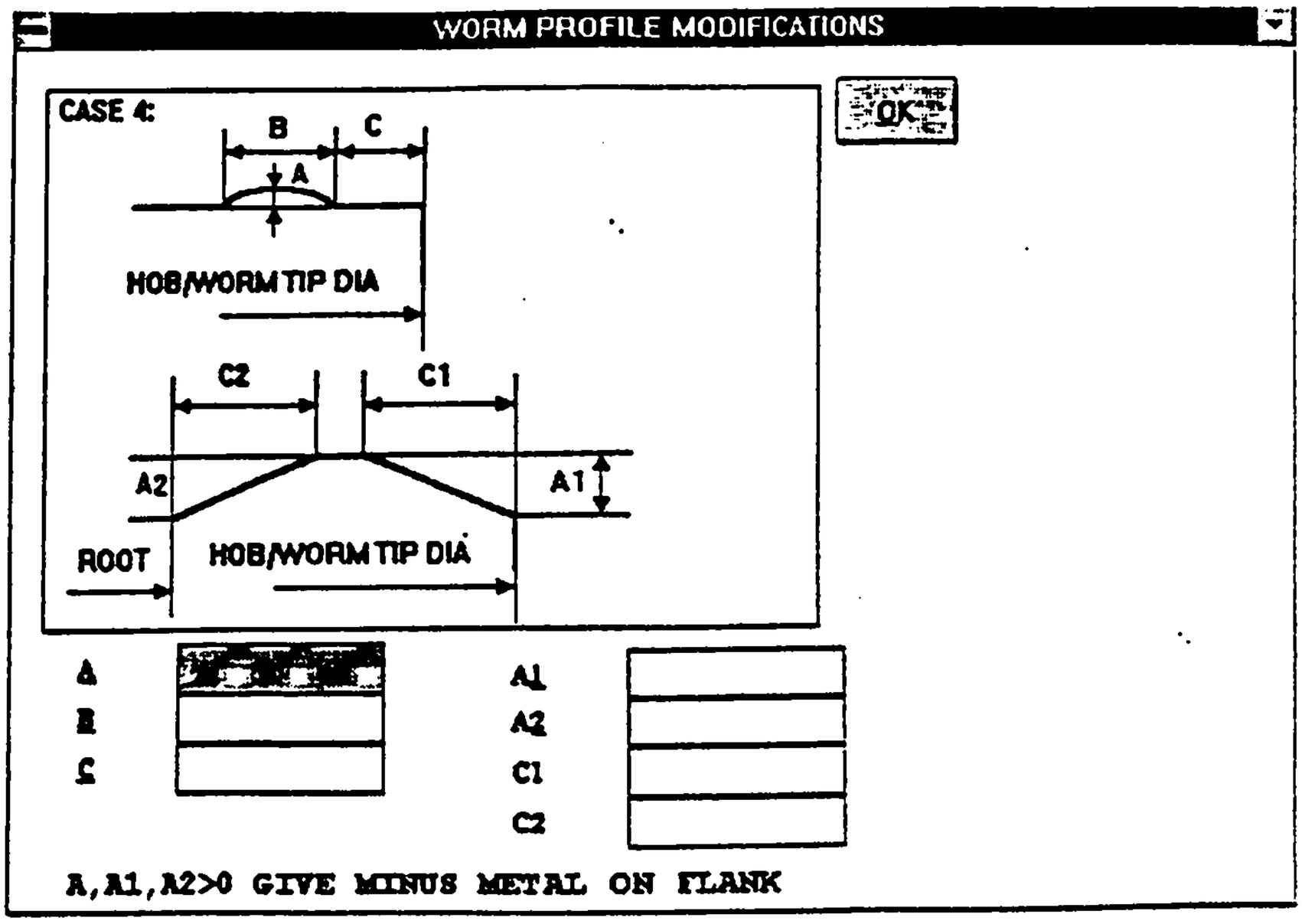


Figure A 2.1 Worm profile modification

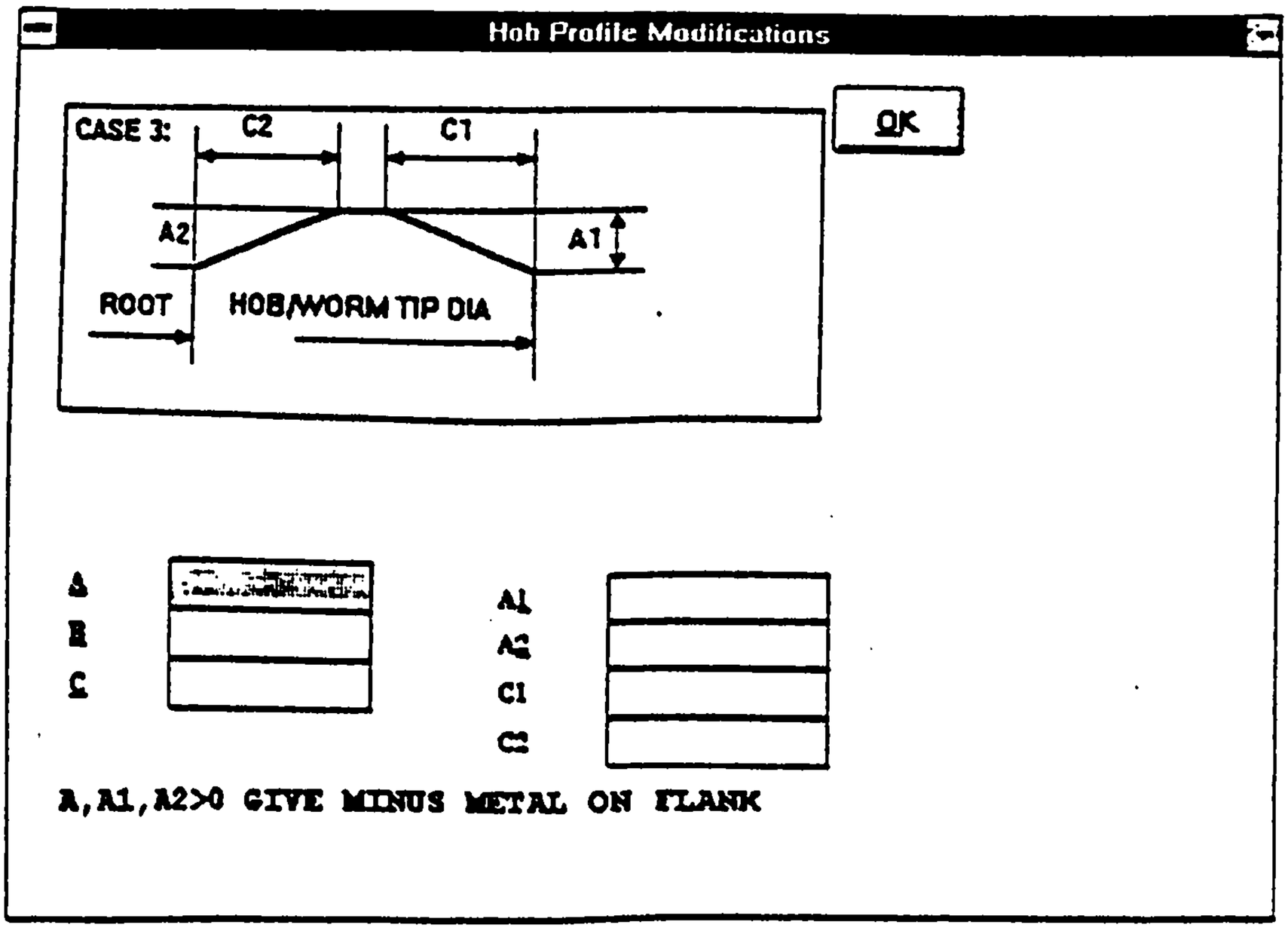


Figure A 2.2 Hob profile modification

APPENDIX 3

NORMAL PITCH METHOD FOR HOB DESIGN

This method of hob design is practically regarded as the industry standard. The principle of this method is to use the same normal pitch and normal pressure angle on both the worm and the hob. The hob data is thus determined by the following procedure:

$$p_{nH} = p_{nw} = p_n$$

$$\alpha_{nH} = \alpha_{nw} = \alpha_n$$

$$z_{1H} = z_{1w} = z_1$$

$$d_H = d_w + O_s$$

$$\sin \gamma_H = p_{nH} \cdot z_{1h} / (d_H \cdot \pi)$$

$$p_{xH} = p_n / \cos \gamma_H$$

where

p_{nH} - hob normal pitch at its mean diameter

p_{nw} - worm normal pitch at its mean diameter

z_{1w} - number of worm starts

z_{1H} - number of hob starts

d_H - hob mean diameter

d_w - worm mean diameter

O_s - hob oversize, determined by the worm gear manufacturer

γ_H - hob lead angle at its mean diameter

p_{xH} - hob axial lead at its mean diameter

The full set of data for the hob thread can then be calculated using BS 721 specifications.

APPENDIX 4

GEOMETRY OF THE "THIN SPUR GEAR" SECTION OF A WORM WHEEL

A 4.1 The "thin involute spur gear" section of a ZI wheel

A ZI worm wheel cut by a non-mismatch hob has a "very thin spur gear" section, which can be measured as a spur gear. To prove the existence of this "very thin spur gear" the following two statements must be proved:

Statement 1: There is an axial section of straight-line profile on the worm wheel hob flank

It is widely known that a convolute worm helicoid (ZA, ZI or ZN) is formed by a straight-line generatrix (see Chapter 1). The generatrix is the intersection between the base tangent plane and the helicoid, as represented by

$$r_f^{(1)} = \begin{bmatrix} -u\cos\lambda \\ -r_b \\ K\cdot(\frac{3}{2}\pi+\phi_1)-u\sin\lambda \\ 1 \end{bmatrix} \quad \ldots\ldots \quad (A4.1)$$

$$\text{where } K = \frac{p_z}{2\cdot\pi}$$

Equation (A 4.1) can be obtained from formula (4.7) by replacing θ with $(3\pi) / 2$ for $x_1^{(1)}$ and $y_1^{(1)}$ and with $(3\pi / 2 + \Phi_1)$ for $z_1^{(1)}$. This is a plane formed by the straight-line generatrix, by translating it at a uniform speed as Φ_1 increases. Since the hob swing angle η is zero (for no-mismatch), the translation of this straight-line generatrix will, like a basic rack, generate an involute profile on the worm wheel flank. For the wheels,

shown in Figure A 1.1, this basic rack lies a distance r_b below the worm wheel mid plane, as shown in Figure A 4.1.

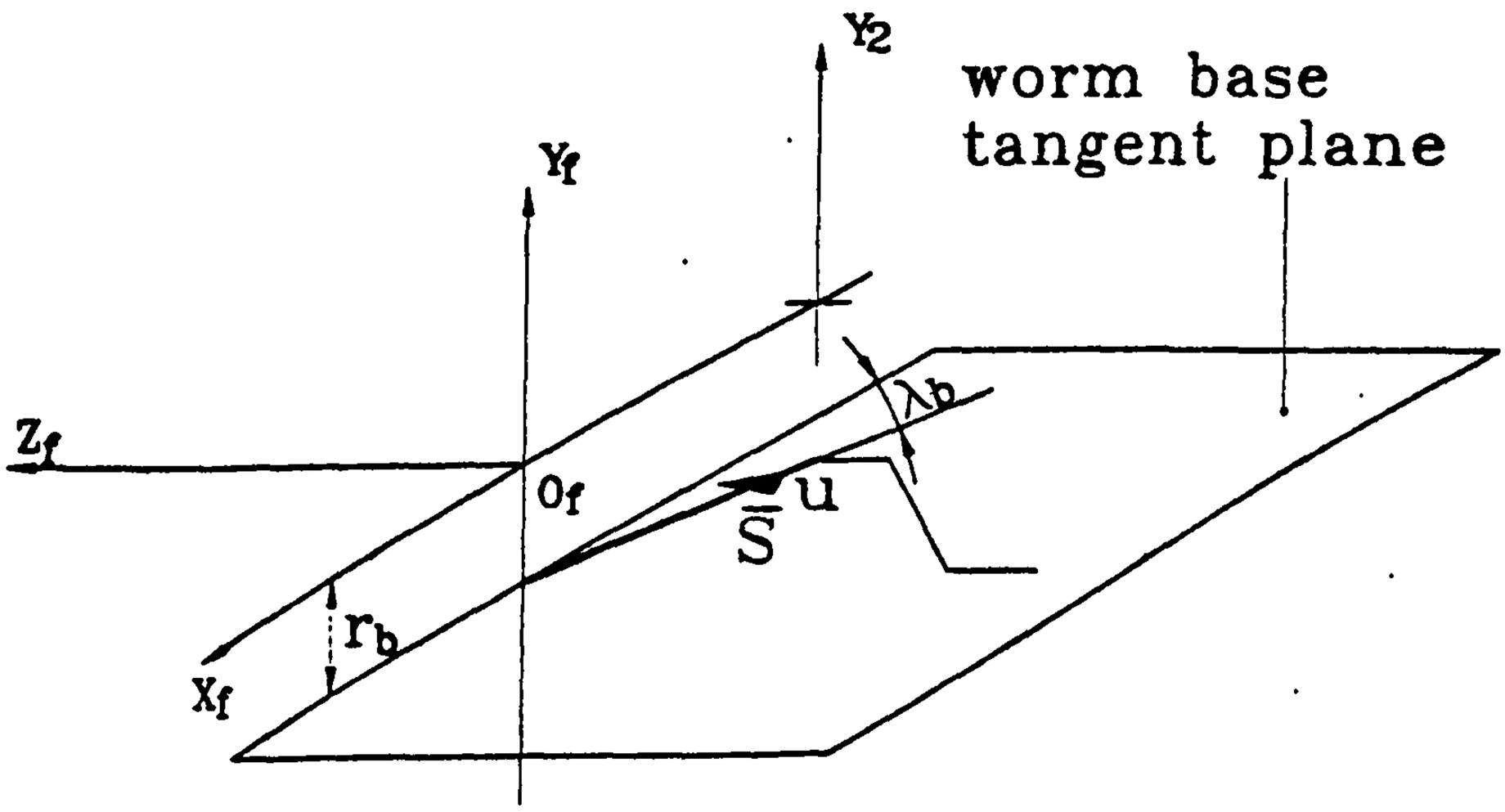


Figure A 4.1

Any type of convolute helicoid worm will generate an involute profile in one transverse plane of the worm wheel, but only the ZI (involute) worm wheel has a "thin spur gear" section. This is proved in statement 2:

Statement 2: The normal to the involute profile section of a ZI worm wheel generated by the non-mismatch worm wheel hob, lies in the wheel transverse plane.

When $\theta = (3\pi)/2$, the normal to the hob helicoid at is, from formula (4.8)

$$N_f^{(1)} = cos\lambda \cdot \begin{bmatrix} -usin\lambda \\ K - r_b tan\lambda \\ ucos\lambda \\ 0 \end{bmatrix} \quad \ldots\ldots \quad (A4.2)$$

$$where\ K = \frac{p_z}{2 \cdot \pi}$$

This has a constant component along y_f (parallel to the worm wheel axis y_2 as shown in

Figure A 4.1) and a variable component u acting in the worm base tangent plane at the generating angle λ (along the line of action normal to the generatrix).

It follows that the direction of this normal will, in general, vary as parameter u varies. When measuring the involute profile on a convolute worm wheel flank, this will cause the contact point between the measuring probe and the wheel tooth flank to move round the probe tip, necessitating a correction proportional to the probe tip radius ρ [Hu and Pennell, 1994].

For involute helicoid worms, however, $\lambda = \gamma_b = \tan^{-1}(p_z / (2\pi r_b))$, so, from formula (A 4.2), the component of the normal parallel to worm wheel axis y_2 vanishes. The normal then lies in the base tangent plane, ie, in the transverse plane of the worm wheel, which is, locally, equivalent to a "thin spur gear".

It thus follows that no correction for stylus contact radius is needed when measuring the involute section of a ZI worm wheel as a "thin spur gear".

A 4.2 Determination of base radius r_{b2} of the "thin spur gear"

The normal of the "thin spur gear" is along the line of action and tangential to the base circle of the "thin spur gear", whose radius r_{b2} is given by

$$r_{b2} = r_2 \cdot \cos\gamma_b = m_x \cdot \frac{z_2}{2} \qquad \text{(A 4.3)}$$

where

r_2 - the wheel reference radius

m_x - worm axial module

γ_b - worm base lead angle